Marketing Research

An Applied Approach

Updated Second European Edition

Naresh K. Malhotra

David F. Birks

Includes CD-Rom

FT Prentice Hall
FINANCIAL TIMES

Additional **student support** at
www.pearsoned.co.uk/malhotra_euro

SNAP and XSight can help you improve your marks!

Your purchase of the updated second European edition of Malhotra and Birks, *Marketing Research: An Applied Approach,* includes a CD-ROM containing valuable SNAP and XSight software demos, to enhance your understanding of quantitative and qualitative aspects of marketing research.

Marketing Research
An Applied Approach

Updated Second European Edition

snap
surveys

xsight

Serial number to access XSight software:
XSTØ1-LZØZØ-EH2HØ-1RIS4-FAULH

Naresh K. Malhotra and David F. Birks

ISBN 0-273-69530-4
© Pearson Education 2006

FT Prentice Hall
FINANCIAL TIMES

DON'T THROW IT AWAY!

What are SNAP and XSight and how will they help you?

SNAP is a user-friendly program for marketing research, providing you with help to design surveys, create questionnaires, prepare data for analysis, collect data and to perform analyses. This powerful survey software is an industry standard that has been helping researchers and educators in more than 50 countries for over 20 years.

SNAP consists of a core product, Snap Professional, and specialist modules that may be added to extend its capabilities to surveys via the Internet, Personal Digital Assistants (PDAs), scanning and telephone interviewing.

You will find direct references to the use of SNAP in Chapter 10 (Survey and quantitative observation techniques), Chapter 13 (Questionnaire design), Chapter 17 (Data preparation) and Chapter 18 (Frequency distribution, cross-tabulation and hypothesis testing).

XSight is new qualitative data analysis software, customised for marketing researchers.

Conventional qualitative data analysis software was designed for and used mainly by academic researchers. XSight was designed for marketing researchers by marketing researchers who understood the particular problems faced by their profession.

Created by QSR International (whose product NVivo is reviewed in Chapter 9, Qualitative research: data analysis), a qualitative research software company with years of experience in developing solutions to a wide array of research problems, XSight is seen as a breakthrough tool for every qualitative marketing researcher. It will enable you to explore unstructured qualitative data gathered via focus groups, interviews or open ended surveys much more easily. This will allow you much more time to devote to the real art of qualitative research – interpretation.

With XSight you will be able to compile, compare and make logical connections in qualitative data almost instantaneously. It can help you identify even the most subtle data patterns.

Using SNAP and XSight

With the enclosed free trial version of SNAP, simply install it on your personal computer at a time that's convenient. You'll then be able to design a 'mini-survey' of up to 9 questions and up to 25 respondents. You'll also be able to see examples of much larger surveys designed with SNAP. **There is no time limit to the use of SNAP.**

With XSight, again simply install it on your personal computer at a time that's convenient and you'll be able to enjoy the full suite of XSight features. **You'll be able to use XSight for a period of 90 days once you've installed it.**

Once you have experienced the benefits of using SNAP and XSight, they will become integral to your work in the design, data collection and analysis of quantitative and qualitative data. Working without them may become inconceivable.

To use the CD:

1. Put the CD in your machine.

2. When prompted, input the password printed on the CD.

3. Follow the instructions on screen.

Visit the *Marketing Research, updated second edition* Companion Website at **www.pearsoned.co.uk/malhotra_euro** to find valuable **student** learning material including:

- Annotated weblinks to relevant, specific Internet resources to facilitate in-depth independent research
- Extra Case Studies
- Full online version of the GlobalCash research project

MARKETING RESEARCH
An Applied Approach

Naresh K. Malhotra

David F. Birks

BRIEF CONTENTS

Dedicated to the memory
of Kevin Fogarty

CONTENTS

Contents

Contents

Supporting resources

Visit **www.pearsoned.co.uk/malhotra_euro** to find valuable online resources

Companion Website for students

- Annotated weblinks to relevant, specific Internet resources to facilitate in-depth independent research
- Extra Case Studies
- Full online version of the GlobalCash research project

For instructors

- Customisable colour PowerPoint slides, including key figures from the main text
- Extensive Solutions Manual, including sample answers for all question material in the book
- Solutions for Extra Case Studies

Also: The Companion Website provides the following features:

- Search tool to help locate specific items of content
- E-mail results and profile tools to send results of quizzes to instructors
- Online help and support to assist with website usage and troubleshooting

For more information please contact your local Pearson Education sales representative
or visit **www.pearsoned.co.uk/malhotra_euro**

PREFACE

Being a marketing researcher is a very creative task. This creativity is nurtured by an environment that makes many demands on the researcher. They must be able to cope with the technical challenges to plan, gather, analyse and interpret information. They must be aware of the challenges faced by the array of decision-makers who trust sound marketing research. They must be able to empathise with the people they aim to question and observe, and to treat them with care and respect.

Trying to develop the technical skills and to balance an appreciation of decision-makers and respondents may seem daunting. Finding the confidence to conduct research and to interpret the findings may seem difficult with so many options to consider. This is where we believe *Marketing Research, An Applied Approach, 2nd European edition* can help. Founded on the enormously successful US editions and the 1st European edition, this text aims to be comprehensive, authoritative and applied. This edition includes an array of European and international examples, practices and illustrations. It portrays a balance of qualitative and quantitative approaches to conducting research that allows the creative support of decision-makers. It will guide the reader through the challenges faced in conducting marketing research of the highest quality. This is achieved through an appropriate blend of scholarship with a highly applied and managerial orientation.

Marketing Research: An Applied Approach

The book is written for use at both the undergraduate and postgraduate levels. The coverage is comprehensive and the depth and breadth of topics are well suited to both levels. The material is presented in a manner that is easy to read and understand. There are numerous diagrams, tables and examples to help explain and illustrate the basic concepts. If you would like to explore a particular topic in more depth, there are numerous references to follow a line of enquiry. The Web addresses presented throughout allow for further illustration of ideas and, in many instances, demonstration versions of software. The Companion Website (**www.pearsoned.co.uk/ malhotra_euro**) presents more European case studies, projects based on the book's running case and annotated Web addresses linking directly to relevant Internet resources. It also now includes a Professional Perspectives section that details more practical applications and issues facing the marketing research industry.

Marketing research is defined as a topic with a clear demonstration of how it may support effective decision-making. The limitations of marketing research are recognised, as are other means of information support that decision-makers turn to. The process of conducting marketing research is described in six stages. Each of these stages is fully described and illustrated. Working out an approach and research design is discussed as a foundation to administering techniques. The justification for the use of qualitative and quantitative techniques is made, with a demonstration of how they can mutually support each other. The administration of a broad range of qualitative and quantitative techniques is presented. Sampling issues for both qualitative and quantitative studies are discussed. Issues of ensuring the quality of data and preparing for data analysis are covered. Qualitative data analysis is discussed in detail. Quantitative data analysis, from simple descriptive statistics to sophisticated multi-

variate techniques, is worked through in seven chapters. Presenting research findings in both oral and written formats is covered, with an emphasis on ensuring that the interpretation of research findings really supports decision-makers. Applications of marketing research in an international context are tackled in most chapters; this vital topic also has a distinct chapter in its own right. The ethical practice of marketing research and the dilemmas that researchers face is also tackled in most chapters. The impact of the Internet and IT on marketing research is integrated throughout the text but also dealt with at the end of each chapter.

Being an excellent marketing researcher requires an appreciation and continual practice of all the elements discussed in this text. This book firmly sets you on that path of creativity and excellence.

New for the updated 2nd European edition

New chapter

There is a new chapter in this 2nd edition update, entitled 'Business-to-business (b2b) marketing research'. Throughout the text there are many examples of the successful use of marketing research set in the context of b2b relationships. As marketing research is obviously used in b2b contexts, we start by questioning whether there is any real distinction between b2b and consumer marketing research. We examine the differences between b2b and consumer marketing and the impact these differences have on the way b2b marketing research is applied. There are very distinctive challenges faced by the b2b marketing researcher in terms of sampling, gaining access to respondents and formulating research designs. We felt that a distinctive chapter was warranted for this large and most challenging sector of the marketing research industry. b2b marketing research also deserves attention in that all manner of organisations, agencies and consultants have found great rewards in the practice of 'business research'. The practices of these organisations, especially those that conduct competitor intelligence, challenge our notions of what marketing research is or should be, in a world where marketing decision-makers continue to seek information support from a burgeoning array of sources.

CD demonstration software

The CD that accompanies this edition contains demonstration software for quantitative and qualitative researchers.

- Snap8 (**www.snapsurveys.com**) is powerful user-friendly software that has driven standards in survey design and analysis in the marketing research industry for over 20 years. It is now used in over 50 countries worldwide. With Snap8, questionnaires can be created and published for surveys that use paper, Web, email, computer assisted personal and telephone interviews. It then offers options to key or scan data or even to collect data automatically. Survey results can be analysed, statistics calculated and tables and charts produced. These features briefly summarise how the software helps to manage a great array of survey designs. By working through the demo, the detail of support in survey design, i.e. the array of functions available, will become apparent. The demo will bring to life the challenges to questionnaire and survey designers and the solutions they use in practice.
- XSight (**www.qsrinternational.com**) is a new product that offers customised data analysis software for qualitative marketing researchers. The software has been developed by QSR International who are world leaders in qualitative research software as used in a variety of applications. Academic marketing researchers that use

qualitative analysis software would be familiar with QSR's NVivo package, which we discuss in Chapter 9. In this chapter, we discuss why for many practitioner qualitative marketing researchers, manual methods or basic computer-based methods have been the norm. XSight has been developed in close consultation with practitioner marketing researchers to bridge the gap between the packages designed for the needs of academics and traditional manual methods. The demo will bring to life the challenges faced by qualitative researchers in the amount and nature of qualitative data they may collect and how they make sense of that data.

New for the 2nd European edition

Revisions to existing chapters

We have retained the most desirable features of the first three US editions and the first European edition, according to our readers and reviewers. In all chapters that tackle quantitative techniques and analysis, none of the well-regarded material has been lost. Changes have been made to improve the clarity of explanations, with many new examples and updated references. Measurement and scaling, once covered in two chapters, is now covered in one chapter. This has been done, without losing any of the substantive explanations, by reorganising and simplifying the approach. Developments in the use of databases have made substantial changes in how decision-makers are supported. These developments are tackled in the major revisions to the chapter that covers internal secondary data and the use of databases. In the chapter devoted to depth interviewing, laddering and repertory grid techniques are described and illustrated.

All chapters have benefited from more up-to-date references and, where appropriate, more relevant examples that reflect the European and international focus. Many new questions that probe issues in more depth have been added to the end of each chapter.

New chapters

There are two new chapters in this edition. The first is entitled 'Qualitative research: its nature and approaches', and the second is 'Qualitative research: data analysis'. This means that this edition has a more balanced approach to qualitative and quantitative marketing research. The text now devotes four chapters to the very important areas of understanding why qualitative research is conducted, group techniques, individual techniques and the manner in which qualitative data is handled and interpreted.

Qualitative and quantitative research should be viewed as complementary. Unfortunately, many researchers and decision-makers do not see this, taking dogmatic positions in favour of either qualitative or quantitative research. The defence of qualitative approaches for a particular marketing research problem, through the positive benefits it bestows and explaining the negative alternatives of a quantitative approach, should be seen as healthy, as should the defence of quantitative approaches. Business and marketing decision-makers use both approaches and will continue to need both. The first of the new chapters should help readers to appreciate these positions and understand why decision-makers use and trust both qualitative and quantitative approaches.

The second new chapter on qualitative data analysis will help researchers cope with the wide array of data that can be generated from qualitative techniques. Working with data 'beyond numbers' means qualitative analysis has less structure and fewer rules, and this chapter helps the researcher to cope with the data generated.

Web links

The pace of new technological developments in the marketing research industry is breathtaking. Trying to keep up with them is a thankless task. It can be guaranteed that as soon as any publication hits the shelves, much material on technology can be out of date. Many of the new developments that have a demonstrable impact on the practice of marketing research, especially through the use of the Internet, have been incorporated into each chapter. In the new Professional Perspectives section, which appears on the Companion Website (**www.pearsoned.co.uk/malhotra_euro**), further illustrations of technological developments are presented. To maintain the currency of technology, the Web addresses of the major research associations, national statistical bodies, and research organisations are presented. The Web addresses of major suppliers of qualitative analysis packages, questionnaire design, survey processing and statistical analysis are also presented. Care has been taken to choose those where good case illustrations and demo versions are available, in a variety of formats to suit individual users.

Photography

The photographs that accompany any management textbook may be perceived as trivial and mere adornments. I do not take this view. I see textbook photographs as a means to encapsulate the essence of a subject and its challenges. In putting together the photographs for this edition, I wanted images that portray the creativity and the fun that can be experienced in marketing research. Some of the images in this edition have a serious intent, linking a particular image to concepts and aiding the recall of a concept. Others are there to lighten the sometimes difficult technical passages. In their own right, each image has many technical and artistic merits.

Companion Website

The Companion Website, at **www.pearsoned.co.uk/malhotra_euro**, has been updated to reflect the changes in this edition. There are new European case studies with discussion points and questions to tackle. All the referenced Websites in the text are described with notes of key features to look for on a particular site.

For the lecturer

Instructor's Manual. The Instructor's Manual is very closely tied to the text, but is not prescriptive in how the material should be handled in the classroom. The manual offers teaching suggestions, answers to all end-of-chapter questions, Professional Perspective discussion points and case study exercises.

Solutions. Very detailed suggested solutions to the end-of-chapter questions, to the Professional Perspective questions and to the Web-based case studies. These solutions assist lecturers in the development of examination papers and marking schemes.

Professional Perspectives. These 20 articles can be used as mini case studies and discussion papers. Each Professional Perspective ends with two questions to initiate debate.

PowerPoint slides. The PowerPoint slides are not just replicas of the figures and tables from the text. Care has been taken to ensure that they are not packed with too much information which would render them unintelligible in a lecture theatre.

For the student

Professional Perspectives. There is a new and distinct section entitled Professional Perspectives which is presented on the Companion Website. This section is made up of 20 articles written by leading practitioners, users and writers of marketing research. They show an array of applications and how different techniques work together. They reflect the realities of marketing decision-making and the nature of research support required. As issues emerge in individual chapters, links are made to specific articles in this section. However, these articles have been chosen for the variety of interconnected issues they tackle; they are written to illustrate more than one point. As well as developing issues in individual chapters, they can be used as case studies and as discussion papers in their own right. Points to discuss are presented at the end of each article.

Weblinks. This section offers many further sources of support for any essay or project work. Of particular note are the links to demo versions of a range of industry standard marketing research software packages.

Case studies. These are European in focus or relevant globally. Being longer than the examples set in the text, these set the challenges of applying marketing research in a broader context.

GlobalCash project. This section contains many of the key documents used to manage the project from the initial brief and proposal through to examples of the questionnaire and the data collected in an SPSS format. The questionnaire is presented in English, French, German and Spanish versions, enabling a vision of the challenges faced in translating questionnaires. The GlobalCash documentation will help with project work by showing how key documents help to manage the marketing research process in a professional manner.

Acknowledgements

Wendy Gordon, Marsha Hemmingway, Arno Hummerston, Tim Macer, Jeff Miller, Virginia Monk, Trevor Merriden, Lex van Meurs, Marcel van der Kooi and Ron Whelan.

I would also like to thank other authors who gave me permission to use their ideas: Mary Goodyear, Peter Cooper and Paul Bate.

The reviewers of the 1st edition provided many constructive and valuable suggestions, and their help is gratefully acknowledged.

Reviewers for the first edition:

Howard Jackson, University of Huddersfield
Paul Baines, Middlesex University
Martin Wetzels, University of Maastricht
David Longbottom, University of Derby
John Beaumont-Kerridge, University of Luton
Marc De Laet, Hogeschool, Antwerp
Kare Sandvik, Buskerud College, Norway.

The reviewers of this edition have also been of enormous help. Their advice was most thorough, balanced and very well founded. I have done my best to incorporate the many valuable suggestions they offered. A big thank you for your help.

Reviewers for the second edition:

Martin Wetzels, Technical University of Eindhoven
Heather Skinner, University of Glamorgan
Eddie Rohan, Dublin Institute of Technology.

In putting together the images for this edition, I enjoyed enormously the task of working with the photographer Michael Hewitt. Together we worked on the concepts of the images, Michael created the images and I positioned them in the text and created the captions (sorry for any toe-curling jokes). I worked with Michael on two 'shoots', and witnessed the great amount of preparation, vision and technical skills involved; I was very impressed. So a big thank you to Michael and his partner Angela Williams for all their hard work and end results that I think are amazing.

To Abigail Woodman at Pearson Education, I don't know how to thank you enough. Abigail managed the whole process of developing this new edition superbly. She has cajoled and supported me at just the right level. Her advice has been first class throughout.

For this updated 2nd edition I have been very fortunate to work with David Cox at Pearson Education. Through many personal and professional changes, David has been very patient, kind and helpful. His advice has always been first class, and given with great humour and very incisive guidance. I could not have asked for better support. I would also like to thank Thomas Sigel for championing the development of this text and for his continual support in so many ways.

At Pearson Education the book has also come together with the help of Karen Mclaren, Senior Desk Editor; Michelle Morgan, who designed the cover; Kelly Meyer, who helped with the internal design; and Peter Hooper, Editorial Assistant.

Last but by no means least, to be able to find the time and space to write, the love, support and understanding of your family is vital. So thank you so much, mum, for all that you have given to me, and to my nephew James, sorry for confusing you in your school and college projects that involved questionnaire design. To my partner Helen, huge amounts of thanks and love for all you put up with in giving me the space and peace to write and even bigger thanks for bringing our beautiful son Jesse into the world.

David F. Birks

An **Overview** summarises the topics discussed in the chapter

Stage flowcharts show which of the six stages of the marketing research process the present chapter fits into

Objectives at the beginning of each chapter detail what students should be able to do after reading the chapter

---

CHAPTER 12

Measurement and scaling: fundamentals, comparative and non-comparative scaling

Objectives

After reading this chapter, you should be able to:

1 introduce the concepts of measurement and scaling and show how scaling may be considered an extension of measurement;

2 discuss the primary scales of measurement and differentiate nominal, ordinal, interval and ratio scales;

3 classify and discuss scaling techniques as comparative and non-comparative and describe the comparative techniques of paired comparison, rank order, constant sum and Q-sort scaling;

4 explain the concept of verbal protocols and discuss how they could be employed to measure consumer response to advertising;

5 describe the non-comparative scaling techniques, distinguish between continuous and itemised rating scales, and explain Likert, semantic differential and Stapel scales;

6 discuss the decisions involved in constructing itemised rating scales;

7 discuss the criteria used for scale evaluation and explain how to assess reliability, validity and generalisability;

8 discuss the considerations involved in implementing the primary scales of measurement in an international setting;

9 understand the ethical issues involved in selecting scales of measurement.

'When you can measure what you are speaking about and express it in numbers, you know something about it'. – *Lord Kelvin*

Stage 1 Problem definition
Stage 2 Research approach developed
Stage 3 Research design developed
Stage 4 Fieldwork or data collection
Stage 5 Data preparation and analysis
Stage 6 Report preparation and presentation

Overview

Once the marketing researcher has a clear understanding of what they wish to understand in their target respondents, they should consider the concepts of scaling and measurement. These concepts are vital in developing questionnaires or 'instruments of measurement' that will fulfil their research objectives in the most accurate manner. This chapter describes the concepts of scaling and measurement and discusses four primary scales of measurement: nominal, ordinal, interval and ratio. We describe and illustrate both comparative and non-comparative scaling techniques in detail. The comparative techniques, consisting of paired comparison, rank order, constant sum and Q-sort scaling, are discussed and illustrated with examples. The non-comparative techniques are composed of continuous and itemised rating scales. We discuss and illustrate the popular itemised rating scales – the Likert, semantic differential and Stapel scales – as well as the construction of multi-item rating scales. We show how scaling techniques should be evaluated in terms of reliability and validity and consider how the researcher selects a particular scaling technique. Mathematically derived scales are also presented. The considerations involved in implementing scaling techniques when researching international markets are discussed. The chapter concludes with a discussion of several ethical issues that arise in scale construction. We begin with an example of how the use of different types of scale can give quite different powers of analysis and interpretation.

Will your instrument of measurement do any better in revealing consumer preferences for soups?

291

Photographs encapsulate the essence of the subject in a creative way

---

International marketing research shows how the topics discussed fit into an international context

Chapter 3 • Research design

Figure 3.2 Understanding respondents – to help choose optimum research techniques

Context or environment where respondent is questioned

Nature of issue under investigation, as perceived by respondent

Nature of respondent

edge bars'. This context could have helped the target respondents to relax, to develop a better rapport with interviewers and other respondents, and to think more about the issues and express their feelings more clearly. If the interviews were conducted over the Internet, the same levels of relaxation and rapport may not work. If the interviews were targeted at older respondents, they may have felt very self-conscious in 'leading-edge bars', which may restrict their responses. Researchers therefore must understand characteristics of respondents, how they react to particular issues and how they react in different contexts or environments. These factors are illustrated in Figure 3.2, which acts as a reminder of the understanding of respondents that researchers must develop, in order to choose and apply the best research technique.

Marginal definitions provide a summary of key terms highlighted in the text and are repeated in the Glossary at the back of the book

Exploratory research A research design characterised by a flexible and evolving approach to understand marketing phenomena that are inherently difficult to measure.

Conclusive research A research design characterised by the measurement of clearly defined marketing phenomena.

Research design classification

Research designs may be broadly classified as exploratory or conclusive (see Figure 3.3). The differences between exploratory research and conclusive research are summarised in Table 3.1.

Flow charts order the discussion which follows to allow for quick reference

Figure 3.3 A classification of marketing research designs

Research design
- Exploratory design
  - Qualitative exploration
  - Quantitative exploration
- Conclusive design
  - Descriptive research
    - Cross-sectional design
      - Single cross-sectional
      - Multiple cross-sectional
    - Longitudinal design
  - Causal research

62

International marketing research

The selection of appropriate interviewing techniques is much more difficult in foreign countries because of the challenges of conducting research there. Given the differences in the economic, structural, informational, technological and socio-cultural environment, the feasibility and popularity of the different interviewing techniques vary widely. In the United States and Canada, for example, the telephone has achieved almost total penetration of households. Consequently, telephone interviewing is a dominant mode of questionnaire administration. The same situation exists in some European countries, such as Sweden. In many other European countries, however, the telephone interview gets confused with telephone sales. This results in high refusal rates and scepticism of what the purpose of a survey is. In developing countries, the problem with the telephone is the low number of households that have telephones.

Because of the low cost, mail interviews continue to be used in most developed countries where literacy is high and the postal system is well developed, for example in Britain, Canada, Denmark, Finland, Iceland, the Netherlands, Norway, Sweden and the United States. In many parts of Africa, Asia and South America, however, the use of mail surveys and mail panels is low because of illiteracy and the large proportion of the population living in rural areas.

The following example illustrates how CAPI technology has developed to allow consistent approaches to survey techniques across Europe. It shows that, by developing an appreciation of the cultural differences between countries, a research design can be built that allows accurate and comparable surveys to be conducted.

example **CAPIBUS Europe** [40]

The concept of the International Omnibus is not new; many research groups offer an international service. In reality, however, these have been little more than brokering services, bolting together whatever omnibus is available in each country, with little real standardisation.

Research Services Limited and the IPSOS group have tackled this problem by calling on the benefits of CAPI technology to introduce CAPIBUS Europe, a weekly omnibus survey covering the six major markets of Europe (Britain, France, Germany, Italy, Netherlands and Spain).

The use of computer technology means that questionnaires can be scripted in one location and transmitted electronically to other countries. While the need to ensure accurate translation remains, the problems involved in having different questionnaire formats, classification systems and data maps are minimised. At the end of a project, data are again transmitted electronically, to be aggregated by the lead agency in a standardised format for all markets. This can then be weighted to provide information on the European market as well as for each local market. ■

Selection of survey techniques

No questionnaire administration method is superior in all situations. Table 10.4 presents a comparative evaluation of the major modes of collecting quantitative data in the context of international marketing research. In this table, the survey techniques are discussed only under the broad headings of telephone, personal, mail and electronic interviews. The use of CATI, CAPI, electronic survey and mail panels depends heavily on the state of technological development in the country. Likewise, the use of street interviewing is contingent upon the dominance of shopping centres in the retailing environment.

249

Examples pinpoint specific companies to illustrate the points made

A **Summary** at the end of each chapter lists the main points covered

Ethics in marketing research focus on the ethical practice of marketing research and the dilemmas that researchers face

Internet and computer applications integrated throughout the text as well as at the end of each chapter tackle the impact of the Internet and IT on marketing research

---

Ethics in marketing research

Possible ethical dilemmas exist when using internal or external secondary data. Some ethical issues that are pertinent include:

- The unnecessary collection of primary data when the problem can be addressed based on secondary data alone
- The use of only secondary data when primary data are needed
- The use of secondary data that are not applicable
- The use of secondary data that have been gathered through morally questionable means
- Compromising the anonymity of customer details held on databases.

As was discussed in Chapter 2, the unnecessary collection of expensive primary data when the research problem can be addressed using only secondary data is unethical. In this case, the researcher is using a more expensive method that is less appropriate. Similarly, the exclusive reliance on secondary data when the research problem requires primary data collection could raise ethical concerns. This is particularly true if the researcher is charging a fixed fee for the project and the research design was not specified in advance. Here again, the researcher's profit goes up, but at the expense of the client.

The researcher is ethically obliged to ensure the relevance and usefulness of secondary data to the problem at hand. The secondary data should be evaluated by the criteria discussed earlier in this chapter. Only data judged to be appropriate should be used.

Internet and computer applications

The World Wide Web as an online source of secondary data

The World Wide Web is a vital source of secondary data and intelligence for the marketing researcher. The speed of the Internet can aid rapid problem diagnosis and data collection at various stages of the research process. Given the global nature of the technology, the Internet is a vital tool for the international marketing researcher. Searching the Web is facilitated by using generalist search engines such as Google, Yahoo! or AltaVista, which require a few keywords to get hundreds of sites related to one subject. One can go directly to the Websites of traditional suppliers of secondary data from government or business sources. Many of those sites also have inside search engines that sort data from the supplier's internal database. Information on the Web is of great value as generally it is current, though care must be taken to note when Web pages have been updated. It should be noted that not all secondary data on the Web is free. The Web may reveal the existence of data on a particular subject or industry, but remember the costs involved in conducting quality research. Hence the Web may be used to give an awareness and a 'taste' of secondary data but it does not necessarily mean 'free' data. See the Companion Website for Professional Perspective 9 by Trevor Fenwick. Trevor's article 'How online retrieval is devaluing research' is a reminder of the quality issues of secondary data gathered from the Internet.

►►► *See Professional Perspective 9.*

Summary

A questionnaire has three objectives. It must translate the information needed into a set of specific questions the respondents can and will answer. It must motivate respondents to complete the interview. It must also minimise response error.

Designing a questionnaire is more of a craft than a science. This is primarily caused by the interrelationship of stages and the trade-offs that questionnaire designers make in balancing the source of ideas, question purposes, actual questions and question analyses. The steps involved in the questionnaire design process involve:

1 Specifying the information needed. Understanding what information decision-makers need.
2 Specifying the type of interviewing method. Understanding which means of eliciting the information will work best, given the research design constraints that the researcher has to work with.
3 Determining the content of individual questions. Understanding the purpose of each question and working out how a posed question may fulfil that purpose.
4 Overcoming the respondents' inability and unwillingness to answer questions. Understanding the process of approaching and questioning respondents – from their perspective. Knowing what benefits they get from taking part in the survey process.
5 Choosing the question structure. Understanding how individual questions help to elicit information from respondents and help them to express their feelings.
6 Choosing the question wording. Understanding the meaning of words from the perspective of the respondent.
7 Arranging the questions in a proper order. Understanding what 'proper' means from the perspective of the respondent. Recognising that, as each question is posed to a respondent and they think about their response, the respondent changes. Information is not only drawn out of respondents, it is communicated to them as each question is tackled.
8 Identifying the form and layout of the questionnaire. Understanding how in a self-completion scenario the form and layout motivate and help the respondent to answer the questions properly and honestly. Understanding how the form and layout help the interviewer to conduct and record the interview.
9 Reproducing the questionnaire. Understanding how the professional appearance of a questionnaire affects the perceived credibility and professional ability of researchers.
10 Eliminating problems by pilot-testing. Understanding that no matter how much experience the researcher has in designing questionnaires – the issues, respondent characteristics and context of questioning make each survey unique – pilot-testing is vital.

Questions

1 What is the purpose of the questionnaire?
2 What expectations does the marketing researcher have of potential questionnaire respondents – in terms of how they will react to the experience of completing a questionnaire?
3 What does the marketing researcher have to offer potential questionnaire respondents? Why should this question be considered?
4 How would you determine whether a specific question should be included in a questionnaire?

As issues emerge in individual chapters, **links** are made to specific articles within the Professional Perspectives section on the Companion Website

Questions test the student's understanding of the chapter's content

---

5 What are the reasons why respondents may be (a) unable to answer and (b) unwilling to answer the question asked?
6 Explain the errors of omission, telescoping and creation. What can be done to reduce such errors?
7 Explain the concepts of aided and unaided recall.
8 What can a researcher do to make the request for information seem legitimate?
9 What are the advantages and disadvantages of unstructured questions?
10 What are the issues involved in designing multiple-choice questions?
11 What are the guidelines available for deciding on question wording?
12 What is a leading question? Give an example.
13 What is the proper order for questions intended to obtain basic, classification and identification information?
14 What guidelines are available for deciding on the form and layout of a questionnaire?
15 Describe the issues involved in pilot-testing a questionnaire.

Notes

1 Dowding, P., Research, Fieldwork Supplement, July 2000, 4.
2 Abel, S., 'Finding the elusive young through co-operative parents', *ResearchPlus* (February 1996).
3 The founding reference to this subject is Payne, S.L., *The Art of Asking Questions* (Princeton, NJ: Princeton University Press, 1951).
4 These guidelines are drawn from several books on questionnaire design: Schuman, H. and Presser, S., *Questions and Answers in Attitude Surveys* (Thousand Oaks, CA: Sage, 1996); Fink, A., *How to Ask Survey Questions* (Thousand Oaks, CA: Sage, 1995); Sudman, S. and Bradburn, N.M., *Asking Questions* (San Francisco, CA: Jossey-Bass, 1983); Erdos, P.L., *Professional Mail Surveys* (Malabar, FL: Robert E. Krieger, 1983); Fowler Jr., F.J., Backstrom, C.H. and Hursh-Csar, G., *Survey Research* (Cambridge, MA: Wiley, 1981); Labau, P., *Advanced Questionnaire Design* (Orlando, FL: Abt Books, 1981); Dillman, D., *Mail and Telephone Surveys: The Total Design Method* (New York: Wiley, 1978); Blankenship, A.B., *Professional Telephone Surveys* (New York: McGraw-Hill, 1977), 94–5; Koehauser, A. and Sheatsley, P.B., 'Questionnaire construction and interview procedure', in Selltiz, C., Wrightsman, L.S. and Cook, S.W., (eds) *Research Methods in Social Relations*, 3rd edn (New York: Holt, Rinehart & Winston, 1976), 541–73.
5 Bourque, L.B. and Fielder, E.P., *How to Conduct Self-Administered and Mail Surveys* (Thousand Oaks, CA: Sage, 1995); Frey, J.H. and Oishi, S.M., *How to Conduct Interviews by Telephone and in Person* (Thousand Oaks, CA: Sage, 1995).
6 Semon, T.T., 'Asking "how important" is not enough', *Marketing News* 31(16) (4 August 1997), 19; Hague, P., 'Good and bad in questionnaire design', *Industrial Marketing Digest* 12, Third Quarter 1987, 161–70.
7 Boyd Jr, H.W., Westfall, R. and Stasch, S.E., *Marketing Research: Text and Cases*, 7th edn (Homewood, IL: Irwin, 1989), 277.
8 Stapel, J., 'Observations: a brief observation about likability and interestingness of advertising', *Journal of Advertising Research* 34(2) (March/April 1994), 79–80; Bishop, G.E., Oldendick, R.W. and Tuchfarber, A.J., 'Effects of filter questions in public opinion surveys', *Public Opinion Quarterly* 46 (Spring 1982), 66–85.
9 Schneider, K.C. and Johnson, J.C., 'Link between response-inducing strategies and uninformed response', *Marketing Intelligence and Planning* 12(1) (1994), 29–36.
10 Dutka, S. and Frankel, L.R., 'Measuring response error', *Journal of Advertising Research* 37(1) (January/February 1997), 33–9; Haller, T., *Danger: Marketing Researcher at Work* (Westport, CT: Quorum Books, 1983), 149.
11 Menon, G., Raghubir, P. and Schwarz, N., 'Behavioural frequency judgments: an accessibility-diagnosticity framework', *Journal of Consumer Research* 22(2) (September 1995), 212–28; Cook, W.A., 'Telescoping and memory's other tricks', *Journal of Advertising Research* (February–March 1987), 5–8; Sudman, S., Finn, A. and Lannom, L., 'The use of bounded recall procedures in single interviews', *Public Opinion Quarterly* (Summer 1984), 520–4.
12 Hill, R.P., 'Researching sensitive topics in marketing – the special case of vulnerable populations', *Journal of Public Policy and Marketing* 1(1) (Spring 1995), 143–8.
13 Tourangeau, R. and Smith, T.W., 'Asking sensitive questions: the impact of data collection mode, question format, and question context', *Public Opinion Quarterly* 60(20) (Summer 1996), 275–304; Marquis, K.H., Marquis, M.S. and Polich, M.J., 'Response bias and reliability in sensitive topic surveys', *Journal of the American Statistical Association* (June 1986), 381–9.
14 Peterson, R.A., 'Asking the age question: a research note', *Public Opinion Quarterly* (Spring 1984), 379–83; Sheth, J.N., LeClaire Jr, A. and Wachspress, D., 'Impact of asking race information in mail surveys', *Journal of Marketing* (Winter 1980), 67–70.

Notes are sources for further reading

Dr Naresh K. Malhotra is Regents' Professor, DuPree College of Management, Georgia Institute of Technology. He is listed in *Marquis Who's Who in America*, 51st Edition (1997), 52nd Edition (1998), 53rd Edition (1999), and in the *National Registry of Who's Who* (1999).

In an article by Wheatley and Wilson (1987 AMA Educators' Proceedings), Professor Malhotra was ranked number one in the country based on articles published in the *Journal of Marketing Research* from 1980 to 1985. He also holds the all-time record for the most publications in the *Journal of Health Care Marketing*. He is ranked number one based on publications in the *Journal of the Academy of Marketing Science* (JAMS) from its inception through volume 23, 1995. He is also number one based on publications in JAMS from 1986 to 1995. He is listed as one of the best researchers in marketing in John Fraedrich, 'The best researchers in marketing', *Marketing Educator* (Summer 1997), p. 5.

He has published more than 75 papers in major refereed journals including the *Journal of Marketing Research*, *Journal of Consumer Research*, *Marketing Science*, *Journal of Marketing*, *Journal of Academy of Marketing Science*, *Journal of Retailing*, *Journal of Health Care Marketing*, and leading journals in statistics, management science and psychology. In addition, he has also published numerous refereed articles in the proceedings of major national and international conferences. Several articles have received research awards.

He was Chairman, Academy of Marketing Science Foundation, 1996–1998, and was President, Academy of Marketing Science, 1994–1996, and Chairman of the Board of Governors from 1990 to 1992. He is a Distinguished Fellow of the Academy and Fellow of the Decision Sciences Institute. He serves as an Associate Editor of *Decision Sciences Journal* and has served as Section Editor, Health Care Marketing Abstracts, *Journal of Health Care Marketing*. Also, he serves on the Editorial Boards of eight journals.

His book entitled *Marketing Research: An Applied Orientation*, Third Edition, was published by Prentice Hall, Inc. An International Edition and an Australian Edition of his book have also been published, along with a Spanish translation. The book has received widespread adoption at both the graduate and undergraduate levels with more than 100 schools using it in the United States.

Dr Malhotra has consulted for business, non-profit and government organisations in the United States and abroad and has served as an expert witness in legal and regulatory proceedings. He is the winner of numerous awards and honours for research, teaching and service to the profession.

Dr Malhotra is a member and Deacon, First Baptist Church of Atlanta. He lives in the Atlanta area with his wife, Veena, and children, Ruth and Paul.

Dr David Frederick Birks is a Senior Lecturer in Marketing at the Institute for Entrepreneurship in the School of Management, University of Southampton, England. He has taught marketing research and management research on undergraduate, taught postgraduate and research degree programmes. David's industrial experience was gained in the construction and housing industry in England and Germany. Before university lecturing he worked in purchasing, planning, marketing and research. He has Masters degrees in Marketing Management and in Social Statistics and a Ph.D. in Marketing Information Systems.

David has continued to practise marketing research throughout his university career, managing projects in financial institutions, retailers, industrial organisations, local authorities and charities. He has managed projects as part of the undergraduate, M.Sc., M.B.A. and Ph.D. degree programmes at the University of Bath. He was the Head of Marketing Research on the GlobalCash Project detailed throughout this text. He helped to develop the research design, research instruments and analyses of this project. This involved coordinating the demands of major pan-European and American banks, and the research requirements of 19 business schools throughout Europe. He edited and made a major contribution to the text *Global Cash Management in Europe* that resulted from the combined research efforts of colleagues in Britain, Denmark, Germany, Ireland, Italy, Norway and Sweden. He has published further research from GlobalCash with colleagues from Aarhus, Denmark. In addition to his Cash Management work, David's publications have covered the fields of Housing, Statistics and Marketing.

David lives in the beautiful village of Vobster in the Mendip Hills of Somerset. From this very English setting, he brings a cosmopolitan background to this European edition, having an English father, German mother, French grandfather and Belgian great-grandparents.
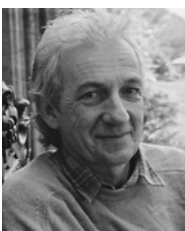
Michael George Hewett graduated in Fine Art at the University of Newcastle upon Tyne. He is the Curator of the Angela Williams Archive of Norman Parkinson Vintage Photography. Through this role, Michael has developed a great interest in fashion photography in the 1950s and 60s, particularly the work of Norman Parkinson at *Vogue*. The history of advertising and the role of photography in post-war advertising are also major research interests for Michael. He is currently studying the formative years of the author Mary Shelley through original documentation.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Introduction to marketing research

Objectives

After reading this chapter, you should be able to:

1 understand the nature and scope of marketing research and its role in designing and implementing successful marketing programmes;

2 explain the role of marketing research in marketing information systems and decision support systems;

3 discuss the types and roles of research suppliers, including internal and external, full-service and limited-service suppliers;

4 describe a conceptual framework for conducting marketing research as well as the steps of the marketing research process;

5 understand why some marketers may be sceptical of the value of marketing research;

6 acquire an appreciation of the complexities involved in international marketing research;

7 gain an understanding of the ethical aspects of marketing research and the responsibilities that marketing research stakeholders have to themselves, each other, and to the research project;

8 appreciate the potential opportunities and threats of the Internet to marketing researchers.

Marketing researchers support decision-makers by collecting, analysing and interpreting information needed to identify and solve marketing problems.

Overview

Marketing research comprises one of the most important and fascinating facets of marketing. In this chapter, we describe the nature and scope of marketing research, emphasising its role of supporting marketing decision-making, and provide several real-life examples to illustrate the basic concepts of marketing research. We give a formal definition of marketing research and subdivide marketing research into two areas: problem identification and problem-solving research. We show that marketing research may be conducted on an *ad hoc* basis but is also an integral part of marketing information systems or decision support systems. Next, we provide an overview of marketing research suppliers and services. We go through a simple six-stage linear description of the marketing research process. This description is extended to illustrate many of the interconnected activities in the marketing research process. There are many successful marketing decisions that have used marketing research support; however, marketing research does not replace decision-making. The limitations of marketing research are established. To illustrate the marketing research process, we examine the GlobalCash study, a live pan-European marketing research project that is conducted biennially. GlobalCash will be used as a running example throughout this book. The topic of international marketing research is introduced. International marketing research will be discussed systematically in the subsequent chapters and will be tackled in a dedicated chapter. The ethical aspects of marketing research and the responsibilities that marketing research stakeholders have to themselves, to each other, and to the research project are presented and developed in more detail throughout the text.

What does marketing research encompass?

The term 'marketing research' is broad in meaning; it is related to all aspects of marketing decision-making. The following examples provide a flavour of the varied nature and applications of marketing research.

e x a m p l e

Listening to the customer – research at Royal Ahold[1]

Royal Ahold is a world supermarket leader. In the Netherlands Ahold operates six chains with over 1,750 outlets, including the flagship Albert Heijn supermarkets. Worldwide, Ahold serves 20 million customers weekly in 3,400 stores in 17 countries across the US, Central and Western Europe, Latin America and Asia. Customer orientation is at the top of the fundamental principles of the company. Its credo is that the customer comes first: 'However big we become, however international, it is ultimately the customer who determines our success'.

The following summarises the ways that Royal Ahold 'listens' to its customers to maintain its customer orientation and continued success:

■ Produce economic analyses and forecasts. They do this by gathering secondary data and intelligence that give them an understanding of retail developments, competitive threats and market changes.
■ Use and contribute to audit data from A.C. Nielsen to obtain global data about developments in their markets.
■ See their stores as a major 'market research laboratory' to study customers. They know when they come in, how often and what they buy. Fundamental to these observations is the use of scanner systems and loyalty cards.
■ Use focus groups as a major source of information about how customers and non-customers feel. They insist that management teams watch and listen to these discussions.

- Use observation approaches to watch how customers behave in store, using protocol tapes (where people think aloud about the purchases they are making).
- Select some researchers to supply raw data, where Ahold perform their own analyses and interpretations, whilst in other projects they use specialised strategic input from researchers to provide added value.
- Share their best practices and know-how. They have an electronic market research platform where researchers discuss their projects and any problems they have. For example, if a successful approach has been used in the Netherlands, they can consider using it in the US.
- Circulate a digest to their Corporate Executive Board and management teams all over the world. *Market Research Findings* is issued twice a year, giving details of important papers, articles and reports on retail research from inside or outside the Ahold company.

Bert L. J. van der Herberg, Vice-President Market Research, summarises the role of market researchers at Royal Ahold:

'Researchers are backroom consultants. I see the market researcher as a philosopher who can take a critical view of the internal and external world. The researcher can act as "the serious fool" to the court. The Board takes on our ideas and our language, but the market researcher is not the spokesman for the company. They are listeners and interpreters.' ∎

e x a m p l e

Banner advertising – more than clicks [2]

The Internet has the potential to fulfil advertising objectives beyond a direct response to a banner advertisement. It has the potential to generate awareness and to build brands. For media planners, it is very important to gain as much knowledge as possible of the effect of Internet advertising in a media plan. In order to understand these effects, OMD Denmark and A.C. Nielsen online initiated an Internet survey entitled 'Banner Advertising Survey 2000'. The survey was carried out on 16 major Danish Websites. The importance of the Websites involved can be emphasised by the fact that 80% of Danes with Internet access from their home visit at least one of the sites on a monthly basis. Fourteen banners for a total of 14 different brands were tested. In total, 23,823 respondents participated in the survey. The approach to the study was based on a classic causal research design involving a test group and a control group. The groups were differentiated only by one known factor – one banner impression – with the purpose of examining whether an effect could be measured. The results of the survey showed that as well as generating direct responses, banner adverts build brand awareness. The vast majority of respondents saw the banners without clicking through to the Website, giving the banner great potential as a broad tool of communication. ∎

e x a m p l e

From Armani to AIDS: the hopes, the fears, of Europe's youth [3]

Today's youth live in a fast-changing world, and their attitudes, behaviour and exposure to the media and to brands differ from those of their predecessors of just a few years ago. A major qualitative research study, interviewing more than 500 young people across 16 countries, has been conducted. Four workshops of seven to nine participants were held in each capital city, each lasting up to five hours. The countries covered were Austria, the Czech Republic, Denmark, Finland, France, Germany, Hungary, Italy, the Netherlands, Norway, Poland, Russia, Spain, Sweden, Switzerland and the UK. The target of the research was 'trend-setters', that is young opinion leaders and leading-edge youth aged 15–19 years, those who decide what is 'in', the early adopters and disseminators of opinions and tastes. Participants were given disposable cameras to take photographs ahead of the groups; these were used, with other materials, to create collages illustrating what makes them 'happy' or 'sad', and what it means to be young and living in their country. A self-completion questionnaire provided data on favourite pastimes, music and media stars, sports teams and so on, and 'awareness', 'usage' and 'preferences' of brands.

On clothing, for example, valuable insight was gained on the rationale behind both aspirational desires and actual behaviour. Male participants aspired to the most obviously status-giving brands, the opulent and clearly expensive (Versace, Armani), reflecting their

desire to buy into a world of success, money and power that everyone (they feel) would appreciate. Female participants, however, were far more interested in standing out from the crowd. They were attracted to the non-mainstream brands which also came with their own strong set of values (e.g. Mambo, Stussy), as well as the more expensive (e.g. Prada and Gucci). What was especially interesting to the researchers was the way that these girls wore brands, mixing and matching, personalising their outfits. As both the boys and girls matured and their repertoire of brands and their financial status grew, they were more inclined to temper their 'ideals'. They moved towards the more High Street designer brands of Diesel, Calvin Klein, Ralph Lauren and DKNY and 'older' stores such as Next and Gap. ■

The previous examples illustrate only a few of the methods used to conduct marketing research, which may range from highly structured surveys with large samples to in-depth interviews with small samples; from the collection and analysis of readily available data to the generation of 'new' data; from direct interaction with consumers to the distant observation of consumers. This book will introduce you to the full complement of marketing research techniques. These examples also illustrate the crucial role played by marketing research in designing and implementing successful marketing plans.[4] Perhaps the role of marketing research can be better understood in light of the basic marketing paradigm depicted in Figure 1.1.

The emphasis in marketing, as illustrated in the Royal Ahold example above, is on the identification and satisfaction of customer needs. To determine customer needs and to implement marketing strategies and plans aimed at satisfying those needs, marketing managers need information about customers, competitors and other forces in the marketplace. In recent years, many factors have increased the need for more and better information. As firms have become national and international in scope, the need for information on larger, and more distant, markets has increased. As consumers have become more affluent, discerning and sophisticated, marketing managers need better information on how they will respond to products and other marketing offerings. As competition has become more intense, managers need information on the effectiveness of their marketing tools. As the environment is changing more rapidly, marketing managers need more timely information.[5]

Marketers make decisions about what they see as potential opportunities and problems, i.e. a process of identifying issues. They go on to devise the most effective ways to realise these opportunities and overcome problems, i.e. a process of solving the issues they have identified. They do this based on a 'vision' of the distinct characteristics of the target markets and customer groups. From this 'vision' they develop,
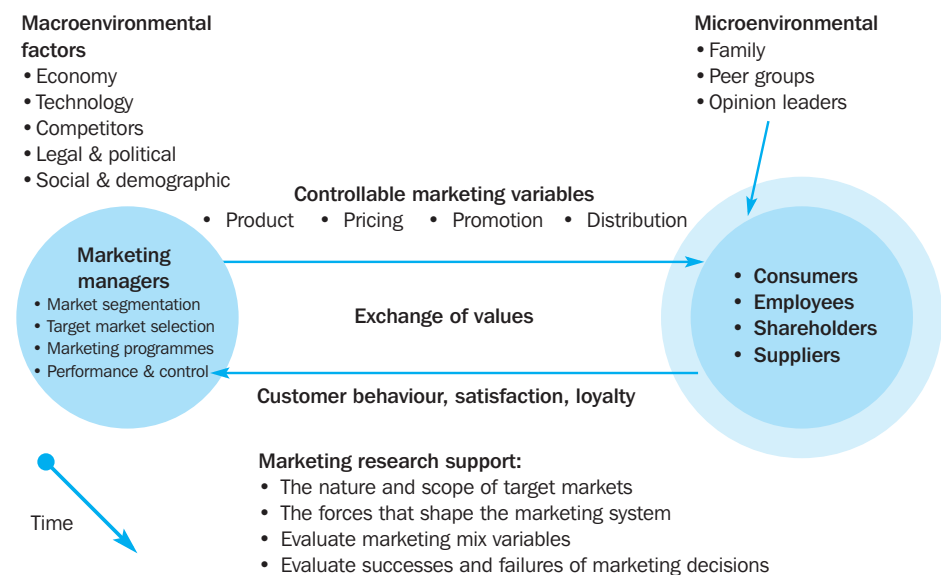
Figure 1.1
The role of marketing research within the marketing system

implement and control marketing programmes. This 'vision' of markets and subsequent marketing decisions may be complicated by the interactive effects of an array of environmental forces that shape the nature and scope of target markets. These forces also affect the marketers' ability to satisfy their chosen target markets.

Within this framework of decision-making, marketing research helps the marketing manager link the marketing variables with their environment and customer groups. It helps remove some of the uncertainty by providing relevant information about marketing variables, environment and consumers.

The role of the marketing researcher in supporting the marketing decision-maker can therefore be summarised as helping to:

- describe the nature and scope of customer groups
- understand the nature of forces that shape customer groups
- understand the nature of forces that shape the marketer's ability to satisfy targeted customer groups
- test individual and interactive marketing mix variables
- monitor and reflect upon past successes and failures in marketing decisions.

Traditionally, marketing researchers were responsible for assessing information needs and providing the relevant information, while marketing decisions were made by the managers. These roles are changing, however, and marketing researchers are becoming more involved in decision-making; conversely, marketing managers are becoming more involved with research. This trend can be attributed to better training of marketing managers and advances in technology. There has also been a shift in the marketing research paradigm, where increasingly marketing research is being undertaken on an ongoing basis rather than in response to specific marketing problems or opportunities on an *ad hoc* basis.[6]

This crucial role of marketing research is recognised in its definition.

The marketing researcher may have to support decision-makers as they reflect upon their failures.

Definition of marketing research

The European Society for Opinion and Marketing Research (ESOMAR) definition of marketing research is given below. For the purpose of this book, which emphasises the need for information in the support of decision-making, marketing research is defined as follows:

Marketing research *is a key element within the total field of marketing information. It links the consumer, customer and public to the marketer through information which is used to identify and define marketing opportunities and problems; to generate, refine and evaluate marketing actions; and to improve understanding of marketing as a process and of the ways in which specific marketing activities can be made more effective.*

Several aspects of this definition are noteworthy. First, it is worth noting the term 'total field of information'. This recognises that marketing decisions are not exclusively supported by marketing research. There are other means of information support for marketers that are now competing with a 'traditional' view of marketing research. For many years, marketing researchers have recognised the competition they face from an array of management consultants, who hope to add value to the data they gather by claiming unique insights and interpretations.[7] In recent years, additional competition has emerged from raw data providers such as call centres, direct marketing, database marketing and telebusinesses.[8] The methods of these competitors may not be administered with the same scientific rigour and/or ethical standards applied in the marketing research industry. Nonetheless, many marketers use these other sources. Marketing research students should be aware of these other sources and appreciate their strengths and weaknesses. Second, it reinforces the notion of basing marketing decisions upon a strong understanding of target customers. It stresses the role of 'linking' the marketer to the consumer, customer and public to help improve the whole process of marketing decision-making.

ESOMAR further qualifies its definition of marketing research by stating:

Marketing research specifies the information required to address these issues [of linking the consumer, customer and public to the marketer]*; designs the method for collecting information; manages and implements the data collection process; analyses the results; and communicates the findings and their implications.*

The above qualification of the definition of marketing research encapsulates the **marketing research process**. The process is founded upon an understanding of the marketing decision(s) needing support. From this understanding, research aims and objectives are defined. To fulfil defined aims and objectives, an approach to conducting the research is established. Next, relevant information sources are identified and a range of data collection methods are evaluated for their appropriateness, forming a research design. The data are collected using the most appropriate method; they are analysed and interpreted, and inferences are drawn. Finally, the findings, implications and recommendations are provided in a format that allows the information to be used for marketing decision-making and to be acted upon directly.

Marketing research should aim to be objective. It attempts to provide accurate information that reflects a true state of affairs. It should be conducted impartially. Although research is always influenced by the researcher's research philosophy, it should be free from the personal or political biases of the researcher or the management. Research motivated by personal or political gain involves a breach of professional standards. Such research is deliberately biased to result in predetermined findings. The motto of every researcher should be 'Find it and tell it like it is'.

Marketing research
A key element within the total field of marketing information. It links the consumer, customer and public to the marketer through information which is used to identify and define marketing opportunities and problems; to generate, refine and evaluate marketing actions; and to improve understanding of marketing as a process and of the ways in which specific marketing activities can be made more effective.

Marketing research process
A set of six steps which define the tasks to be accomplished in conducting a marketing research study. These include problem definition, developing an approach to the problem, research design formulation, fieldwork, data preparation and analysis, and report generation and presentation.

ESOMAR's view of the distinctive contribution of marketing research

ESOMAR distinguishes marketing research from other competitive forms of data gathering, through the issue of the anonymity of respondents. It stresses that in marketing research the identity of the provider of information is not disclosed. It makes a clear distinction between marketing research and database marketing where the names and addresses of the people contacted are to be used for individual selling, promotional, fund-raising or other non-research purposes. These issues will be explored more fully in the 'Ethics in marketing research' section in this chapter. The distinction between marketing research and the database as a research tool is not so clear. There is a growing amount of support given to marketing decision-makers from database analyses that are not 'respondent specific'. It is possible to perform database analyses with the same level of professional standards as is applied in the marketing research industry.

There are many instances where database analyses can add clarity and focus to marketing research activities. For example, since the start of 1995 the highly respected marketing research agency, Taylor Nelson AGB, has been building the European Toiletries & Cosmetics Database (ETCD). Some 14,000 usage diaries of personal care products are collected each year, across Britain, France, Germany, Italy and Spain. Given the huge impact that database analyses are having upon marketing decision-making, these issues will be developed more fully in Chapter 5. In the meantime, the maxim stated by ESOMAR of preserving the anonymity of respondents is vital for the continuing support of respondents and the ultimate health of the marketing research industry.

A classification of marketing research

The ESOMAR definition encapsulates two key reasons for undertaking marketing research: (1) to identify opportunities and problems, and (2) to generate and refine marketing actions. This distinction serves as a basis for classifying marketing research into problem identification research and problem-solving research, as shown in Figure 1.2.

Problem identification research is undertaken to help identify problems that are, perhaps, not apparent on the surface and yet exist or are likely to arise in the future. Examples of problem identification research include market potential, market share,

Problem identification research
Research undertaken to help identify problems that are not necessarily apparent on the surface, yet exist or are likely to arise in the future.

Figure 1.2
A classification of marketing research

Marketing research

Problem identification research

Problem-solving research

Examples:
- market potential research
- market share research
- image research
- market characteristics research
- sales analysis research
- forecasting research
- business trends research

Examples:
- segmentation research
- product research
- pricing research
- promotion research
- distribution research

brand or company image, market characteristics, sales analysis, short-range forecasting, long-range forecasting and business trends research. Research of this type provides information about the marketing environment and helps diagnose a problem. For example, a declining market potential indicates that the firm is likely to have a problem achieving its growth targets. Similarly, a problem exists if the market potential is increasing but the firm is losing market share. The recognition of economic, social or cultural trends, such as changes in consumer behaviour, may point to underlying problems or opportunities.

<div style="border:1px solid #00AEEF;display:inline-block;padding:4px 12px;color:#00AEEF">e x a m p l e</div>

Beer research shows overspill [9]

The proliferation of new brands in Europe's booming beer market could result in shorter product lifecycles and, ultimately, destroy brand loyalty. The report 'Western European Beer' reveals that the industry is in danger of an overspill of brands, with the number of brands available exceeding demand. As a result, consumers are likely to fall back on 'tried and tested brands', normally owned by large manufacturers, at the expense of lesser known brands, which are likely to fall by the wayside. In Western Europe, there is an average per head consumption of 84 litres per year, the highest level in the world. ■

The above research presents either a problem or an opportunity for beer manufacturers, depending largely upon whether your brand is 'tried and tested' or not. The research could be classified as 'business trends research' or even 'market potential research'. It could be followed by individual beer manufacturers conducting their own 'image research' to reveal the extent to which their brand(s) are perceived as being 'tried and tested'.

Problem-solving research
Research undertaken to help solve specific marketing problems.

Once a problem or opportunity has been identified, **problem-solving research** may be undertaken to help develop a solution. The findings of problem-solving research are used to support decisions that tackle specific marketing problems. Problem-solving research is illustrated by the following example of revamping the image of Kellogg's Frosties.

<div style="border:1px solid #00AEEF;display:inline-block;padding:4px 12px;color:#00AEEF">e x a m p l e</div>

Kellogg's Frosties [10]

During 2000, Kellogg wanted to revamp the image of its breakfast cereal Frosties and, more particularly, brand icon Tony the tiger. Tony has been around since 1956, and according to *Advertising Age* magazine is one of the top ten icons of the twentieth century. Kellogg wanted to make sure that Tony was not tired and jaded but a fresh proposition to children.

Their research involved a series of group discussions with 7 to 10 year olds. This involved showing them different ideas and concepts, putting together different images and advertisements. The researchers then observed the children's reactions, to see what made them laugh and got further input, using picture-led questionnaires. The research revealed that there was still a huge affection for Tony but that they wanted him to be a bit sharper and witty while retaining his fun character. The findings also pointed to a repositioning for Tony, away from his traditional sports coach image to that of an action hero. ■

The Frosties example illustrates the refinement of marketing actions. With a brand that has been so successful for many years, Kellogg had to be extremely careful in making changes to the image of Tony. The example illustrates promotions research supporting decisions on advertising copy. This example is clearly one of problem-solving research. A problem-solving perspective enabled management to focus on the promotions development of Frosties. Table 1.1 shows the different types of issues that can be addressed using problem-solving research.

Problem identification research and problem-solving research go hand-in-hand, however, and a given marketing research project may combine both types of research. A

Table 1.1 Examples of problem-solving research

Segmentation research	Determine basis of segmentation
	Establish market potential and responsiveness for various segments
	Select target markets and create lifestyle profiles: demography, media, and product image characteristics
Product research	Test concept
	Determine optimal product design
	Package tests
	Product modification
	Brand positioning and repositioning
	Test marketing
Pricing research	Importance of price in brand selection
	Pricing policies
	Product line pricing
	Price elasticity of demand
	Initiating and responding to price changes
Promotions research	Optimal promotional budget
	Optimal promotion mix
	Copy decisions
	Creative advertising testing
	Evaluation of advertising effectiveness
Distribution research	Attitudes of channel members
	Intensity of wholesale and retail coverage
	Channel margins
	Retail and wholesale locations

marketing research project for a European beer manufacturer that sees its market share diminish may determine through image research that its brand is perceived in a most positive manner. This may indicate that the brand be extended into other types of beer or even into clothes and fashion accessories! Appropriate target markets may be selected, with detailed profiles of potential customers and an associated media and product image. These decisions can clearly be supported with problem-solving research.

The role of marketing research in MkIS and DSS

Marketing information system (MkIS)
A formalised set of procedures for generating, analysing, sorting and distributing pertinent information to marketing decision-makers on an ongoing basis.

Information obtained through the process of marketing research and from sources such as internal records stored on databases and marketing intelligence becomes an integral part of a firm's marketing information systems (MkIS). A **marketing information system (MkIS)** is a formalised set of procedures for generating, analysing, storing and distributing pertinent information to marketing decision-makers on an ongoing basis. At face value, it may seem that the definition of a MkIS is similar to that of marketing research. The difference lies in the MkIS providing information on a continuous basis, guided by a marketing plan. Marketing research forms a component of the MkIS, providing primarily *ad hoc* studies that focus upon new and problematic issues faced by decision-makers. The design of the MkIS takes a much broader view of the support for decision-makers. The MkIS focuses upon each marketing decision-maker's responsibilities, information needs and decision style. The MkIS is designed around information gathered from a variety of sources, such as

customer analyses from invoices and competitor intelligence reported by sales personnel. The continuously gathered information is combined and presented in a format that can be readily used in decision-making. More information can be obtained from an MkIS than from *ad hoc* marketing research projects. However, an MkIS may be limited in the nature of information it provides and the manner in which the information is structured, and cannot easily be manipulated.

Developed to overcome the limitations of the MkIS, decision support systems (DSS) enable marketing decision-makers to interact directly with databases and analysis models.[11] A **decision support system (DSS)** is an integrated information system, including hardware, communications network, database, model base, software base and the marketing decision-maker, that collects and interprets information for decision-making. Marketing research contributes research data to the database, marketing models and analytical techniques to the model base, and specialised programs for analysing marketing data to the software base. DSSs differ from MkISs in various ways (see Figure 1.3).[12] A DSS can combine the use of models or analytical techniques with the traditional access and retrieval functions of an MkIS. They are easier to use in an interactive mode and can be adapted as changes in the marketing system occur, as well as to the decision-making approach of the user. In addition to improving efficiency, a DSS can also enhance decision-making effectiveness by using 'what if' analyses, allowing the creativity of the marketing decision-maker to be reflected in the search and connection between different internal and external data sources. An example of a DSS is the use of geodemographic information systems, which will be covered in more detail in Chapter 5 but is illustrated in the following example.

Decision support system (DSS)
An information system that enables decision-makers to interact directly with both databases and analysis models. The important components of a DSS include hardware and a communication network, database, model base, software base and the DSS user (decision-maker).

Geographic marketing at Banco Central Hispanoamericano[13]

There are two important geographic features of the Spanish banking market: the concentration of the population in cities and the large number of bank branches. Forty-two per cent of the population is concentrated in cities with more than 100,000 inhabitants. In these cities, a DSS in the form of a geodemographic information system (GIS) (which combines customer data with the geographic data on a map of a city) is very useful to define customer distribution areas based on their banking behaviour.

Banco Central Hispanoamericano have designed a model within their GIS which allows them to assign customers to branches, maximising the number of customers assigned to each branch while minimising the distance they have to travel. To do this, the system considers the maximum distance (by street) that a customer is capable of travelling to go to a branch, the maximum number of branches along which to distribute the customers, the total number of branches to analyse, and the maximum distance over which to carry out the analysis.

The system also allows the bank to select new locations for automated teller machines (ATMs) in an 'intelligent' way. When opening a new ATM, it is possible to select customers of the Bank who use the machines of competitor banks close by and who also live in its area of influence. This facilitates promotions activities, targeted at customers with the greatest likelihood of using the new service. ∎

As shown by the experience of Banco Central Hispanoamericano, a DSS can greatly enhance the support given to marketing decision-makers. To make an MkIS or a DSS work properly, the data used in the systems have to be accurate and totally

Figure 1.3
Marketing information systems versus decision support systems

MkIS	DSS
• structured problems	• unstructured problems
• reports 'fed' to decision-makers	• use of models
• rigid structure	• user-friendly interaction
• information display restricted	• adaptability
• combines quantitative and qualitative data	• uses 'what if' analyses

dependable. To generate accurate data, management relies upon well-devised internal data systems and on the generation of external data through marketing research suppliers and services.

Marketing research suppliers and services

Internal supplier
Marketing research department located within a firm.

Marketing research suppliers provide most of the information needed for making marketing decisions. Figure 1.4 classifies marketing research suppliers and services. Broadly speaking, research suppliers can be classified as internal or external. An **internal supplier** is a marketing research department or function located within a firm. Many firms, particularly large ones, maintain in-house marketing research departments. A marketing research department's place in an organisation structure may vary quite considerably. At one extreme, the research function may be centralised and located at the corporate headquarters, allowing the development of a range of skills and expertise to be built up in a team. At the other extreme is a decentralised structure in which the marketing research function is organised along divisional lines, allowing specific and focused expertise to be developed but without the breadth of experience that may lie in a centralised team. In a decentralised scheme, the company may be organised into divisions by products, customers or geographical regions, with marketing research personnel assigned to the various divisions. These personnel generally report to a division manager rather than to a corporate-level executive. In addition, between these two extremes is a variety of types of organisation. Even if a firm has its own marketing research specialists, it may still turn to external suppliers to perform specific marketing research tasks.

External suppliers
Outside marketing research companies hired to supply marketing research services.

External suppliers are outside firms hired to supply marketing research data. These external suppliers collectively comprise the marketing research industry. These suppliers range from small (one or a few persons) operations to very large global corporations. Table 1.2 lists the top 12 marketing research suppliers in the world. Table 1.3 lists the spending on marketing research in different countries; the growth in spend and population sizes are shown to indicate the relative importance of the industry to the economy. One country of note in Table 1.3 is Hungary, whose developing economy is supported by a high growth and dynamic marketing research
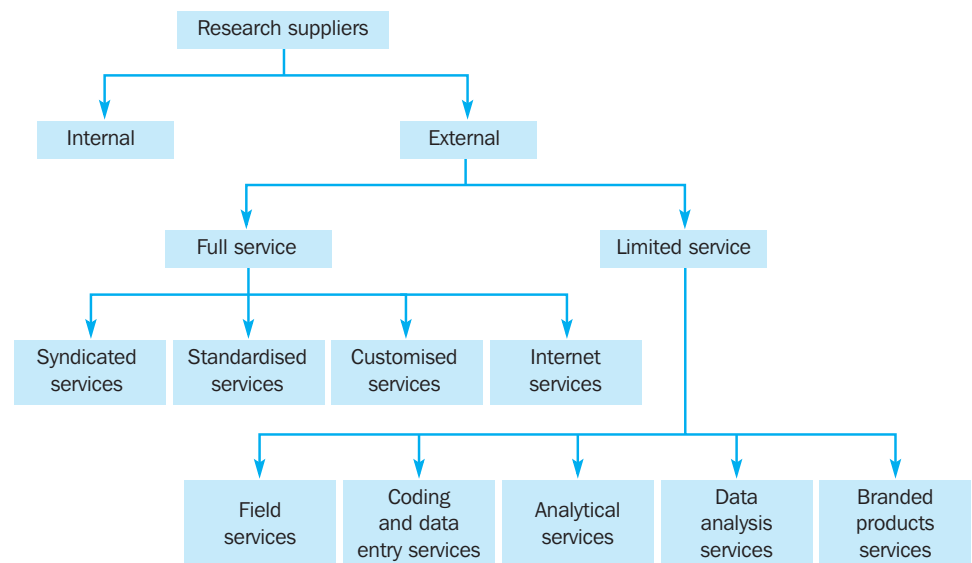
Figure 1.4
Marketing research suppliers

Table 1.2 **World top 12 market research companies, 2000**[14]

Company	Global research revenues (euro m)	Countries with offices	Head office location	Number of full-time employees	Percentage of revenues from outside home country
1 AC Nielsen Corp.	1,680	80	USA	21,000	67%
2 IMS Health, Inc.	1,205	74	USA	8,000	63%
3 The Kantor Group	989	59	UK	5,800	71%
4 Taylor Nelson Sofres Plc	756	41	UK	7,125	75%
5 Information Resources, Inc.	566	17	USA	4,000	25%
6 VNU, Inc.	561	21	USA	2,916	3%
7 NFO WorldGroup, Inc.	496	38	USA	3,500	63%
8 GfK Group	473	34	Germany	4,212	64%
9 Ipsos Group SA	324	24	France	2,437	78%
10 Westat, Inc.	282	1	USA	1,430	0%
11 NOP World	262	6	UK	1,302	60%
12 Aegis Research	247	12	UK	1,980	60%

industry. Below Table 1.3 is an example which explains the role of marketing research in Hungary's development in more detail. Table 1.4 lists where marketing research monies are spent worldwide through external suppliers. External suppliers can be classified as full-service or limited-service suppliers.

Table 1.3 **World marketing research turnover in euros, with growth, adspend and population, 2000**[15]

Country	Turnover (euro m)	Percentage growth in local currencies compared with 1999 (adjusted for inflation)	Advertising spend (euro m)	Population in millions
USA	6,307	5%	128,015	273.2
UK	1,728	6%	19,065	58.7
Germany	1,374	3%	21,339	82.1
Japan	1,284	8%	35,742	126.5
France	1,020	8%	10,293	59.1
Italy	442	9%	7,553	57.3
Spain	291	0%	5,660	39.4
Netherlands	243	−1%	4,136	15.8
Sweden	243	2%	2,033	8.9
China	193	25%	4,428	1266.8
Switzerland	126	9%	2,781	7.1
Belgium	121	6%	2,115	10.2
Austria	84	−2%	2,063	8.2
Denmark	82	−1%	1,460	5.3
Finland	81	11%	1,236	5.1
Poland	75	10%	1,412	38.7
Norway	73	9%	1,190	4.4
Portugal	49	8%	1,128	10
Greece	47	9%	1,268	10.6
Ireland	42	9%	650	3.8
Hungary	34	17%	567	10.1
Czech Rep.	32	12%	616	10.3

Table 1.4 Europe still has the largest research market in the world – just! [16]

Region	Turnover (euro m)*	Percentage of world turnover	Increase based on euro 1999/2000 (%)
Europe	6,452	39	11.1%
USA	6,428	39	26.0%
Central/South America	766	5	27.7%
Japan	1,309	8	31.3%
Asia Pacific	969	6	28.6%
Other	619	4	35.8%
Total world	16,543	100	20.3%

* Excludes in-house research by marketing departments, advertising agencies, governmental and academic institutions.

Full-service suppliers
Companies that offer the full range of marketing research activities.

Syndicated services
Companies that collect and sell common pools of data designed to serve information needs shared by a number of clients.

Standardised services
Companies that use standardised procedures to provide marketing research to various clients.

Customised services
Companies that tailor research procedures to best meet the needs of each client.

Internet services
Companies which specialise in the use of the Internet to collect, analyse and distribute marketing research information.

Limited-service suppliers
Companies that specialise in one or a few phases of a marketing research project.

Full-service suppliers offer the entire range of marketing research services, for example defining a problem, developing a research design, conducting focus group interviews, designing questionnaires, sampling, collecting, analysing and interpreting data, and presenting reports. They may also address the marketing implications of the information they present. The services provided by these suppliers can be further broken down into syndicated services, standardised services and customised services (see Figure 1.4).

Syndicated services collect information that they provide to subscribers. Surveys, diary panels, scanners and audits are the main means by which these data are collected.

Standardised services are research studies conducted for different clients but in a standard way. For example, procedures for measuring advertising effectiveness have been standardised so that the results can be compared across studies and evaluative norms can be established.

Customised services offer a variety of marketing research services specifically designed to suit a client's particular needs. Each marketing research project is treated uniquely.

Internet services offer a combination or variety of secondary data and intelligence gathering, survey or qualitative interviewing, and the analysis and publication of research findings, all through the Internet.

Limited-service suppliers specialise in one or a few phases of a marketing research project. Services offered by such suppliers are classified as field services, coding and data entry, analytical services, data analysis, and branded products.

e x a m p l e

Marketing and opinion research in Hungary grows more than 20% each year [17]

Marketing and opinion research is a fast-growing industry in an emerging market such as Hungary, and shows more dynamic growth than the economy itself. Ten years ago, marketing research was a small sector and its contribution to the decision-making process in business and society was not very well known in Hungary. Basic market research did exist in some state-owned organisations, but few people were knowledgeable about the key research methods – quantitative research, qualitative research, focus groups, or what they meant.

Contrast this with the situation today where the turnover of the Hungarian research industry reached about 30 million US dollars in 2000, according to ESOMAR's latest global market research study. Following a period of dynamic growth, the sector grew by 20% in 2000 (17% adjusted for inflation) over the previous year. The research industry in Hungary employs around 5,000 people of whom 4,000 are interviewers and 1,000 are researchers, data processors and support staff. The market research industry in Hungary is young with all the advantages and potential of a young sector. It is dynamic, talented, is open for new things and learning fast. ■

Field services
Companies whose primary service offering is their expertise in collecting data for research projects.

Coding and data entry services
Companies whose primary service offering is their expertise in converting completed surveys or interviews into a usable database for conducting statistical analysis.

Analytical services
Companies that provide guidance in the development of research design.

Data analysis services
Firms whose primary service is to conduct statistical analysis of quantitative data.

Branded marketing research products
Specialised data collection and analysis procedures developed to address specific types of marketing research problems.

Field services collect data through mail, personal interviews or telephone interviews, and firms that specialise in interviewing are called field service organisations. These organisations may range from small proprietary organisations that operate locally to large multinationals. Some organisations maintain extensive interviewing facilities across the country for interviewing shoppers. Many offer qualitative data collection services such as focus group interviewing (discussed in detail in Chapter 7).

Coding and data entry services include editing completed questionnaires, developing a coding scheme, and transcribing the data on to diskettes or magnetic tapes for input into a computer.

Analytical services include designing and pretesting questionnaires, determining the best means of collecting data, and designing sampling plans, as well as other aspects of the research design. Some complex marketing research projects require knowledge of sophisticated procedures, including specialised experimental designs (discussed in Chapter 10) and analytical techniques such as conjoint analysis and multidimensional scaling (discussed in Chapter 24). This kind of expertise can be obtained from firms and consultants specialising in analytical services.

Data analysis services are offered by firms, also known as tab houses, that specialise in computer analysis of quantitative data such as those obtained in large surveys. Initially, most data analysis firms supplied only tabulations (frequency counts) and cross-tabulations (frequency counts that describe two or more variables simultaneously). Now, many firms offer sophisticated data analysis using advanced statistical techniques. With the proliferation of microcomputers and software, many firms now have the capability to analyse their own data, but data analysis firms are still in demand.

Branded marketing research products and services are specialised data collection and analysis procedures developed to address specific types of marketing research problems. These procedures may be patented, given brand names, and marketed like any other branded product. *Microscope* by Retail Marketing (In-Store) Services is an example of a branded product. It is a test marketing package for new product development that supplies cost-effective measurements of new product performance.

The marketing research process

The marketing research process consists of six broad stages. Each of these stages is discussed in detail in subsequent chapters; thus, the discussion here is brief. The process illustrated in Figure 1.5 is of a simple linear nature. Figure 1.6 takes the process a stage further to show the many iterations and connections between stages. This section will explain the stages and illustrate the connections between the stages.

Step 1: Problem definition. The logical starting point in wishing to support the decision-maker is trying to understand what marketing problem is being tackled. Marketing problems are not simple 'givens', as will be discussed in Chapter 2, and the symptoms and causes of a problem are not as neatly presented as they may be in a case-study such as those found in marketing texts. In Figure 1.6, the first three stages show the iterations between *environmental context of the problem, management decision problem* and *marketing research problem*. Understanding the environmental context of the problem has distinct stages that will be discussed in Chapter 2. It involves discussion with decision-makers, in-depth interviews with industry experts, and the collection and analysis of readily available published information (from both inside and outside the firm). Once the problem has been precisely defined, the researcher can move on to designing and conducting the research process with confidence.
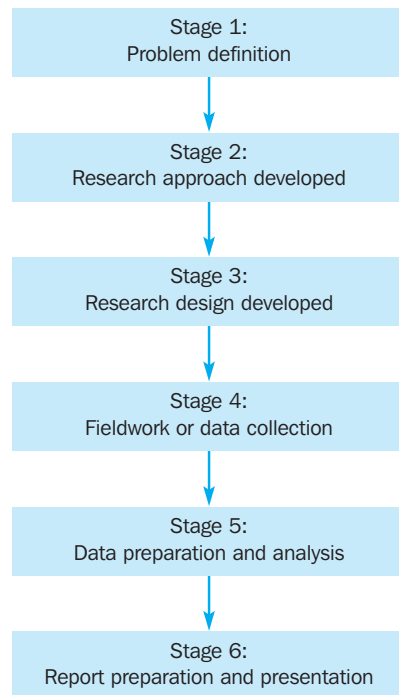
Figure 1.5
Simple linear description of the marketing research process

The figure shows a vertical flowchart with the following stages connected by downward arrows:

Stage 1: Problem definition

Stage 2: Research approach developed

Stage 3: Research design developed

Stage 4: Fieldwork or data collection

Stage 5: Data preparation and analysis

Stage 6: Report preparation and presentation

Step 2: Development of an approach to the problem. The development of an approach to the problem involves identifying factors that go on to influence research design. A key element of this stage involves the selection, adaptation and development of an appropriate theoretical framework. Understanding the interrelated characteristics of the nature of target respondents, the issues to be elicited from them and the context in which this will happen rely upon 'sound' theory. 'Sound' theory helps the researcher to decide 'what should be measured or understood' and 'how best to encapsulate and communicate the measurements or understandings'. In deciding what should be either measured or encapsulated, the researcher also develops a broad appreciation of how the data they collect will be analysed. The issues involved in developing an approach are tackled in more detail in Chapter 2.

Step 3: Research design developed. A research design is a framework or blueprint for conducting a marketing research project. It details the procedures necessary for obtaining the required information. Its purpose is to establish a study design that will either test the hypotheses of interest and/or determine possible answers to set research questions, and ultimately provide the information needed for decision-making. Conducting any exploratory techniques, precisely defining variables to be measured, and designing appropriate scales to measure variables can also be part of the research design. The issue of how the data should be obtained from the respondents (for example, by conducting a survey or an experiment) must be addressed. These steps are discussed in detail in Chapters 3 to 13.

Step 4: Fieldwork or data collection. In Figure 1.5, this stage could be simplified to 'collecting the required data'. In Figure 1.6, a whole array of relationships between stages of data collection is shown, starting at *Secondary data collection and analysis* through to *Quantitative research* or *Qualitative research*. The process starts with a more thorough collection and analysis of secondary data sources. Secondary data are data collected for some other purpose than the problem at hand. They may be held within the organisation such as databases that detail the nature and frequency of customer purchases, through to surveys that may have been completed some time ago
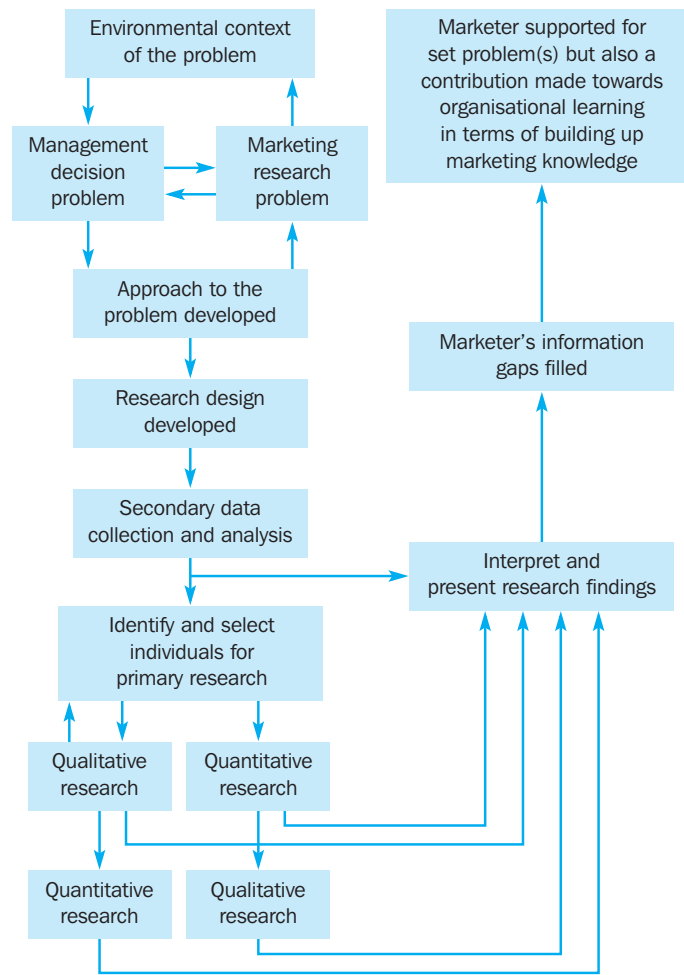
Figure 1.6
The marketing research process, detailing iterations between stages

that may be accessed through libraries, CD-ROMs or the Internet. Going through this stage avoids replication of work and gives guidance in sampling plans and in deciding what to measure or encapsulate using quantitative or qualitative techniques. Secondary data collection and analysis may complete the research process, i.e. sufficient information may exist to interpret and report findings to a point whereby the information gaps that the decision-maker has are filled. The case may be that the secondary data form a vital foundation and a clear focus to primary data collection.

In Figure 1.6, the stage of *Identify and select individuals for primary research* covers sampling issues for both quantitative and qualitative studies. This stage may include the selection of individuals for in-depth qualitative research. In qualitative research, issues of 'representativeness' are less important than the quality of individuals targeted for investigation and the quality of response elicited. However, as can be seen from the line leading up from *Qualitative research* to *Identify and select individuals for primary research*, the qualitative research process may help in the identification and classification of individuals who may be targeted using more formal sampling methods. These sampling methods are covered in detail in Chapters 14 and 15.

Beyond issues of identifying and selecting individuals, the options available for primary data collection vary considerably. A stage of *Qualitative research* alone may be sufficient to support the decision-maker, as indeed could a stage of *Quantitative research*. The following example illustrates the use of qualitative observation to support marketing decision-making. This example mirrors one of the research

techniques used by Royal Ahold (presented at the start of this chapter), helping them to build up a 'picture' of their customers. As a technique in its own right, it does not necessarily have to be followed by a survey or quantitative work to confirm the observations. This technique will be developed under the heading 'Ethnographic research' in Chapter 6.

Supermarket sweep [18]

Video camera analysis can be used to gain vital information about the way in which supermarket consumers spend their money. People behave in certain ways within the supermarket environment, according to store design specialist company ID Magasin. Patterns of consumer behaviour can be used to maximise profits. Store managers who work on store design without video footage are sometimes wrong about their conclusions. Consumers, for example, often ignore the products placed on shelves at eye-level, despite popular belief to the contrary. ■

The research problem may require a stage of qualitative and quantitative research to run concurrently, perhaps measuring and encapsulating different characteristics of the problem under investigation.

A stage of qualitative research could be used to precede a stage of quantitative research. For example, a series of focus groups may help to generate a series of statements or expectations that are subsequently tested out in a survey to a representative sample. Conversely, a survey may be conducted and, upon analysis, there may be clear statistically significant differences between two distinct target markets. A series of qualitative in-depth interviews may follow to allow a more full exploration and understanding of the reasons for the differences between the two groups.

Step 5: Data preparation and analysis. Data preparation includes the editing, coding, transcription and verification of data. In Figure 1.6, this stage is not drawn out as a distinct stage in its own right, but is seen as integral to the stages of *Secondary data collection and analysis* through to *Quantitative research* or *Qualitative research.* The process of data preparation and analysis is essentially the same for both quantitative and qualitative techniques, for data collected from both secondary and primary sources. Considerations of data analysis do not occur after data has been collected; such considerations are an integral part of the development of an approach, the development of a research design, and the implementation of individual quantitative or qualitative methods. If the data to be collected are qualitative, the analysis process can occur as the data are being collected, well before all observations or interviews have been completed. An integral part of qualitative data preparation and analysis requires the researcher to reflect upon their own learning and the ways they may interpret what they see and hear. These issues will be developed in Chapters 6 to 9.

If the data to be analysed are quantitative, each questionnaire or observation form is inspected or edited and, if necessary, corrected. Number or letter codes are assigned to represent each response to each question in the questionnaire. The data from the questionnaires are transcribed or keypunched into a proprietary data analysis package. Verification ensures that the data from the original questionnaires have been accurately transcribed, whereas data analysis gives meaning to the data that have been collected. Univariate techniques are used for analysing data when there is a single measurement of each element or unit in the sample; if there are several measurements of each element, each variable is analysed in isolation (see Chapter 18). On the other hand, multivariate techniques are used for analysing data when there are two or more measurements of each element and the variables are analysed simultaneously (see Chapters 18 to 24).

Step 6: Report preparation and presentation. The entire project should be documented in a written report that addresses the specific research questions identified, describes the approach, research design, data collection and data analysis procedures adopted, and presents the results and major findings. Research findings should be presented in a comprehensible format so that they can be readily used in the decision-making process. In addition, an oral presentation to management should be made using tables, figures and graphs to enhance clarity and impact. This process is encapsulated in Figure 1.6 with the reminder that the marketer's information gaps are filled and that the marketer is supported for the set problem but also a contribution is made towards organisational learning in terms of building up marketing knowledge (see Chapter 25).

Our description of the marketing research process is typical of the research being done by major corporations. The following case of Land Rover illustrates some of the activities involved in the marketing research process. Most importantly it shows the vital link between marketing decision-makers and marketing researchers in generating information that can give a clear competitive edge.

e x a m p l e   **Land Rover brand values across Europe** [19]

For Land Rover, it is vitally important to regularly measure consumers' knowledge, image and desirability of the Land Rover brand. To this end, Land Rover monitors its brand values across Europe via its image tracking study, CATS, which provides a measure of the health of the brand and its composite values in each country surveyed. Since Europe is the most important region for Land Rover in terms of product sales, this defined the focus of their research. The countries covered on a continuous basis were Britain, France, Germany, Italy, Portugal and Spain.

The need for research to address the issue of brand values within Rover Group Marketing Research stemmed from the fact that they were embarking on a global strategy of brand value focus. The challenge was how to convert the strategy into something that was measurable, in other words, how to make the marketing information requirements operational from a research standpoint. Translating this challenge into the development of a research design set the challenge of how to quantify the overtly emotional brand values and those values that are based on the more concrete products. This had to be achieved, keeping the study as simple and elegant as possible.

The data collection methodology adopted used Computer Assisted Personal Interviewing (CAPI). New car buyers were interviewed covering the sectors that Land Rover competes in:

- 4x4 sector – with Range Rover, Discovery and Defender
- Luxury sector – with Range Rover
- 4x2 sector – with Discovery and Range Rover.

With a clear research design, the most important element to address Land Rover's set of objectives was to design a workable international questionnaire. The most important needs in developing the questionnaire were as follows:

- To develop common measures to be adopted across all markets
- To distinguish between the emotional and rational elements of brand image
- To measure the Land Rover brand in a meaningful context.

These requirements resulted in a questionnaire development process of four stages:

1 Contents determination, e.g. image dimensions
2 Questionnaire piloting – using a qualitative studio facility
3 Questionnaire translation
4 Agency face-to-face briefing (in individual countries where the fieldwork was to be conducted).

A great deal of time was allocated to the structure of the reports to ensure that they focused on the most salient findings of Land Rover's progress in its competitive context. The information was carefully marshalled to provide brand management with input into the brand planning cycle without swamping them with masses of data.

The philosophy of taking data to wisdom is at the heart of Rover Group Marketing Research. The objective was to transform data along the continuum to information, then knowledge and ultimately wisdom. Land Rover clearly sees the wisdom of being focused and driven by the consumer. Such a focus requires regular, timely and actionable consumer-based research. By being passionate about the brand and the customer, Land Rover survives and indeed thrives in today's fiercely competitive environment. ■

The limitations of marketing research

If decision-makers have gaps in their knowledge, if they perceive risk and uncertainty in their decision-making and cannot find support at hand within their organisation, they can gain support from marketing research. However, many marketers can recount cases where the use of marketing research has resulted in failure or where decisions based upon gut feeling or intuition have proved to be successful. Such cases present a challenge to marketing researchers, especially in light of the competition faced by the industry from alternative information sources as discussed earlier. Reflecting upon such cases should remind marketing researchers to maintain a focus of offering real and valuable support to decision-makers. Understanding what real and valuable support means should underpin the whole array of creative data collection and analysis procedures available to the marketing researcher. The following example starts this reflection process with a case that is very close to home!

Uncle Jim heads off to the Caribbean for his summer holidays, using the money he saved by not conducting marketing research.

example | **What's this marketing research then, Dave?**

James Birks founded and successfully ran a kiln construction company for over 40 years. He designed, built and maintained kilns for some of the most demanding porcelain and ceramics manufacturers worldwide, including Wedgwood, Doulton and Spode. At retirement age he sold his company as a going concern – a very wealthy man.

James was presented with a copy of the first edition of this text by his nephew, David Birks. He was very pleased with the present but was intrigued by the title and asked 'what's this marketing research then, Dave?'. He certainly had a clear idea of what marketing meant to his business and what was involved in being a successful marketer in his industry, but the notion of researching marketing activities was alien to him. ■

The intriguing aspect of this question is that James Birks had run a successful business on an international basis for over 40 years without the need to be aware of or to use marketing research. Had he used marketing research, could he have been even more successful, or would it have been a wasted investment? Could he have been practising marketing research 'activities' in a very informal manner to support marketing decisions? In his business-to-business marketing situation, he knew his customers and competitors well, he knew what shaped their demands. This knowledge he acquired on a day-to-day basis, nurturing a curiosity about opportunities and how to realise them – without resorting to support from formal *ad hoc* marketing research. The example of James Birks shows that decision-makers do not rely solely upon marketing research, and in certain circumstances can survive and perform well without it.

19

Another view to reflect upon is the damning comment from Anita Roddick of The Body Shop, who said that 'market research is like looking in the rear view mirror of a speeding car'.[20] This may be a valid point if one sees the relationship of marketing and marketing research from the perspective illustrated by the respected research practitioner, Wendy Gordon:

> *Traditional marketers delegate responsibility to the processes of marketing research. They believe that you can ask people what they want and need in the future and then deliver it to them. It is a fallacy. Marketing research removes people from experiencing reality, where the signs of change bubble up in unexpected places. Sitting in comfort behind a one-way mirror, listening to a debrief from researchers describing the world 'out there' or reading statistical reports on markets and the 'aggregate consumer' is not real, it is sanitised and second-hand.*[21]

Given the above criticisms, it is a fair point to acknowledge that there are cases where the use of marketing research has resulted in poor decision-making or even failure. Ultimately, this examination should lead to a stronger justification of what ensures strong and valuable marketing research support. It may be a painful path to tread but this journey has to be made!

There are two areas of misconception of the role of marketing research:[22]

1 *Marketing research does not make decisions.* The role of marketing research is not to make decisions. Rather, research replaces hunches, impressions or a total lack of knowledge with pertinent information.
2 *Marketing research does not guarantee success.* Research, at best, can improve the odds of making a correct decision. Anyone who expects to eliminate the possibility of failure by doing research is both unrealistic and likely to be disappointed. The real value of research can be seen over a long period where increasing the percentage of good decisions should be manifested in improved bottom-line performance and in the occasional revelation that arises from research.

The last point shows the long-term benefits of conducting marketing research, i.e. that the results of a study may help decision-makers with an immediate problem, but by building their knowledge they can also have long-term benefits.

A great proportion of marketing research money has been spent on developing and testing new products. Questions have been posed that examine the role marketing research plays in poor success rates of new products across a wide range of industries.[23] Clearly, marketing research cannot be blamed for every failure, and the following two reasons may explain why decision-makers can make poor decisions when sound research has been conducted:[24]

1 *Blind optimism/disbelief in research.* Many patently bad products have been launched because marketing management did not believe research findings.
2 *Political pressures.* Given how many personal reputations may be at stake in the lengthy and costly process of new product development, there may be political pressures to launch a 'borderline' case.

The following example illustrates where research was used in new product development but the designer and entrepreneur chose to ignore the findings, and ultimately achieved immense levels of success.

example

Doing a Dyson[25]

Just 23 months after its launch in the UK, the Dyson bagless vacuum cleaner became Britain's best seller, overtaking sales of Hoover, Electrolux, Panasonic, Miele and all other vacuum cleaners. The Dyson clear bin was given a resounding thumbs-down in marketing research. People said they did not like the dirt being visible in the bin in case their neighbours

saw how much dirt had been picked up in their homes. Some retailers said they would not want to have dust on display in demonstration machines. Yet, the dust was there because they began using Dyson display machines to clean their shops. Dyson felt compelled to launch their vacuum cleaner with a clear bin, believing that it is important to see when it is full. Moreover, what better way was there to show stockists, sales staff and customers proof of its increased efficiency than to see the dirt being collected?

How would consumers react to a new vacuum cleaner with totally radical styling, revolutionary internal engineering and a price tag almost twice that of the current brand leader? The public's response proved immediately that innovative products do sell, even at a premium price. However, marketing research did not point to this product having the potential to be a success. Dyson argued that 'marketing research will only tell you what has happened. No research can tell you what is going to happen'. ∎

Other researchers offer reasons why decision-makers may reject research findings. These go beyond the confines of developing and testing new products. Examples of these reasons include:[26]

1 *Invalidity of research methods.* If the decision-maker suspects the accuracy or appropriateness of the methods for the problem faced, lack of confidence would lead to a rejection of findings.
2 *Faulty communication.* This would lead to the findings being difficult to comprehend or utilise, or being unconvincing.
3 *Irrelevance.* The decision-maker may see the findings as irrelevant to the perceived marketing problem.

If decision-makers reject research findings, and have been taken to task as a consequence, they can easily support their stance. They could question the time taken to conduct research (and the possible consequences of delays in decision-making), and the cost required (not only in paying for marketing research projects but also in managerial effort). They could point to the few reliable methods of evaluating the return on investment. They may also question the amount of support that can be given from a researcher who may have little understanding of the context and environment in which the decisions are made.

Collectively, the examples in this section illustrate that out of the array of research and information support approaches, there is no one guaranteed approach, research design or technique that can create the perfect means to support decision-makers. Given the vast array of marketing decisions and styles of marketing decision-maker, the hallmark of marketing decision-making is creativity. Creativity in decision-making can be supported well by accurate and relevant information. Generating accurate and relevant information is also a creative act. The diagnosis of problems, the measurement of consumers and the interpretation of those measurements are all creative acts, not a set system.

If decision-makers complain that research is misleading or is only telling them what they already know, the marketing researcher may argue that the fault lies with managers who pose the wrong questions or problem in the first place. If one takes the narrow view that the decision-maker poses the questions and the researcher finds the answers, there may be some validity in such an argument. It does not hold if one considers that the decision-maker and the researcher have a joint commitment to solve problems, that they have quite distinct but complementary creative skills that they can bring together to understand what problem they should be researching, how they conduct the research, and how they interpret their findings.

➤➤➤
See Professional Perspectives 1, 2, 3 and 13.

The above perspective is neatly encapsulated in the views of two marketing research practitioners in the Professional Perspectives section on the Companion Website. See Professional Perspective 1 for Wendy Gordon's article 'Be creative to innovate: market

research can be the enemy of innovation' and Professional Perspective 2 for Miranda Forestier-Walker's article 'Research is not a substitute for talent and skills'.

Professional Perspective 3 by Magnus Kristjansson and Magnus Baldursson contends that marketing research is not just about questioning what consumers 'want'. They go on to argue, and illustrate, the role of marketing research in 'creating' customers. See also Professional Perspective 13, based on a speech by the Chairman of Heineken to an audience of European marketing researchers. His speech poses questions about the role that marketing researchers play in supporting decision-makers.

Can marketing researchers survive in an age of increasing competition from other information providers? How can the industry fend off the challenge from the armies of consultants and other information suppliers? To achieve this, the industry has to offer clients insights rather than just data analysis. The marketing researcher's input must be seen to benefit the bottom line. Initiatives that bring clients and researchers closer together are needed – initiatives that educate buyers that marketing research has as much if not more to offer than very expensive consultancy firms.[27]

Supporting decision-makers in pan-European banking

e x a m p l e

GlobalCash Project

Supporting pan-European banking decisions

A research project designed to understand the cash management and electronic banking practices of Europe's largest companies is used to provide examples throughout this text. The examples illustrate particular techniques and data analysis procedures. They may be read as a case that runs throughout the book or individually; all that is needed is to appreciate the context of the study presented here.

GlobalCash has been directed by the English co-author, coordinating 19 business schools throughout Europe in 1996, 1998, 2000 and 2002. It is a study conceived, financed and administered by The Bank Relationship Consultancy using a mixture of quantitative and qualitative techniques. The targets to be researched are the largest companies in 19 European countries, including the Eastern European countries of the Czech Republic, Hungary and Poland. The whole project is funded by major banks from Britain, France, Germany, Ireland, Spain and the USA, with additional sponsorship from Treasury Management Associations in individual countries. The subjects tackled in the study focus upon cash management practices, service quality, electronic banking and the impact of European Monetary Union. Of particular importance to banks are questions that indicate changes in market share in individual countries, evaluations by companies of their service delivery, and companies' plans for the future.

The main methods used were:

- a postal survey of companies
- a telephone survey focused upon the impact of the euro
- an Internet survey focused upon innovations in response to the euro
- in-depth interviews in each European country
- a feedback workshop in Brussels, whereby the survey findings were presented to questionnaire respondents followed by qualitative workshops.

The findings of the GlobalCash studies were used to develop the marketing, operational and IT strategies of major banks operating in Europe. As the study is conducted every two years, measurement of the effect or outcome of past strategies can be conducted. ■

International marketing research

With the spread of marketing and research skills has come a noticeable decline in a 'national research culture'. There was a time when each country had a stubbornly distinctive approach to research, making it extremely difficult to get a consistent research design across markets. Most people are aware now that there are different, equally legitimate, ways to approach research problems and that no one school of thought has absolute authority for all types of problem. This greater flexibility has made multi-country coordinated projects much more feasible – not easier, as they represent intellectually, logistically and diplomatically the most demanding of problems.

Conducting international marketing research is much more complex than conducting domestic marketing research. All research of this kind will be discussed and illustrated in individual chapters as individual techniques are developed and in greater detail in Chapter 26. The marketing, government, legal, economic, structural, socio-cultural and informational environments prevailing in target international markets, and the characteristics of target consumers that are being studied, influence the manner in which the marketing research process should be performed. Examples of these environmental factors and their impact on the marketing research process are illustrated in detail in subsequent chapters. The following example illustrates the problems and challenges faced by researchers in the European context.

example

Crossing borders [28]

What is this thing called Europe? Over a decade after the collapse of the Berlin Wall and with the levelling force of globalisation at play, arriving at a meaningful definition is harder than ever. Despite binding influences like a single currency, Europe's national marketing research markets have distinct cultural, cost and quality characteristics. For example, German marketing research still feels the effects of reunification, while Spanish agencies are trying to leverage Spain's historical links with Latin America to capture a slice of the coordination market there.

For research buyers and users, the diversity of Europe is very real, and the trickiest part of their remit is to get comparable data systems across the different country markets. Andrew Grant, European Research Director at Ford, says: 'From our perspective, actually getting to a single methodology, a single questionnaire, a single data collection method and a single set of attributes that means the same thing across markets is a hugely difficult part of what we do.'

Europe may be moving towards becoming a single market, but from a research cost perspective it is still far from being a level playing field. In the absence of tax and social cost harmonisation across Europe, countries like France and Sweden are two of the region's most expensive places to conduct research. The euro will have an increasingly positive effect, agencies say, for those countries that participate. 'It makes contracting and financing easier,' says Klaus Wübbenhorst, Chief Executive of German-based group GfK. 'The euro won't mean that prices in all European countries will be the same, you have to take into account different levels of productivity and efficiency. But it makes pricing more transparent between countries.'

Against a background of centralisation, standardisation and increasing cooperation, Europe's researchers believe that the region's defining diversity will survive. 'The fact that we're multi-cultural means that we have great creative potential,' says Meril James of GIA. ◼

Ethics in marketing research

Ethical issues arise in marketing research for several reasons. Marketing research often involves contact with the respondents and the general public, usually by way of data collection, dissemination of the research findings, and marketing activities such as advertising campaigns based on these findings. Thus, there is the potential to abuse or misuse marketing research by taking advantage of these people, for example by misrepresenting the research findings in advertising. As explained earlier, marketing research is generally conducted by commercial (i.e. for-profit) firms which are either independent research organisations (external suppliers) or departments within corporations (internal suppliers). Most marketing research is conducted for clients representing commercial firms. The following example summarises the basic principles of the Code of Conduct as presented by ESOMAR. The full array of ESOMAR codes of conduct can be viewed on: **www.esomar.nl/codes_and_guidelines.html.**

The principles detailed in these codes will be developed and illustrated at the end of most chapters.

e x a m p l e

ESOMAR Code of Practice: basic principles

Effective communication between the suppliers and the consumers of goods and services of all kinds is vital to any modern society. Growing international links make this even more essential. For a supplier to provide in the most efficient way what consumers require they must understand their differing needs, how best to meet these needs, and how they can most effectively communicate the nature of the goods or services under offer. This is the objective of marketing research.

Such research depends upon public confidence: confidence that it is carried out honestly, objectively, without unwelcome intrusion or disadvantage to respondents, and that it is based upon their willing cooperation. This confidence must be supported by an appropriate professional Code of Practice, which governs the way in which marketing research projects are conducted.

The first such Code was published by ESOMAR in 1948. This was followed by a number of Codes prepared by national marketing research societies and by other bodies such as the International Chamber of Commerce (ICC), which represents the international marketing community. In 1976 ESOMAR and the ICC decided that it would be preferable to have a single International Code instead of two differing ones, and a joint ICC/ESOMAR Code was therefore published in the following year (with revisions in 1986).

Subsequent changes in the marketing and social environment, new developments in marketing research methods and a great increase in international activities of all kinds, including legislation, led ESOMAR to prepare a new version of the International Code in 1994. This new version sets out as concisely as possible the basic ethical and business principles which govern the practice of marketing and social research. It specifies the rules which are to be followed in dealing with the general public and with the business community, including clients and other members of the profession. ∎

Sugging
The use of marketing research to deliberately disguise a sales effort.

Frugging
The use of marketing research to deliberately disguise fundraising activities.

Classic examples of breaches of these principles come in the guise of **sugging** and **frugging**. Sugging occurs when surveys are used to gain access to respondents to deliver a sales pitch or to generate sales by other means. Frugging occurs when surveys are used to raise funds for a charity or other cause under the guise of conducting research.

Internet and computer applications

According to most estimates, the global online population has grown to over 400 million and is destined to grow even further with the increasing use of palm-tops, WAP phones, interactive television and other access modes. In line with this development, online research is growing rapidly with triple-digit growth in the last few years and turnover estimated to exceed €320 million last year.[29] This online population cannot be treated as one homogeneous mass, as there are some clear differences between the European and US Internet markets. In Europe the Internet has become a feature of a new-media environment that includes other interactive media such as digital television, whereas in the US the research focus is on the Internet, which has grown faster than other new media. Internet growth in Europe has also been shaped by environmental factors unheard of in the States: data protection legislation and telecoms charging.[30]

The Internet has generated massive changes in how business is conducted and how different businesses, stakeholders and consumers communicate. Given the level of growth and turnover highlighted above, the Internet has generated a great array of opportunities and threats for the marketing research industry. There are few in the marketing research industry that would dispute the potential of the Internet. From initially being used in desk research, the Internet is being applied increasingly in both quantitative and qualitative techniques. As well as being a very effective tool for gathering data, the Internet, intranets and extranets are being used very effectively to disseminate information.[31] The key benefits of using the Internet when compared with traditional data collection methods lie in the speed, economy, control and elements of data quality, e.g. 'People are more likely to give an honest opinion on the Internet. A recent study showed that people are willing to disclose up to four times as much over the Net because they can't see each other – they are more focused and less concerned about being judged'.[32] The key limitations lie in sampling issues, technical problems and elements of data quality. Given the strength of the undoubted benefits of the Internet in conducting marketing research, the initial trickle of research agencies dabbling in Internet research has turned into a flood. This 'flood' has been precipitated by fears that agencies which continued to watch and wait could find themselves out of business.[33] So, as well as opportunities, the marketing research industry is faced by new threats, not just from traditional competitors but also from completely new competitors in the marketplace. The one factor that will intensify competition for the marketing research industry is the Internet. The Internet enables anybody to be a 'data collector', including marketing clients, i.e. research buyers and users. More generally, a growing market for automated research tools potentially makes it easy for anybody to become a researcher.

With the relative ease of becoming a 'researcher', the growth rate of the number of online surveys is generating concern and sparking debate. At risk are the levels of quality and ethical standards that have been nurtured and valued in the industry over many years. Millward Brown Intelliquest, for example, recently published figures estimating that up to 95% of marketing research in the European Community by 2005 could be conducted through the Internet.[34] Because it has become cheap and relatively straightforward to produce surveys through the Internet, those companies with an eye on the bottom line may be tempted to use organisations that do not apply the quality standards prevalent in traditional forms of research. It is a relatively simple task to put together an email questionnaire, send it out to addresses compiled from any number of newsgroups and wait for the response. Most likely in this scenario are results that have the potential to be grossly misleading.

Responding to the threats, it has been argued that the availability of data on the Web means that survival of the marketing research industry will depend more on what it can do with the information it collects and less with how it is collected.[35] The marketing research industry must therefore evolve its current offering, adapting to the demands of decision-makers that recognise the value of good research. The industry must understand the technology and use it effectively, but it must be careful not to lose sight of the human element of the accumulated knowledge and added value that good researchers can bring.[36] What must not be compromised are the levels of quality in data collection and analysis and the ethical standards that have been nurtured and valued in the industry over many years.

Given the importance of the Internet and developing technologies to the marketing research process and industry, key issues and debates in this subject will be continually addressed. Throughout this book we show how the stages of the marketing research process are facilitated by the Internet and other developing technologies and software. On the Companion Website to this text, demonstration versions of marketing research software will illustrate the use of computing technology in the process of questionnaire design, data entry, data analysis and reporting, and the application of geodemographic information systems.

Summary

Marketing research provides support to marketing decision-makers by helping to describe the nature and scope of customer groups, understand the nature of forces that shape the needs of customer groups and the marketer's ability to satisfy those groups, test individual and interactive controllable marketing variables, and monitor and reflect upon past successes and failures in marketing decisions. The overall purpose of marketing research is to assess information needs and provide the relevant information in a systematic and objective manner to improve marketing decision-making. Marketing research may be classified into problem identification research and problem-solving research. Information obtained using marketing research can become an integral part of a Marketing Information System (MkIS) and a Decision Support System (DSS). The contribution of marketing research to such systems illustrates that research has a role in developing marketing learning within organisations.

Marketing research may be conducted internally (by internal suppliers) or may be purchased from external suppliers. Full-service suppliers provide the entire range of marketing research services, from problem definition to report preparation and presentation. The services provided by these suppliers can be classified as syndicated services, standardised services or customised services. Limited-service suppliers specialise in one or a few phases of the marketing research project. Services offered by these suppliers can be classified as field services, coding and data entry, data analysis, analytical services or branded products.

The marketing research process consists of six broad steps that must be followed systematically. The process involves problem definition, research approach development, research design formulation, fieldwork or data collection, data preparation and analysis, and report preparation and presentation. Within these six broad steps are many iterations and routes that can be taken, reflecting the reality of practising marketing research.

Marketing research is not a panacea for all marketing problems. There are examples where marketing research has not adequately supported decision-makers. Many of the problems that arise from poor marketing research derive from poor communications

between decision-makers and researchers. International marketing research is much more complex than domestic research because the researcher must consider the environments prevailing in the international markets being researched. Research is founded upon the willing cooperation of the public and of business organisations. Ethical marketing research practices nurture that cooperation allowing a more professional approach and more accurate research information. Marketing research makes extensive use of the great opportunities afforded by the Internet. There are also many competitive threats to the marketing research industry that have been exacerbated by the Internet.

Questions

1 Describe the task of marketing research.

2 What decisions are made by marketing managers? How does marketing research help in supporting these decisions?

3 Define marketing research.

4 What problems are associated with using consumer databases in marketing research?

5 Describe one classification of marketing research.

6 What is a marketing information system and how does it differ from a decision support system?

7 Explain one way to classify marketing research suppliers and services.

8 What are syndicated services?

9 Describe the steps in the simple linear marketing research process.

10 Explain why there may be the need for iterations between stages of the marketing research process.

11 What arguments can be used by sceptics of marketing research?

12 What arguments would you use to defend investment in marketing research?

13 What factors fuel the growth of international marketing research?

14 Discuss the ethical issues in marketing research that relate to (a) the client, (b) the supplier, and (c) the respondent.

15 Summarise the nature of threats and opportunities that the Internet offers the marketing researcher.

Notes

1 Vangelder, P., 'Listening to the customer – research at Royal Ahold', *ESOMAR Newsbrief* 4 (April 1999), 14–15.

2 Johansen, T. and Herlin, K., 'Banner advertising – more than clicks', *Admap* (October 2001), 15–17.

3 Bradford, S., 'From Armani to AIDS: the hopes, the fears, of Europe's youth', *ResearchPlus* (November 1997), 4.

4 For the strategic role of marketing research, see Zabriskie, Noel B. and Huellmantel, Alan B., 'Marketing research as a strategic tool', *Long-Range Planning* 27 (February 1994), 107–18.

5 For relationships between information processing, marketing decisions and performance, see Glazer Rashi, and Weiss, Allen M., 'Marketing in turbulent environments: decision process and the time-sensitivity of information', *Journal of Marketing Research* 30 (November 1993), 509–21.

6 Malhotra, N.K., 'Shifting perspective on the shifting paradigm in marketing research', *Journal of the Academy of Marketing Science* 20 (Fall 1992), 379–87; and Perreault, William, 'The shifting paradigm in marketing research', *Journal of the Academy of Marketing Science* 20 (Fall 1992), 367–75.

7 Cervi, B. and Savage, M., 'Innovate to survive', *Research* (October 1999), 16–17.

8 Savage, M., 'Downstream danger', *Research* (May 2000), 25–7.

9 'Beer research shows overspill', *Marketing* (9 October 1997), 6.

10 'Research can be child's play', *Marketing* (10 May 2001), 35.

11 Turban, E., *Decision Support and Expert Systems* (New York: Macmillan, 1990); McCann, J., Tadlaoui, A. and Gallagher, J., 'Knowledge systems in merchandising: advertising designs', *Journal of Retailing* (Fall 1990), 257–77.

12 Mohanty, R.P. and Deshmukh, S.G., 'Evolution of a decision support system for human resource planning in a petroleum company', *International Journal of Production Economics* 51(3) (September 1997), 251–61.

13 Calmet, A.V., 'Geographic marketing in financial institutions: the experience of Banco Central Hispanoamericano', *Marketing and Research Today* (August 1996), 182–9.

14 ESOMAR Annual Study on the Market Research Industry 2000.

15 *Ibid.*

16 ESOMAR Press Release, Amsterdam, 7 September 1997.

17 ESOMAR Press release, October 2001.

18 'Supermarket sweep', *The Grocer* (13 July 1996), 28.

19 Bull, N. and Oxley, M., 'The search for focus – brand values across Europe', *Marketing and Research Today* (November 1996), 239.

20 Lury, G., 'Market research cannot cover for the "vision thing"', *Marketing* (9 November 2000), 34.

21 Gordon, W., 'Be creative to innovate', *Research* (January 2000), 23.

22 Lehmann, D.R., *Market Research and Analysis*, 3rd edn (Homewood, IL: Irwin, 1994), 14.

23 Sampson, P. and Standen, P., 'Predicting sales volume and market shares', in *New Product Development Research Contributions to Strategy Formulation, Idea Generation and Screening Product, Product Testing and Final Marketing*, ESOMAR (November 1983).

24 *Ibid.*

25 Muranka, T. and Rootes, N., *Doing a Dyson* (Dyson Appliances, 1996), 22.

26 Luck, D.J., Wales, H.C., Taylor, D.A. and Rubin, R.S., *Marketing Research*, 5th edn (Englewood Cliffs, NJ: Prentice Hall, 1978).

27 Chervi, B. and Savage, M., 'Innovate to survive', *Research* (October 1999), 16.

28 McElhatton, N., 'Crossing borders', *Research, Guide to Europe* (May 2000), 4–5.

29 ESOMAR Press Release, 'New ESOMAR guideline on Internet privacy policies and privacy statements', September 2001.

30 Savage, M., 'Netting the online gains', *Research* (November 1999), 27.

31 Johnson, A., 'Welcome to the wired world', *Research* (November 1999), 23.

32 Haig, M., 'Netting consumer opinion', *Marketing Week* (26 October 2000), 71.

33 Savage, M., 'Beware the net traps', *Research* (May 2000), 22.

34 West, R., 'Tools you can trust?', *Marketing Week* (12 October 2000), 69.

35 Savage, M., 'Downstream danger', *Research* (May 2000), 27.

36 Johnson, A., 'Welcome to the wired world', *Research* (November 1999), 25.

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Defining the marketing research problem and developing a research approach

Objectives

After reading this chapter, you should be able to:

1 understand the importance of, and the process used, in defining marketing research problems;

2 describe the tasks involved in problem definition;

3 discuss in detail the nature and various components of a research brief and a research proposal;

4 discuss the environmental factors affecting the definition of the research problem;

5 clarify the distinction between the marketing decision problem and the marketing research problem;

6 explain the structure of a well-defined marketing research problem, including the broad statement and the specific components;

7 understand the role of theory in the development and execution of applied marketing research;

8 acquire an appreciation of the complexity involved in defining the problem and developing a research approach in international marketing research;

9 understand the ethical issues and conflicts that arise in defining the problem and developing a research approach.

The correct diagnosis of marketing research problems cannot be compromised. Regardless of how well subsequent stages are conducted, the whole process may be worthless if the diagnosis is weak.

Overview

This chapter covers the first two of the six steps of the marketing research process described in Chapter 1: defining the marketing research problem and developing a research approach to tackle the problem. Defining the problem is the most important step, since only when a problem has been clearly and accurately identified can a research project be conducted properly. Defining the marketing research problem sets the course of the entire project. Regardless of how well a research plan is designed and subsequent stages are carried out, if the problem is not correctly diagnosed, research findings could be misleading or even dangerous. In this chapter, we allow the reader to appreciate the complexities involved in defining a problem, by identifying the factors to be considered and the tasks involved.

In practical terms, the means to communicate and facilitate the diagnosis of research problems is achieved through the preparation of a research brief and research proposal. The rationale and components of the research brief and research proposal are presented. We provide guidelines for appropriately defining the marketing research problem and avoiding common types of errors. We also discuss in detail the characteristics or factors influencing the research design and components of an approach to the problem: objective/theoretical framework, analytical models, research questions and hypotheses. The special considerations involved in defining the problem and developing a research approach in international marketing research are discussed. Finally, several ethical issues that arise at this stage of the marketing research process are considered.

We introduce our discussion with an example. The ABC Global Kids study illustrates that starting from broad research questions, a clear set of research questions can be developed. There are many difficulties in designing a research approach that will allow comparable findings from six countries in three continents and in questioning children. Without a precise diagnosis of marketing and research problems, such a complicated research approach would be doomed to failure.

example

Is there one global village for our future generation?[1]

The ABC Global Kids Study was conducted in the spring of 1996 among 2,400 children aged 7 to 12 years and their mothers, in China, Japan, France, Britain, Germany and the United States. The study was sponsored by multinational corporations such as Mars, Kodak and McDonald's. The key topic areas were as follows:

- Child wealth and spending patterns
- Influence on household purchases
- Media habits
- Technology ownership and usage
- Social issues and concerns
- Daily activities
- Food and beverages
- Toys and games
- Cartoon and movie awareness and attitudes
- Recreation.

The study was designed to quantify the size of business opportunities in given product categories and to provide fact-based direction on how to create an exciting child product portfolio and marketing programme.

While uniformity of survey, method and instrument were maintained for cross-country comparison, adaptations were sought locally to capture unique retail and cultural characteristics. A pictorial response scale was used when interviewing children on emotions and preferences. Product usage frequency was asked of mothers, instead of children.

The survey revealed considerable spending power for children aged 7–12, as well as active participation from children in making family purchase decisions in a number of product categories. It was apparent that, worldwide, children basically have many common dreams and aspirations. ■

Importance of defining the problem

Problem definition
A broad statement of the general problem and identification of the specific components of the marketing research problem.

Although each step in a marketing research project is important, problem definition is the most important step. As mentioned in Chapter 1, for the purpose of marketing research, problems and opportunities are treated interchangeably. **Problem definition** involves stating the general problem and identifying the specific components of the marketing research problem. Only when the marketing research problem has been clearly defined can research be designed and conducted properly. 'Of all the tasks in a marketing research project, none is more vital to the ultimate fulfilment of a client's needs than an accurate and adequate definition of the research problem. All the effort, time, and money spent from this point on will be wasted if the problem is misunderstood and ill-defined.'[2] An analogy to this is the medical doctor prescribing treatment after a cursory examination of a patient – the medicine may be even more dangerous than the condition it is supposed to cure!

The importance of clearly identifying and defining the research problem cannot be overstated. The foundation of defining a research problem is the communication that develops between marketing decision-makers and marketing researchers. In some form or another the marketing decision-makers must communicate what they see as being the problems they face and what research support they need. This communica-

Another big fish that wriggles off the hook.

tion usually comes in the form of a research brief. The marketing researcher responds to the research brief with a research proposal, which encapsulates their vision of a practical solution to the set research problem. The following example illustrates that a research brief may not always be particularly well thought out. The marketing researcher is expected to develop the brief into a research proposal and in doing so has a vital role to play in the diagnosis of research problems.

example

How to bait the interview hook for those Top 1000 big fish[3]

The groans from researchers when another brief arrives asking for 100 or 200 interviews with Chief Executive Officers (CEOs) or equivalents within the *Times* Top 1000 companies typify the attitude generated by business-to-business marketers' constant demand to reach this audience. When the research brief arrives, it is certainly worth examining whether, practically, what is requested can actually be done. The research proposal developed must reflect the practicalities of questioning managers who are constantly bombarded with requests to respond to research questions. The number of interviews, the time scale, the nature of questions and the structure of the sample all need to be taken into account. For example, is it really worth undertaking 200 interviews within any single European country? If we were limited to one per organisation, we would be interviewing to strike rates of between 1 in 2.5 and 1 in 5. If the research targets companies throughout Europe, individual countries such as Britain and France may have few large companies compared with the USA, while Italy has a limited number of very large companies and a great many smaller ones. In actually reaching the target audience, a number of issues need to be taken into account. International business-to-business research with senior business audiences brings with it not only the particular difficulties of reaching them but also the need to understand both country and cultural issues that impact on the research. Telephone interviews (even if possible) are considered inappropriate for these audiences in many Far East and Middle East markets, especially South Korea and Japan, while in Singapore and Hong Kong, provided the interviews are not too long (over 15 minutes), telephone is fine. ■

The marketing research brief

Research brief
A document produced by the users of research findings or the buyers of a piece of marketing research. The brief is used to communicate the perceived requirements of a marketing research project.

The marketing **research brief** is a document produced by the users of research findings or the buyers of a piece of marketing research. The brief may be used to communicate the perceived requirements of a marketing research project to external agencies or internally within an organisation to marketing research professionals. It should act as the first step for decision-makers to express the nature of a marketing and research problem as they see it. This first step is vital in developing an agreement of an appropriate research approach. As a first step, *the marketing research brief should not be carved in tablets of stone!*

It has been contended that the greatest form of potential error in marketing research lies in the initial relationship between marketing decision-makers and marketing researchers.[4] In developing a sound initial relationship, the research brief plays a vital role. Without some formal method of communicating the nature of a marketing problem, there is great potential for ambiguities, illogical actions (by both parties), misunderstandings and even forgetfulness.

The purpose of a written marketing research brief may be summarised as:

■ It makes the initiator of the brief more certain of how the information to be collected will support decision-making.

■ It ensures an amount of agreement or cohesion among all parties who may benefit from the research findings.

■ It helps both the marketer and the researcher to plan and administer the research programme.

- It helps to reduce disputes that can occur when the gaps in marketers' knowledge are not 'filled' as intended.
- It can form the basis for negotiation with a variety of research organisations.

In all, the research brief saves resources in time and money by helping to ensure that the nature of the problem or opportunity under investigation has been thought through.

Components of the marketing research brief

The rationale for a marketing research brief may seem logical, but actually generating a brief from marketing decision-makers can be extremely difficult. These difficulties will be tackled later in this chapter. The following format for a research brief has advantages from the perspective of the decision-maker and the researcher. First, it does not demand that decision-makers have a great deal of technical knowledge about research. Their focus can remain upon the gaps in their knowledge, the nature of support they need, not the technicalities of how data are to be collected and analysed. Second, it allows the researchers the opportunity to demonstrate their creative abilities. Using their experiences from problems faced by other decision-makers, perhaps from a great variety of contexts and industries, researchers have the possibility of examining the marketing and research problem from many different perspectives. They can create, develop and adapt a research design to the research problem that supports the marketing decision-maker within clear time and cost parameters.

1 *Background information.* The background serves to put research objectives into context, helping the researcher to understand *why* certain research objectives are being pursued. The marketer would detail what they see as being the main events that have caused or contributed to the problem under study. Such a background gives a framework for the researcher to investigate other potential events, contributory factors or causes.

2 *Objectives.* The first part of this section would detail which marketing decisions are to be completed once the research has been undertaken. This requires the decision-maker to explain what they see as the focus of the decisions they plan to make. They then go on to explain what gap(s) they see in their knowledge. Those gaps create the focus to planned research activities and set the research objectives. The formulation of the marketing objectives can encompass two areas: organisational

Figure 2.1
Components of the marketing research brief

Background information

→

Objectives (a) marketing
(b) research

→

Target to research

→

Who is to use findings
(a) analysis
(b) format

→

Constraints

→

Administration

objectives and personal objectives of the decision-maker. For a research project to be successful, it must serve the objectives of the organisation and of the decision-maker. For the researcher this may not be explicit or obvious to discern. It may take some time working with a decision-maker or a particular organisation to see potential conflicts in organisational and personal objectives. The problem faced by researchers is that decision-makers may not formulate marketing objectives clearly. Rather, it is likely that objectives tend to be stated in terms that have no operational significance, such as 'to improve corporate image'. Ultimately this does not matter, as this 'first-step' brief offers the opportunity for the researcher to draw out and develop a much clearer vision of marketing and research objectives. Drawing out and developing decision-makers' perspectives of objectives, even if they have no operational significance, helps the process of developing a common understanding of what the decision-maker is trying to achieve.

3   *Target to research.* Any marketing research project will measure or observe a target group of individuals. These may be distinct groups of consumers, channel members such as retailers or competitors, or company employees. In this section, details of the characteristics of the target group(s) can help in many research design decisions. These cover areas of identification, gaining access to conduct research, understanding which techniques are appropriate to measure or understand these individuals, and the best environment or context in which to conduct research.

4   *Who is to use the findings.* This section would outline brief details of the decision-makers who will use the research findings. For example, certain decision-makers may be entrepreneurial and introspective, looking for short-term tactical advantages. Presenting research findings that make tactical advantages apparent would be the best way to communicate to such managers. Managers with a background and training in statistics may expect results to be analysed and presented in a particular manner to have any credibility. Other managers may not have such training or even be distrustful of statistical analyses and seek a more qualitative interpretation. These issues have an impact upon the nature and extent of analysis conducted upon the data collected and the style and format in which research findings will be presented.

5   *Constraints.* The main limitation to marketing researchers carrying out what they may perceive as being the correct way to research a problem is the time and money that a marketer can afford. Proposing a large-scale project that would cost €200,000 when only €50,000 has been budgeted obviously will not meet management approval. In many instances, the scope of the marketing research problem may have to be reduced to accommodate budget constraints. With knowledge of time and cost constraints, the researcher can develop a research design to suit these needs. The researcher may also demonstrate other courses of action that could demand greater amounts of money or time, but could have clear benefits that the marketer may be unaware of. Other constraints, such as those imposed by the client firm's personnel, organisational structure and culture, or decision-making styles, should be identified to determine the scope of the research project. Yet, constraints should not be allowed to diminish the value of the research to the decision-maker or to compromise the integrity of the research process. In instances where the resources are too limited to allow a project of sufficient quality, the firm should be advised not to undertake formal marketing research. Thus, it becomes necessary to identify resources and constraints, a task that can be better understood when examined in the light of the objectives of the organisation and the decision-maker.

6   *Administrative considerations.* These would lay out administrative details in completing the research project. Examples could be the expected delivery of interim reports, contacts in an organisation who may be able to help supply further information, or reference to sources of materials and individuals that are needed to successfully complete the research.

With a formal marketing research brief and perhaps preliminary discussions with the organisation that is to commission the research, the marketing researcher has the necessary material to develop a research proposal.

The above outline of a marketing research brief is a format that allows the researcher to display their wide range of expertise and creative abilities. Other formats of research briefs may be far more detailed and explicit in terms of specifying methods of measurement and analysis. Such research briefs act as tender documents. In many instances, however, the marketing researcher does not enjoy the luxury of a written research brief. The marketing decision-maker may outline their ideas in an oral manner, perhaps on an informal basis. This can happen if the decision-maker is not aware of the personal benefits of producing a written research brief detailed above. They may see the brief as a time-consuming process that really is the job of the researcher. If the marketing researcher is faced with an oral brief, they can use the proposed brief outline above as a guideline to the issues they should elicit in informal discussions in order to develop an effective proposal.

The marketing research proposal

Research proposal
The official layout of the planned marketing research activity.

In response to a research brief, the marketing researcher will develop a research plan (covered in detail in Chapter 3) and will develop a **research proposal** to communicate this plan. The marketing research proposal contains the essence of the project and, in its final format, serves as a contract between the researcher and management. The research proposal covers all phases of the marketing research process. It allows the researcher to present their interpretation of the problems faced by management and to be creative in developing a research solution that will effectively support decision-makers. Although the format of a research proposal may vary considerably, most proposals address all the steps of the marketing research process and contain the elements shown in Figure 2.2.

1 *Executive summary.* The proposal should begin with a summary of the major points from each of the other sections, presenting an overview of the entire proposal.
2 *Background.* The researcher would be expected to have developed ideas beyond those presented in the brief 'background'. Other potential causes of the problems faced or alternative interpretations of the factors that shape the background in an environmental context should be presented.
3 *Problem definition.* Again, if necessary, the researcher may go beyond the problem definition presented in the brief. If the researcher sees potential to add value for the marketer through alternative diagnoses of the problem presented in the brief, then these should be shown. If the researcher sees a problem in the brief that is ambiguous or unattainable, again alternative diagnoses should be presented. From this section, the marketer's gaps in knowledge should be apparent.
4 *Research objectives.* These may be presented in the form of clear hypotheses that may be tested. They may also cover broader areas in terms of 'research questions' that are to be explored rather than formally measured in a conclusive manner.
5 *Research design.* The research design to be adopted, in broad terms classified as exploratory, descriptive or causal, should be specified. Beyond such a broad classification should be details of the individual techniques that will be adopted and how they will unfold and connect to each other. This means that the reader will clearly see methods of collecting the desired data, justification for these methods, and a sampling plan to include details of sample size(s). This applies to both quantitative and qualitative approaches.

Figure 2.2
Components of the marketing research proposal

6 *Fieldwork/data collection*. The proposal should discuss how the data will be collected and who will collect it. If the fieldwork is to be subcontracted to another supplier, this should be stated. Control mechanisms to ensure the quality of data collected should be described.

7 *Data analysis*. This should describe the kind of data analysis that will be conducted, e.g. content analysis, simple cross-tabulations, univariate analysis or multivariate analysis. If software packages are to be used in these analyses, they should be specified, as they will be indicative of the potential analyses that can be conducted. There should be further description of the extent to which the results will be interpreted in light of the set marketing objectives, beyond the specified analysis techniques.

8 *Reporting*. The proposal should specify the nature of any intermediate reports to be presented, what will be the form of the final report, and whether an oral presentation of the results will be made.

9 *Cost and timetable*. The cost of the project and a time schedule, broken down by phases, should be presented. A critical path method chart might be included. In large projects, a payment schedule is also worked out in advance.

10 *Research organisation and key researchers working on the project*. When an organisation is working with researchers for the first time, some idea of past research projects and clients should be displayed. This can help the marketer to trust the researcher in problem diagnosis, research design and implementation (e.g. how credible the researchers may be seen by the individuals they are to research and how this may affect respondent openness and honesty), and interpretation of the findings.

11 *Appendices*. Any statistical or other information of interest to only a few people should be contained in appendices.

12 *Agreement*. All parties concerned in fulfilling the research plan should sign and date their agreement to the proposal.

Preparing a research proposal has several advantages. It ensures that the researcher and management agree about the nature of the project, and it helps sell the project to a wider array of decision-makers who may contribute to and benefit from the research findings. As preparation of the proposal entails planning, it helps the researcher conceptualise and execute the marketing research project. The following example illustrates the language problems in writing a pan-European research proposal. In this example, a problem lies with English speakers wanting to negotiate and work with other English speakers and having little experience of the countries in which they will undertake research.

e x a m p l e

Some dos and don'ts of pan-European research[5]

Many European studies are commissioned by English speakers who want to work with English speakers. They want to agree questionnaires in English, and they want reports and tables in English. And so they come to Britain.

In addition, it is simpler for someone commissioning international research to accept an English proposal written by a first-language English speaker. Even the best European – used here in its not-British sense – English speaker finds it hard to write a creative proposal in good idiomatic English. *But it is creativity that sells proposals*.

A problem in writing proposals for research in Europe is often lack of familiarity with European habits and behaviour. Most British researchers' experience of Europe is derived from holidays backed up (in all too few cases) by limited work experience in one or more countries. Faced by a proposal on TV viewing, on grocery purchasing, on holiday habits or on headache remedies, few British researchers have any notion of how Italians, Spaniards or Germans treat these subjects. Hence the first step in writing an international proposal on an unfamiliar subject is secondary research. This may also involve calling a dozen homes in the relevant country, and chatting on the phone for an hour or two with locals. Soon you start to identify how Italians, Spaniards and Germans differ from the British and what these main differences are. Designing a piece of European research in an unfamiliar area without taking this first simple step is like taking up power boating without learning to swim. It is a route to disaster, not maybe with the proposal, but with the research when commissioned. ■

The process of defining the problem and developing a research approach

By formally developing and exchanging a marketing research brief and research proposal, the marketing decision-maker and the marketing researcher utilise their distinctive skills. They ensure that the marketing problem and research problems have been correctly defined and an appropriate research approach is developed. The research brief and the research proposal are the formal documents that ensure each party is clear about the nature and scope of the research task. These documents allow decision-makers and researchers to formally present their perspective of the task in hand. The following section details the process that needs to be undertaken in order to produce these documents. The detail of defining the nature of problems and developing an appropriate research approach to the point of creating a research design is shown in Figure 2.3.

The tasks involved in problem definition consist of discussions with decision-makers, qualitative interviews with industry experts and other knowledgeable individuals, and analysis of readily available secondary data. These tasks help the researcher to understand the background of the problem by analysing the environmental context. Certain essential environmental factors bearing on the problem should be evaluated. An understanding of the environmental context facilitates the identification of the marketing decision problem. Then, the marketing decision
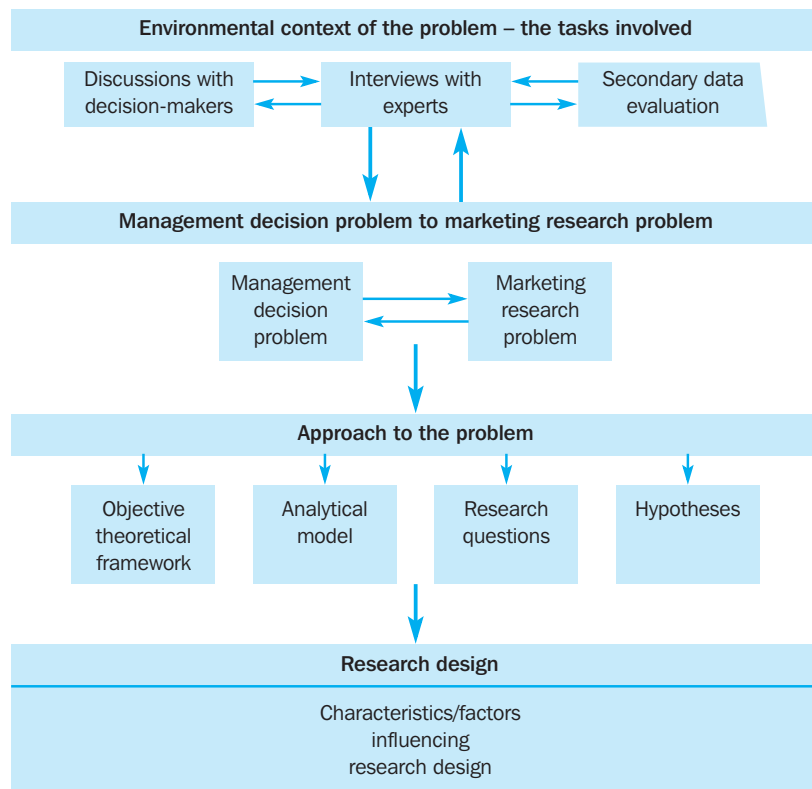
Figure 2.3
The process of defining the problem and developing an approach

problem is translated into a marketing research problem. Based on the definition of the marketing research problem, an approach to the problem is established and an appropriate research design is developed. The components of the approach may consist of an objective/theoretical framework, analytical models, research questions and hypotheses. Further explanation of the problem definition process begins with a discussion of the tasks involved.

Environmental context of the problem

The tasks involved in understanding the environmental context of the marketing and research problem can include discussions with decision-makers, qualitative interviews with industry experts, and secondary data analysis. The purposes of these tasks are to develop an understanding of forces that may affect the nature of decision-makers' problems and related research problems.

Discussions with decision-makers

Discussions with the decision-makers beyond the formal presentation of a research brief and research proposal are usually vital. The decision-maker needs to understand the capabilities and limitations of research.[6] Research provides information relevant to management decisions, but it cannot provide solutions, because solutions require managerial creativity and judgement. Conversely, the researcher needs to understand the nature of the decision the managers face – the marketing problem, and what they hope to learn from the research.

To identify the marketing problem, the researcher must possess considerable skill in interacting with the decision-maker. Several factors may complicate this interaction. Access to the decision-maker may be difficult, and some organisations have complicated protocols for access to top executives. The organisational status of the researcher or the research department may make it difficult to reach the key decision-maker in

Gaining access to key decision-makers can be easier said than done.

the early stages of the project. Finally, there may be more than one key decision-maker, and meeting collectively or individually may be difficult. All of the problems make it difficult to develop a research brief. Despite these problems, though, it is necessary that the researcher interact directly with the key decision-makers.[7]

Problem audit
A comprehensive examination of a marketing problem to understand its origin and nature.

A **problem audit** provides a useful framework to develop ideas from a brief, allowing the researcher to interact with the decision-maker and identify the underlying causes of the problem. A problem audit, like any other type of audit, is a comprehensive examination of a marketing problem with the purpose of understanding its origin and nature.[8] A problem audit involves discussions with the decision-maker on the following issues:

1 The events that led to the decision that action is needed, or the history of the problem.
2 The corporate culture as it relates to decision-making.[9] For example, in some firms, the decision-making process is dominant; in others, the personality of the decision-maker is more important.
3 The alternative courses of action available to the decision-maker. The set of alternatives may be incomplete at this stage, and qualitative research may be needed to identify the more innovative courses of action.
4 The criteria that will be used to evaluate the alternative courses of action. For example, new product offerings might be evaluated based on sales, market share, profitability, or return on investment.
5 What the decision-maker perceives to be gaps in their knowledge.
6 The manner in which the decision-maker will use each item of information in making the decision.

It may be necessary to perform a problem audit, because the decision-maker may have only a vague idea of what the problem is. For example, the decision-maker may know that the firm is losing market share but may not know why; decision-makers may tend to focus on symptoms rather than on causes. Inability to meet sales forecasts, loss of market share and decline in profits are all symptoms. The researcher should treat the underlying causes, not merely address the symptoms. For example, loss of market share may be caused by much better advertising campaigns by the

competition, inadequate distribution of the company's products, or any number of other factors. Only when the underlying causes are identified can the problem be successfully addressed.

A problem audit, which involves extensive interaction between the decision-maker and the researcher, can greatly facilitate problem definition by determining the underlying causes. The interaction between the researcher and the decision-maker is facilitated when one or more people in the client organisation serve to liaise and form a team with the marketing researcher. To be fruitful, the interaction between the decision-maker and the researcher can be characterised by the following:

1 *Communication.* A free exchange of ideas between the decision-maker and the researcher is essential.
2 *Cooperation.* Marketing research is a team project in which both parties (decision-maker and researcher) must cooperate, from problem diagnosis through to the interpretation and presentation of findings.
3 *Confidence.* Mutual trust of each other's distinct skills and contribution should underlie the interaction between the decision-maker and the researcher.
4 *Candour.* There should not be any hidden agendas, and an attitude of openness should prevail.
5 *Closeness.* An understanding of each other's problems should result in a closeness that should characterise the relationship between the decision-maker and the researcher.
6 *Continuity.* The decision-maker and the researcher must interact continually rather than sporadically.
7 *Creativity.* The interaction between the decision-maker and the researcher should be creative rather than formulaic. Though the research process may be laid out in 'easy-to-follow' steps, in reality great amounts of creativity are needed at every stage.

Interviews with industry experts

In addition to discussions with decision-makers, qualitative interviews with industry experts, individuals knowledgeable about the firm and the industry can help in diagnosing the nature of the marketing and research problem.[10] These experts may be found both inside and outside the firm. Typically, expert information is obtained by unstructured personal interviews, without a formal questionnaire. It is helpful, however, to prepare a list of topics to be covered during the interview. The order in which these topics are covered and the questions to ask should not be predetermined. Instead, they should be decided as the interview progresses, which allows greater flexibility in capturing the insights of the experts (see Chapter 8 for full details of depth interviewing techniques). The list of topics to cover and the type of expert sought should evolve as the researcher becomes more attuned to the nature of the marketing problem. The purpose of interviewing experts is to explore ideas, make new connections between ideas, and create new perspectives in defining the marketing research problem. If the technique works well, and an amount of trust and rapport is developed, the potential to generate and test ideas can be immense. Experts may have other contacts that the researcher may not be aware of or may not be able to get access to. They may also have secondary data which, again, the researcher may not be aware of or have access to. Unfortunately, two potential difficulties may arise when seeking advice from experts:

1 Some individuals who claim to be knowledgeable and are eager to participate may not really possess expertise.
2 It may be difficult to locate and obtain help from experts who are outside the client organisation, i.e. access to these individuals may be problematic.

For these reasons, interviews with experts are more useful in conducting marketing research for industrial firms and for products of a technical nature, where it is relatively easy to identify and approach the experts. This method is also helpful in situations where little information is available from other sources, as in the case of radically new products. Experts can provide valuable insights in modifying or repositioning existing products, as illustrated in the following example by the development of brand values at Visa.

Visa to travel [11]

Visa ran a major brand-building campaign spanning Europe. The aim of the campaign was to create the same brand values in different countries. Although the benefits of uniform branding are obvious, many companies have found that targeting a number of countries with the same campaign is fraught with danger. However, with research that revealed perceptions of Visa and cash to be consistent across Europe that 'plastic is better than cash', the message was simple to translate – simple if one is aware of cultural sensibilities and the relationship between countries, the company and the target audience. With industry expertise emphasising the appeal of locally produced campaigns, Visa have succeeded. As an example, they have been running a separate campaign for the Electron debit card in Spain, Portugal and Italy for some time. It has won many awards in Spain where they think it is a Spanish campaign, but then they think the ad is locally produced in Italy and Portugal as well. ■

Initial secondary data analyses

Secondary data
Data collected for some purpose other than the problem at hand.

Primary data
Data originated by the researcher specifically to address the research problem.

Secondary data collection and analysis will be addressed in detail in Chapters 4 and 5. A brief introduction here will demonstrate the worth of secondary data at the stage of problem diagnosis. **Secondary data** are data collected for some purpose other than the problem at hand. **Primary data**, on the other hand, are originated by the researcher for the specific purpose of addressing the research problem. Secondary data include data generated within an organisation, information made available by business and government sources, commercial marketing research firms, and computerised databases. Secondary data are an economical and quick source of background information. Analysis of available secondary data is an essential step in the problem definition process: primary data should not be collected until the available secondary data have been fully analysed. Past information and forecasts of trends with respect to sales, market share, profitability, technology, population, demographics and lifestyle can help the researcher to understand the underlying marketing research problem. Where appropriate, this kind of analysis should be carried out at the industry and firm levels. For example, if a firm's sales have decreased but industry sales have increased, the problems will be very different from if the industry sales have also decreased. In the former case, the problems are likely to be specific to the firm. Past information and forecasts can be valuable in uncovering potential opportunities and problems.

Marketing decision problem
The problem confronting the marketing decision-maker, which asks what the decision-maker needs to do.

Marketing research problem
A problem that entails determining what information is needed and how it can be obtained in the most feasible way.

Marketing decision problem and marketing research problem

The **marketing decision problem** asks what the decision-maker needs to do, whereas the **marketing research problem** asks what information is needed and how it can best be obtained. [12] The marketing decision problem is action oriented. It is concerned with the possible actions the decision-maker can take. How should the loss of market share be arrested? Should the market be segmented differently? Should a new product be introduced? Should the promotional budget be increased?

In contrast, the marketing research problem is information oriented. It involves determining what information is needed and how that information can be obtained effectively and efficiently. Consider, for example, the loss of market share for a particular product line. The decision-maker's problem is how to recover this loss. Alternative courses of action can include modifying existing products, introducing new products, changing other elements in the marketing mix, and segmenting the market. Suppose that the decision-maker and the researcher believe that the problem is caused by inappropriate segmentation of the market and wanted research to provide information on this issue. The research problem would then become the identification and evaluation of an alternative basis for segmenting the market. Note that this process requires much interaction, in the sense that both parties critically evaluate, develop and defend each other's ideas to clarify the nature of decision and research problems, and to ensure there is a clear and logical connection between them. The GlobalCash Project illustrates further the distinction between the marketing decision problem and the marketing research problem as well as the interactive nature of the problem definition process.

e x a m p l e
GlobalCash Project

Defining the problem

Bank X: We are experiencing a loss of market share in Ireland in corporate banking.
Researcher: Is it just Ireland?
Bank X: No, but as we conduct the majority of our business there, the loss is causing us the greatest amount of concern.
Researcher: Why do you think you are losing market share?
Bank X: We wish we knew!
Researcher: How are your competitors coping?
Bank X: We suspect that other Irish banks are also suffering, but that the multinational banks are capturing market share.
Researcher: How do your customers feel about the quality of services you deliver?
Bank X: We recently attained our ISO 9000 for service quality, which we are proud of!
Researcher: But how does your service delivery compare with your competitors?

After a series of discussions with key decision-makers, analysis of secondary data within the bank and from other sources, the problem was identified as follows:

- *Marketing decision problem*
  What should be done to arrest the decline in market share of Bank X?

- *Marketing research problem*
  Determine the relative strengths and weaknesses of Bank X, vis-à-vis other major competitors in Ireland. This would be done with respect to factors that influence a company in its choice of a bank to handle its transactions.

The GlobalCash Project shows the interactive nature of the process to identify the marketing decision problem and the marketing problem. The following examples further distinguish between the marketing decision problem and the marketing research problem. ◼

Marketing decision problem	Marketing research problem
Should a new product be introduced?	To determine consumer preferences and purchase intentions for the proposed new product.
Should the advertising campaign be changed?	To determine the effectiveness of the current advertising campaign.
Should the price of the brand be increased?	To determine the price elasticity of demand and the impact on sales and profits of various levels of price changes.

Defining the marketing research problem

The general rule to be followed in defining the research problem is that the definition should:

1 allow the researcher to obtain all the information needed to address the marketing decision problem;
2 guide the researcher in proceeding with the project.

Researchers make two common errors in problem definition. The first arises when the research problem is defined too broadly. A broad definition does not provide clear guidelines for the subsequent steps involved in the project. Some examples of excessively broad marketing research problem definitions are: developing a marketing strategy for a brand, improving the competitive position of the firm, or improving the company's image. These are not specific enough to suggest an approach to the problem or a research design.

The second type of error is just the opposite: the marketing research problem is defined too narrowly. A narrow focus may preclude consideration of some courses of action, particularly those that are innovative and not obvious. It may also prevent the researcher from addressing important components of the marketing decision problem. For example, in a project conducted for a consumer products firm, the marketing problem was how to respond to a price cut initiated by a competitor. The alternative courses of action initially identified by the firm's research staff were to:

1 decrease the price of the firm's brand to match the competitor's price cut;
2 maintain price but increase advertising heavily; or
3 decrease the price somewhat, without matching the competitor's price, and moderately increase advertising.

None of these alternatives seemed promising. When outside marketing research experts were brought in, the problem was redefined as improving the market share and profitability of the product line. Qualitative research indicated that in blind tests consumers could not differentiate products offered under different brand names. Furthermore, consumers relied on price as an indicator of product quality. These findings led to a creative alternative: increase the price of the existing brand and introduce two new brands – one priced to match the competitor and the other priced to undercut it.

The likelihood of committing either error of problem definition can be reduced by stating the marketing research problem in broad, general terms and identifying its specific components (see Figure 2.4). The **broad statement of the problem** provides perspective and acts as a safeguard against committing the second type of error. The **specific components of the problem** focus on the key aspects and provide clear guidelines on how to proceed further, and act as a safeguard against committing the first type of error. Examples of appropriate marketing research problem definitions are provided in the following two examples.

Broad statement of the problem
The initial statement of the marketing research problem that provides an appropriate perspective on the problem.

Specific components of the problem
The second part of the marketing research problem definition that focuses on the key aspects of the problem and provides clear guidelines on how to proceed further.
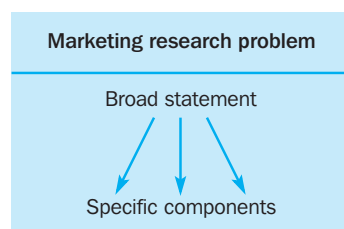
Figure 2.4
Proper definition of the marketing research problem

Marketing research problem

Broad statement

Specific components

Net has yet to prove its value [13]

As the stampede to set up Websites continues, those companies without a clear long-term strategy are looking increasingly rudderless as they search for ways to capitalise on their investment. A recent survey by Gallup and the *Wall Street Journal* of 603 board-level executives in Britain, France and Germany has revealed that Europe's businesses continue to embrace the Internet on the somewhat vague grounds that it is in the interests of their 'competitiveness'. Yet, they remain unsure about how, specifically, to use the medium for generating profits in the longer term.

The notable exception has been the financial services sector. Exploratory secondary data gathering and analyses revealed the opportunities that have been grasped by this sector. It also revealed the strategies adopted by financial companies. To develop these ideas into relevant actions, case histories of exemplary financial companies were built, key decision-makers targeted and questions focused upon technology, design and financial returns were developed.

Specifically the following issues could be addressed (as examples of many) for a company embarking on developing a website:

1 To what extent is the Net used by competitors?
2 What specific information types and services do they offer?
3 How can the Net generate knowledge of customers?
4 How can the Net add to marketing communication strategies?
5 How can the Net help to generate actual sales?
6 What types of existing or potential customers could use our site? ■

Problem definition in the GlobalCash Project followed the following pattern.

Problem definition

In the GlobalCash Project, the marketing research problem was to determine the relative strengths and weaknesses of Bank X, vis-à-vis other major competitors in Ireland. This would be done with respect to factors that influence a company in its choice of a bank to handle its transactions. Specifically, research provided information on the following questions:

1 What criteria do companies use when choosing a bank for the first time?
2 What criteria do companies use to allocate business between existing banks?
3 Which banks are used for domestic business and pan-European business?
4 How do those banks fare on a range of service quality delivery criteria?
5 What is the market share of Bank X and its competitors in Ireland?
6 What is the profile of companies that use Bank X? Does it differ from the profile of customers from competing banks?

Once the marketing research problem has been broadly stated and its specific components identified, the researcher is in a position to develop a suitable approach. ■

Components of the research approach

Once the marketing decision-maker and researcher have clarified the decision problem and established the research problem they face, it has to be decided how to approach this problem. The research problem may be very clear in the sense that there are strong established theories of what should be measured and how to conduct the measurements. Conversely, the research problem may lack theoretical foundation, with the researcher trying to cope with a broad set of issues that have not been sufficiently researched beforehand and unable to trust existing theories. How the researcher perceives the research problem affects the **paradigm** they will adopt in either an implicit or explicit manner. The researcher's adopted paradigm is built upon a set of assumptions. These assumptions consist of 'agreed upon' knowledge, criteria

Paradigm
A set of assumptions consisting of agreed-upon knowledge, criteria of judgement, problem fields and ways to consider them.

of judgement, problem fields and ways to consider them[14] (these factors will be developed further in Chapter 6). What is 'agreed upon' refers to how strong the theories are in defining and encapsulating the issues that make up a research problem. Bringing together the 'agreed upon' knowledge, criteria of judgement, problem fields and ways to consider them can be undertaken by considering the objective/theoretical framework, analytical models, research questions and hypotheses. Each of these components is discussed in the following sections. Collectively they may be considered to be the 'approach' that a researcher will take.

Objective/theoretical framework

Theory
A conceptual scheme based on foundational statements, or axioms, that are assumed to be true.

Objective evidence
Perceived to be unbiased evidence, supported by empirical findings.

In general, researchers should aim to base their investigations upon objective evidence, supported by **theory**. A theory is a conceptual scheme based on foundational statements called axioms that are assumed to be true. **Objective evidence** is gathered by compiling relevant findings from secondary sources. Likewise, an appropriate theory to guide the research might be identified by reviewing academic literature contained in books, journals and monographs. The researcher should rely on theory to determine which variables should be investigated. Past research on theory development and testing can provide important guidelines on determining dependent variables (variables that depend on the values of other variables) and independent variables (variables whose values affect the values of other variables). Furthermore, theoretical considerations provide information on how the variables should be operationalised and measured, as well as how the research design and sample should be selected. A theory also serves as a foundation on which the researcher can organise and interpret the findings: 'nothing is so practical as a good theory.'[15] Conversely, by neglecting theory, the researcher increases the likelihood that they will fail to understand the data obtained or be unable to interpret and integrate the findings of the project with findings obtained by others. The role of theory in the various phases of an applied marketing research project is summarised in Table 2.1.

Theory also plays a vital role in influencing the research procedures adopted in basic research. Applying a theory to a marketing research problem requires creativity on the part of the researcher, however. A theory may not specify adequately how its abstract constructs (variables) can be embodied in a real-world phenomenon. Moreover, theories are incomplete; they deal with only a subset of variables that exist

Table 2.1 The role of theory in applied marketing research

Research task	Role of theory
Conceptualising and identifying key variables	Provides a conceptual foundation and understanding of the basic processes underlying the problem situation. These processes will suggest key dependent and independent variables.
Operationalising key variables	Provides guidance for the practical means to measure or encapsulate the concepts or key variables identified.
Selecting a research design	Causal or associative relationships suggested by the theory may indicate whether a causal, descriptive or exploratory research design should be adopted (see Chapter 3).
Selecting a sample	Helps in defining the nature of a population, characteristics that may be used to stratify populations or to validate samples (see Chapter 14).
Analysing and interpreting data	The theoretical framework and the models, research questions and hypotheses based on it guide the selection of a data analysis strategy and the interpretation of results (see Chapter 17).
Integrating findings	The findings obtained in the research project can be interpreted in the light of previous research and integrated with the existing body of knowledge.

in the real world. Hence, the researcher must also identify and examine other variables that have yet to be published as theories. This may involve the researcher developing 'grounded theory', which will be explained and developed in Chapter 6.

Analytical model

Analytical model
An explicit specification of a set of variables and their interrelationships designed to represent some real system or process in whole or in part.

Verbal models
Analytical models that provide a written representation of the relationships between variables.

Graphical models
Analytical models that provide a visual picture of the relationships between variables.

Mathematical models
Analytical models that explicitly describe the relationships between variables, usually in equation form.

An **analytical model** is a set of variables and their interrelationships designed to represent, in whole or in part, some real system or process. Models can have many different forms. The most common are verbal, graphical and mathematical structures. In **verbal models**, the variables and their relationships are stated in prose form. Such models may be mere restatements of the main tenets of a theory. **Graphical models** are visual. They are used to isolate variables and to suggest directions of relationships but are not designed to provide numerical results. They are logical, preliminary steps to developing mathematical models.[16] **Mathematical models** explicitly specify the relationships among variables, usually in equation form.[17] These models can be used as guides for formulating the research design and have the advantage of being amenable to manipulation.[18] The different models are illustrated in the context of the GlobalCash Project.

As can be seen from this example, the verbal, graphical and mathematical models depict the same phenomenon or theoretical framework in different ways. The phenomenon of 'threshold effect', stated verbally, is represented for clarity through a figure (graphical model) and is put in equation form (mathematical model) for ease of statistical estimation and testing. Graphical models are particularly helpful in clarifying the concept or approach to the problem.

The verbal, graphical and mathematical models complement each other and help the researcher identify relevant research questions and hypotheses.

---

e x a m p l e
GlobalCash Project

Model building

Verbal model

A finance manager first becomes aware of a bank's ability to perform a particular cash management service. That manager then gains a greater understanding of the service by evaluating the bank in terms of the factors comprising the choice criteria. Based on the evaluation, the manager forms a degree of preference for the bank. If preference exceeds a certain threshold level, the manager will switch business to the bank.

Graphical model

Custom
↓
Preference
↓
Understanding: Evaluation
↓
Awareness
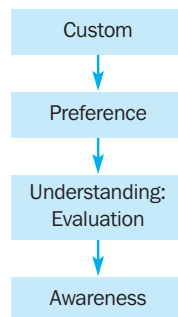
Mathematical model

$$y = a_0 + \sum_{i=1}^{n} a_i x_i$$

where y = degree of preference
a_0, a_i = model parameters to be estimated statistically
x_i = bank custom factors that constitute the choice criteria ∎

Research questions

Research questions are refined statements of the components of the problem. Although the components of the problem define the problem in specific terms, further detail may be needed to develop an approach. Each component of the problem may have to be broken down into subcomponents or research questions. Research questions ask what specific information is required with respect to the problem components. If the research questions are answered by the research, then the information obtained should aid the decision-maker. The formulation of the research questions should be guided not only by the problem definition, but also by the theoretical framework and the analytical model adopted. For a given problem component, there are likely to be several research questions, as in the GlobalCash Project.

e x a m p l e

GlobalCash Project

Research questions

One of the major events occurring during the GlobalCash Project was the introduction of systems to cope with Economic and Monetary Union (EMU). This was of major importance to banks, whose business is based upon the problems of conducting business across borders and paying for goods and services in a variety of currencies. Companies pay banks to cope with these transactions in an environment of complicated legislation. The marketplace changed dramatically when there was one currency for the majority of EC countries, with no currency exchange to conduct.

Key research questions for the banks that may lose business relate to relationships. Examples of these are shown in the following two questions.

Question: How will Economic and Monetary Union (EMU) affect your cash management banking relationships within the expected 'Euro zone'?

Answers: (please circle appropriate numbers)
1 Existing country relationships to be maintained within the Euro zone
2 Fewer banks to be used in each country
3 Fewer banks to be used across the Euro zone
4 One major bank to coordinate all Euro accounts where possible

Question: What process will you go through in selecting your cash management banks?

Answers: (please circle appropriate numbers)
1 Formal tender in each country
2 Formal tender for a regional bank
3 Informal evaluation
4 Other (*please state*)…………………………………… ■

Hypotheses

A **hypothesis** is an unproven statement or proposition about a factor or phenomenon that is of interest to the researcher. For example, it may be a tentative statement about relationships between two or more variables as stipulated by the theoretical framework or the analytical model. Often, a hypothesis is a possible answer to the research question.[19] Hypotheses go beyond research questions because they are statements of relationships or propositions rather than merely questions to which answers are sought. Research questions are interrogative; hypotheses are declarative and can be tested empirically (see Chapter 18). An important role of a hypothesis is to suggest variables to be included in the research design.[20] The relationship between the marketing research problem, research questions and hypotheses, along with the influence of the objective/theoretical framework and analytical models, is described in Figure 2.5 and illustrated by the following example from GlobalCash.

Hypotheses are an important part of the approach to a research problem. When stated in operational terms, as H_1 and H_2 in the GlobalCash example, they provide
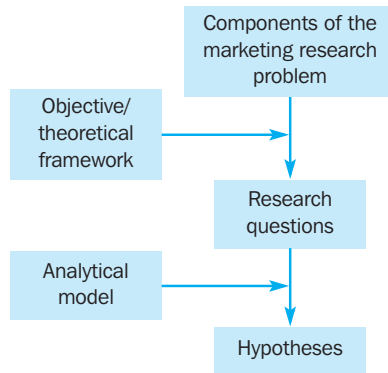
Figure 2.5
Development of research questions and hypotheses

guidelines on what, and how, data are to be collected and analysed. When operational hypotheses are stated using symbolic notation, they are commonly referred to as statistical hypotheses.

---

e x a m p l e

GlobalCash Project

Hypotheses

The following hypotheses were formulated in relation to the research question on bank–company relationships.

H_1: Companies within the Euro zone plan to use fewer banks after the introduction of Monetary Union compared with companies outside the Euro zone.

H_2: Companies that select cash management banks on an informal basis have fewer bank relationships than those that select through formal tenders.

The additional research question that will allow H_1 to be tested relates to the country in which an ultimate parent country operates, e.g. Hitachi, based in Britain and Germany, has a Japanese parent company.

The European Union has assigned countries as being 'within' or 'outside' for the first wave of Monetary Union. The additional question that will allow H_2 to be tested asks how many banks are used for domestic cash management. In countries like Switzerland and Norway, the mean can be around three; in Italy the mean is around 10, with one Italian respondent having 70 bank accounts. ■

These hypotheses guide the research by ensuring that variables which affect relationships are included in the research design. The connection between questions that form hypotheses, and the appropriate analyses needed, are thought out clearly before the research design and questionnaire are designed.

Another example in a consumer goods environment is shown in the following example of Chanel.

---

e x a m p l e

Chanel perfumes[21]

A research question may have more than one hypothesis associated with it, as in the case of Chanel. Chanel was considering advertising its perfumes in magazines it formerly considered too 'down-market' for its prestigious brand. The Chanel brand had a 3% share of department store sales (the leader was Estée Lauder with a 21% share). By expanding its advertising beyond high-fashion magazines, Chanel hoped to improve its share of department store sales. The following research question and hypotheses may be posed:

Research question: Does Chanel have an 'up-market' image?
H_1: Chanel is perceived to be an expensive brand.
H_2: Users of Chanel have higher-than-average incomes.
H_3: Users of Chanel associate this perfume with status. ■

Note that, to test H_1, the researcher would have to operationalise and measure the perceived price associated with Chanel. Empirical testing of H_2 would require that respondents be classified as users or non-users of Chanel and provide information on their incomes. Finally, H_3 tells us that we need to operationalise another variable or a set of variables that measure the status associated with Chanel. The results of this research provided support for Hypotheses 1 and 3 but not for Hypothesis 2. Although Chanel did have an 'up-market' image, its appeal was not limited to 'up-market' buyers or users of the product. Broadening the target market by advertising in magazines formerly considered 'down-market' led to improved department store sales of Chanel.

It is important to note that not all research questions can be developed into hypotheses that may be tested. Certain research questions may be exploratory in nature, with the researcher having no preconceived notions of possible answers to the research questions, nor the ability to produce statements of relationships or propositions. If the researcher is faced with such a situation, it does not mean that their investigation will not be as valid as one where hypotheses are clearly established. It means that they may have to adopt a different approach or paradigm to establish its validity.

International marketing research

The precise definition of the marketing research problem is more difficult in international marketing research than in domestic marketing research. Unfamiliarity with the environmental factors of the country where the research is being conducted can greatly increase the difficulty of understanding the problem's environmental context and uncovering its causes.

example

Heinz Ketchup couldn't catch up in Brazil [22]

Despite good records of accomplishment inland and overseas, the H.J. Heinz Company failed in Brazil, a market that seemed to be South America's biggest and most promising market. Heinz entered a joint venture with Citrosuco Paulista, a giant orange juice exporter, because of the future possibility of buying the profitable company. Yet, the sales of its products, including ketchup, did not take off. Where was the problem? A problem audit revealed that the company lacked a strong local distribution system. Heinz lost control of the distribution because it worked on consignment. Distribution could not reach 25% penetration. The other related problem was that Heinz concentrated on neighbourhood shops because this strategy was successful in Mexico. The problem audit, however, revealed that 75% of the grocery shopping in São Paulo is done in supermarkets and not the smaller shops. Although Mexico and Brazil may appear to have similar cultural and demographic characteristics, consumer behaviour can vary greatly. A closer and intensive look at the Brazilian food distribution system and the behaviour of consumers could have averted this failure. ■

As the Heinz example illustrates, many international marketing efforts fail not because research was not conducted but because the relevant environmental factors were not taken into account and fully appreciated. Generally, this leads to a definition of the problem that is too narrow. A major problem for researchers is that their perception of problems may be reflected through their own social and cultural development. Before defining the problem, researchers should reflect upon their unconscious reference to cultural values. The following steps help researchers to reflect upon their own cultural values.[23]

1 Define the marketing research problem in terms of domestic environmental and cultural factors. This involves an identification of relevant European traits, economics, values, needs or habits.

2 Define the marketing research problem in terms of foreign environmental and cultural factors. Make no judgements. This involves an identification of the related traits, economics, values, needs or habits in the proposed market culture. This task requires input from researchers familiar with the foreign environment.

3 Examine the differences between steps 1 and 2. The unconscious reference to cultural values can be seen to account for these differences.

4 Redefine the problem without the social/cultural influence and address it for the foreign market situation. If the differences in steps 3 are significant, the impact of the social/cultural influences should be carefully considered.

Whilst the above steps may seem at face value to be straightforward, note the words 'unconscious reference' in Step 3. This means that the researcher needs to reflect upon their own values and attitudes, the factors that may bias the way they perceive and what they observe. As these factors may be 'unconscious', this reflection and introspection may take some time to realise. These issues will be more fully developed in Chapter 9.

While developing theoretical frameworks, models, research questions and hypotheses, remember that differences in the environmental factors, especially the socio-cultural environment, may lead to differences in the formation of perceptions, attitudes, preferences and choice behaviour. For example, orientation towards time varies considerably across cultures. In Asia, Latin America and the Middle East, people are not as time-conscious as Westerners. This influences their perceptions of and preferences for convenience foods such as frozen foods and prepared dinners. In developing an approach to the problem, the researcher should consider the equivalence of consumption and purchase behaviour and the underlying factors that influence them. This is critical to the identification of the correct research questions, hypotheses and characteristics/factors that influence the research design.

The following example reveals how the use of focus groups could have helped to reveal social/cultural characteristics of the Japanese. Focus groups would have allowed the correct identification of research questions, leading to a more successful product launch.

***

example

Surf Superconcentrate faces a super washout in Japan[24]

Unilever attempted to break into the Japanese detergent market with Surf Superconcentrate. It initially achieved 14.5% of the market share during test marketing but fell down to a shocking 2.8% when the product was introduced nationally. Where did they go wrong? Surf was designed to have a distinctive pre-measured packet as in teabag-like sachets, joined in pairs because convenience was an important attribute to Japanese consumers. It also had a 'fresh smell' appeal. Japanese consumers, however, noticed that the detergents did not dissolve in the wash, partly because of weather conditions and because of the popularity of low-agitation washing machines. Surf was not designed to work in the new washing machines. Unilever also found that the 'fresh smell' positioning of new Surf had little relevance since most consumers hang their wash out in the fresh air. The research approach was certainly not without flaw, as Unilever failed to identify critical attributes that are relevant in the Japanese detergent market. Furthermore, it identified factors such as 'fresh smell' that had no relevance in the Japanese context. Appropriate qualitative research such as focus groups across samples from the target market could have revealed the correct characteristics or factors leading to a suitable research design. ■

Ethics in marketing research

Ethical situations arising from the process of problem definition and developing an approach are likely to occur between the market researcher and the client. As explained earlier, identifying the correct marketing research problem is crucial to the success of the project. This process can, however, be compromised by the personal agendas of the researcher or the decision-maker. For example, the researcher, after performing the tasks involved in problem definition and analysing the environmental context of the problem, realises that the correct marketing research problem may be defined in a way that makes primary research unnecessary. This would reduce the cost of the project and the research firm's profit margin substantially. Does the researcher define the problem correctly, fabricate a research problem that involves primary data collection, or refuse to proceed with this project in lieu of those more profitable? The researcher is faced with an ethical dilemma, as in the following example.

e x a m p l e

Taste (profits) or image (ethics)? [25]

A marketing research firm is hired by a soft drink company to conduct taste tests to determine why its newly introduced soft drink brand has not captured the expected market share. The researcher, after following the process outlined in this chapter, determines that the problem is not one of taste but of the image and product positioning. The client, however, has already defined the problem as a taste problem and not as the broader, market-share problem. The researcher must also weigh the relatively high profit margin of taste test research against the less lucrative survey research needed to answer questions pertaining to soft drink image. What should researchers do? Should they simply conduct the research the client wants rather than the research they feel the client needs? The guidelines indicate that 'the researcher has a professional obligation to indicate to the client that, in his or her judgement, the research expenditure is not warranted. If, after this judgement has been clearly stated, the client still desires the research, the researcher should feel free to conduct the study. The reason for this is that the researcher can never know for certain the risk preferences and strategies that are guiding the client's behaviour.' ■

Such ethical situations would be satisfactorily resolved if the client/researcher relationship developed with both the client and the researcher adhering to the seven Cs discussed earlier: communication, cooperation, confidence, candour, closeness, continuity and creativity. This would provide a relationship of mutual trust that would check any unethical tendencies.

Ethical situations affecting the researcher and the client may also arise in developing an approach to the problem. When researchers conduct studies for different clients in related industries (e.g. banking and financial services) or in similar research areas (e.g. customer satisfaction) they may be tempted to cut corners in theoretical framework and model development. Take an example where a grocery chain client has on its board of directors the chairman of a bank. The bank had recently conducted customer satisfaction research using a client-specific model, and the bank-affiliated board member has access to this research. The researcher feels that a customer satisfaction model for the bank could be easily adapted to work for the grocery chain. The client feels that it would not be a good business decision to have access to this information and not use it. Is it ethical for the client and researcher to obtain and use this model developed for another company by another research firm? There is an underlying trust between the researcher and the client that the research firm is honour-bound not to reuse client-specific models or findings for other projects.

The client also has an ethical responsibility not to solicit proposals merely to gain the expertise of the research firms without pay. It is unethical for a client to solicit proposals from a few research firms, then adopt one or a combination of the approaches suggested in them, and conduct the project in-house. The client must respect the rights of a firm by releasing that an unpaid proposal belongs to the research firm that generated it. However, if the client firm pays for the development of the proposal, it has a right to use the information contained in it.

Internet and computer applications

There are several ways in which the Internet can support the process of problem diagnosis and research design.

Discussions with the decision-maker

The Internet can help the researcher gain access to a wide variety of marketing decision-makers who may use and benefit from proposed research. Thanks to email, it is possible to reach decision-makers anywhere, at any time. The Internet can also provide chat rooms so that decision-makers and researchers can exchange and test ideas. The availability of the responses to be seen by whoever enters the chat room has the effect of getting all the relevant decision-makers together at the same time without requiring that they be physically present at the same time.

Interviews with industry experts

The Internet can be used to enhance the researcher's ability to obtain advice from experts. It can be searched to find industry experts outside the organisation that commissions a research project.

Secondary data location and analysis

Search engines can be used to locate secondary data quickly and economically. This can be vital in helping to understand the issues to be examined and the approach to examining those issues. We will discuss the availability, acquisition and quality of secondary data on the Internet in more detail in Chapter 4.

In setting the environmental context of the research problem, client-specific information can be gained from the company home page. Generally, companies provide information about their products and services in their home page, making it the ideal starting point for information about the company. Further, the user can also search for competitor and industry information on the Internet. While these searches may not provide complete answers for the marketing researcher, they may raise issues and identify contacts where further exploration may prove beneficial.

Summary

Defining the marketing research problem is the most important step in a research project. Problem definition is a difficult step, because frequently management has not determined the actual problem or has only a vague notion about it. The marketing researcher's role is to help management identify and isolate the problem.

The formal ways in which decision-makers and researchers communicate their perspectives on a research problem and how to solve it are through the development of a research brief and a research proposal. To fully develop these documents, researchers should be proactive in arranging discussions with key decision-makers, which should include a problem audit whenever possible. They should also conduct, where necessary, interviews with relevant experts, and secondary data collection and analyses. These tasks should lead to an understanding of the environmental context of the problem.

Analysis of the environmental context should assist in the identification of the marketing decision problem, which should then be translated into a marketing research problem. The marketing decision asks what the decision-maker needs to do, whereas the marketing research problem asks what information is needed and how it can be obtained effectively and efficiently. The researcher should avoid defining the marketing research problem either too broadly or too narrowly. An appropriate way of defining the marketing research problem is to make a broad statement of the problem and then identify its specific components.

Developing an approach to the problem is the second step in the marketing research process. The components of an approach may consist of an objective/theoretical framework, analytical models, research questions and hypotheses. It is necessary that the approach developed be based upon objective evidence or empirical evidence and be grounded in theory as far as it is appropriate. The relevant variables and their interrelationships may be neatly summarised via an analytical model. The most common kinds of model structures are verbal, graphical and mathematical. The research questions are refined statements of the specific components of the problem that ask what specific information is required with respect to the problem components. Research questions may be further refined into hypotheses. Finally, given the problem definition, research questions and hypotheses should be used to create a method to either measure or elicit an understanding of target respondents.

When defining the problem in international marketing research, the researcher must be aware of the impact of their own cultural values when evaluating the environmental impact upon the nature of a problem. Likewise, when developing an approach, the differences in the environment prevailing in the domestic market and the foreign markets should be carefully considered. Several ethical issues that have an impact on the client and the researcher can arise at this stage but can be resolved by adhering to the seven Cs: communication, cooperation, confidence, candour, closeness, continuity and creativity.

Questions

1 What is the nature of the first step in conducting a marketing research project?

2 Why is it vital to define the marketing research problem correctly?

3 What is the role of the researcher in the problem definition process?

4 What are the components of a marketing research brief?

5 What are the components of a marketing research proposal?

6 How may a marketing researcher be creative in interpreting a research brief and developing a research proposal?

7 What is the significance of the 'background' section of a research brief and research proposal?

8 Describe some of the reasons why management is often not clear about the 'real' problem they face.

9 What interrelated events occur in the environmental context of a research problem?

10 What are some differences between a marketing decision problem and a marketing research problem?

11 Describe the factors that may affect the approach to a research problem.

12 What is the role of theory in the development of a research approach?

13 What are the most common forms of analytical models?

14 What are the differences between research questions and hypotheses?

15 Is it necessary for every research project to have a set of hypotheses? Why or why not?

Notes

1 Carey, G., Zhao, X., Chiaramonte, J. and Eden, D., 'Is there a global village for our future generation? Talking to 7–12-year-olds around the world', *Marketing and Research Today* (February 1997), 12.

2 Butler, P., 'Marketing problem: from analysis to decision', *Marketing Intelligence & Planning* 12(2) (1994) 4–12.

3 Ingledew, S., 'How to bait the interview hook for those Top 1000 big fish', *ResearchPlus* (October 1996), 4.

4 Greenhalgh, C., 'How should we initiate effective research?' The Market Research Society Conference, 1983.

5 Clemens, J., 'Some dos and don'ts of pan-European research', *ResearchPlus* (November 1996), 4.

6 Wierenga, B. and van Bruggen, G.H., 'The integration of marketing problem solving modes and marketing management support systems', *Journal of Marketing* 61(3) (July 1997), 21–37.

7 Cronin, M.J., 'Using the web to push key data to decision makers', *Fortune* 36(6) (29 September 1997), 254.

8 Merrilyn Astin Tarlton, 'Quick marketing audit', *Law Practice Management* 23(6) (September 1997), 18, 63; Berry, L.L., Conant, J.S. and Parasuraman, A., 'A framework for conducting a services marketing audit', *Journal of the Academy of Marketing Science* 19 (Summer 1991), 255–8; Ackoff, Russell L., *Scientific Method* (New York: Wiley, 1961), 71; Ackoff, Russell L., *The Art of Problem Solving* (New York: Wiley, 1978).

9 Saviour, L.S., Nwachukwu and Vitell, S.J. Jr., 'The influence of corporate culture on managerial ethical judgments', *Journal of Business Ethics* 16(8) (June 1997), 757–76.

10 Winett, R., 'Guerilla marketing research outsmarts the competition', *Marketing News* 29(1) (January 1995), 33; Armstrong, J.S., 'Prediction of consumer behaviour by experts and novices', *Journal of Consumer Research* 18 (September 1991), 251–6.

11 Barrett, P., 'Abroad minded', *Marketing* (24 April 1997), 20.

12 Heyl, S.M., 'Decision matrix points the way to better research ROI', *Marketing News* 21(19) (15 September 1997), 18, 30.

13 Shannon, J., 'Net has yet to prove its value', *Marketing Week* (26 June 1997), 25.

14 Potter, G., *The Philosophy of Social Science: New Perspectives* (Harlow: Pearson, 2000), 242.

15 Buchanan, P., 'Putting sales theory into practice', *US Banker* 107(9) (12 September 1997), 112; see also Hunt, S.D. 'Truth in marketing theory and research', *Journal of Marketing* 54 (July 1990), 1–15; Hunt, S.D., 'For reason and realism in marketing', *Journal of Marketing* 56 (April 1992), 89–102.

16 For an illustration of a graphical model of software piracy, see Figure 1 of Givon, M., Mahajan, V. and Muller, E., 'Software piracy: estimation of lost sales and the impact on software diffusion', *Journal of Marketing* 59 (January 1995), 29–37.

17 For an example of developing a theoretical framework and a mathematical model based on it, see Miller, C.M., McIntyre,

S.H. and Mantrala, M.K., 'Toward formalizing fashion theory', *Journal of Marketing Research* 30 (May 1993), 142–57.

18 Lilien, G.L., Kotler, P. and Moorthy, K.S., *Marketing Models* (Englewood Cliffs, NJ: Prentice Hall, 1992).

19 For an example of hypothesis formulation, see Smith, N.C. and Cooper-Martin, E., 'Ethics and target marketing: the role of product harm and consumer vulnerability', *Journal of Marketing* 61 (July 1997), 1–20.

20 Kerlinger, F.N., *Foundations of Behavioral Research*, 3rd edn (New York: Holt, Rinehart & Winston, 1986). See pp. 17–20 for a detailed discussion of the characteristics and role of hypotheses in research. For an alternative view, see Lawrence, R.J., 'To hypothesize or not to hypothesize? The correct "approach" to survey research', *Journal of the Market Research Society* 24 (October 1982), 335-43. For an example of model development and hypothesis formulation see Bitner, M.J. 'Servicescapes: the impact of physical surroundings on customers and employees', *Journal of Marketing* 56 (April 1992), 57–71.

21 'Chanel plans to run ads in magazines with less cachet', *Wall Street Journal* (27 January 1988), 30.

22 Judann, D., 'Why Heinz went sour in Brazil', *Advertising Age* (5 December 1988).

23 Orpen, C., 'Developing international marketing strategies in small companies', *Journal of International Marketing & Marketing Research* 20(2) (June 1995), 89-96; and Douglas, S.P. and Craig, C.S., *International Marketing Research* (Englewood Cliffs, NJ: Prentice Hall, 1983).

24 Kilbum, D., 'Unilever struggles with Surf in Japan', *Advertising Age* (6 May 1991).

25 Laczniak, G.R. and Murphy, P.E., *Ethical Marketing Decisions, the Higher Road* (Boston: Allyn & Bacon, 1993), 64.

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Research design

Objectives

After reading this chapter, you should be able to:

1 define research design, classify various research designs, and explain the differences between exploratory and conclusive research designs;

2 compare and contrast the basic research designs: exploratory, descriptive and causal;

3 understand how respondents or the subjects of research design affect research design choices;

4 describe the major sources of errors in a research design, including random sampling error and the various sources of non-sampling error;

5 explain research design formulation in international marketing research;

6 understand the ethical issues and conflicts that arise in formulating a research design.

There are a huge array of alternative research designs that can satisfy research objectives. The key is to create a design that enhances the value of the information obtained, whilst reducing the cost of obtaining it.

Overview

Chapter 2 discussed how to define a marketing research problem and develop a suitable approach. These first two steps are critical to the success of the whole marketing research project. Once they have been completed, attention should be devoted to designing the formal research project by formulating a detailed research design (as a reminder, see Figure 2.3).

This chapter defines and classifies research designs. We describe the two major types of research design: exploratory and conclusive. We further classify conclusive research designs as descriptive or causal and discuss both types in detail. We then consider the differences between the two types of descriptive designs – cross-sectional and longitudinal – and identify sources of errors. The special considerations involved in formulating research designs in international marketing research are discussed. Several ethical issues that arise at this stage of the marketing research process are considered. The reader can develop a better appreciation of the concepts presented in this chapter by first considering the following example, which illustrates the use of a number of interrelated techniques to build a research design.

e x a m p l e

Getting to know you[1]

Building a relationship with consumers is a challenge facing all organisations, but particularly so in the case of 'emergent drinkers', those of legal drinking age up to 25. Allied Domecq Spirits and Wines (ADSW) recognised the danger of being distanced from this crucial group, particularly across geographical markets. ADSW worked with Pegram Walters International (PWi) on a project that went far beyond an exploration of the current usage and attitudes towards spirits. The objectives of the project encompassed an exploration of the target groups' personal values, their feelings about their lives, their universe, their hopes and dreams.

There were three stages to the research design. In the first stage the researchers conducted one-hour depth interviews. There were three clear objectives for this stage: to understand personal viewpoints on marketing and lifestyle issues; to clarify and/or narrow down topics for subsequent exploration; and to recruit appropriate 'information gatherers'. From this stage hypotheses were formulated on issues such as how respondents saw themselves and their future, relationships, self-discovery and opting in or opting out of the system.

In the second stage, from 20 depth interviews, 10 respondents were retained as 'information gatherers'. 'Leading edge' bars were rented out and 50 adult emergent drinkers were invited to participate in workshops. Given a task guideline, the information gatherers led discussions. As an additional record, the workshops were video-recorded. The participants felt comfortable within their peer group and, in the more natural bar environment, fed back real, relevant and honest information.

The third stage occurred on the night following the workshops. Focus groups were used, made up of the 'information gatherers'. They discussed what happened in the workshops and their interpretation of what it actually meant.

In order to ensure that the information remained topical, useful and easily accessible, it was felt important to create a vehicle for an on-going communication and dialogue with the target market. To achieve this, they created a high impact 'magazine' to bring the research to life after the presentation of findings. This was referred to as a magazine and not a research report to reflect the lifestyle of the consumer group in question: it contained images, layouts and fonts typically associated with the generation. ■

The above example illustrates a very creative and useful exploratory research design. As a research design it worked well in that it achieved a balance of the needs and expectations of marketing decision-makers and respondents. Decision-makers helped to set clear research objectives based upon the gaps in their knowledge of the

target market. Respondents related well to the issues posed to them, in a context and environment that was comfortable to them. An understanding of the fundamentals of research design, its components, and the trade-offs between the parties involved in making the design work, enabled the researchers to formulate the most appropriate design for the problem at hand.

Research design definition

Research design
A framework or blueprint for conducting the marketing research project. It specifies the details of the procedures necessary for obtaining the information needed to structure or solve marketing research problems.

A **research design** is a framework or blueprint for conducting a marketing research project. It details the procedures necessary for obtaining the information needed to structure or solve marketing research problems. Although a broad approach to the problem has already been developed, the research design specifies the details – the practical aspects – of implementing that approach. A research design lays the foundation for conducting the project. A good research design will ensure that the marketing research project is conducted effectively and efficiently. Typically, a research design involves the following components or tasks, which will be discussed in detail in the subsequent chapters:

1 Define the information needed (Chapter 2).
2 Decide whether the overall design is to be exploratory, descriptive or causal (Chapter 3).
3 Design the sequence of techniques of understanding and/or measurement (Chapters 4 to 12).
4 Construct and pre-test an appropriate form for data collection or questionnaire (Chapters 7, 8 and 13).
5 Specify the qualitative and/or quantitative sampling process and sample size (Chapters 6, 14 and 15).
6 Develop a plan of qualitative and/or quantitative data analysis (Chapters 9 and 17).

In formulating a research design, the researcher has to balance the perspectives of marketing decision-makers and target respondents. From their education and experience, marketing decision-makers may have certain techniques that they believe to be the most effective and in which they subsequently have more confidence. There is no problem with this, providing the technique is the best means to measure or understand the issue under investigation, from the perspective of respondents. In the example at the start of this chapter, decision-makers had confidence in the qualitative techniques and the data generated. The techniques worked well with the respondents, drawing out a rich picture of respondent behaviour, lifestyle and aspirations. However, should the decision-makers feel that survey techniques are the most effective, giving them the most confidence to support their decisions, the researcher may face a dilemma. If they use survey techniques they may find that respondents may have a different relationship with interviewers, do not reflect in the same manner and ultimately do not reveal so much. Thus, research design involves the researcher developing an understanding of the type of data decision-makers have confidence in, plus an understanding of how respondents may respond to different techniques. The first part of this balancing act involves understanding research design from the decision-makers' perspective; the second part involves understanding the respondents' perspective.

Research design from the decision-maker's perspective

Marketing decision-makers seek support from marketing researchers that is of practical relevance to the decisions they face. To give practical support, decision-makers expect information that is:

- *Accurate*, i.e. the most valid representation of the phenomena under investigation, that has come from the most reliable or consistent form of measurement or understanding, that is sufficiently sensitive to the important differences in individuals being measured or understood. Combining these three criteria refers to the degree to which information may be deemed as 'accurate'.
- *Current*, i.e. as up to date as possible. This is particularly important where consumer attitudes, lifestyle or behaviour changes rapidly, perhaps due to rapid technology changes or new product offerings in a highly competitive market.
- *Sufficient*, i.e. the completeness or clarity of a 'picture' that reflects the characteristics of the marketing problem they face.
- *Available*, i.e. that access to the relevant information can be made when a decision is imminent. This is particularly important where competitive activity forces the decision-maker into making a rapid response.
- *Relevant*, i.e. that the support given 'makes sense' to decision-makers. In very general terms, decision-makers may criticise qualitative techniques for being biased and unrepresentative and quantitative techniques for lacking depth and a holistic perspective. Ideally, whichever technique is adopted, decision-makers should be aware of the benefits, limitations and even alternatives. With this awareness they can use the findings with confidence to build upon their existing experiences and knowledge.

Generating information that fulfils all the above characteristics is extremely difficult, if not impossible to achieve in marketing research. The evaluation of sources of error, presented later in this chapter, and the restrictions of budget and timescales mean that this list represents 'ideals'. Realistically, trade-offs must be made among the above characteristics. Within the first characteristic of accuracy there are further trade-offs which are primarily caused by what the marketing researcher is attempting to measure or understand:[2]

1 The subject of investigation is usually human.
2 The process of measuring or observing humans may cause them to change.
3 It is difficult to assess the effect of extraneous variables in marketing experiments and thus their applications are limited.

Given the complexity of the subjects under study, the context or environment in which measurements are taken, and the skills required to perform and interpret measurements, it is difficult (if not impossible) to gain completely objective and accurate measurements. Of all the potential trade-offs, if one were to remove *relevance* then the whole rationale of supporting the marketing decision-maker has been removed. Therefore this characteristic can never be compromised.

Relevance embraces, *inter alia*, the ability to plan and forecast from research findings, to be able to distinguish real differences in consumer traits, and to know that characteristics are representative of groups of individuals. With relevant information such as this, the decision-maker can build up a stronger understanding or awareness of markets and the forces that shape them. In building up this understanding, the decision-maker cannot turn to a single technique or even body of techniques that may be deemed the 'ideal' in ensuring that information is relevant. In different types of

decision-making scenario, different techniques will offer the best support for that decision-maker. Establishing the best form of support is the essence of research design.

A fundamental starting point in deciding an appropriate design is viewing the process from the point of view of the potential subject or respondent to a marketing research study.

Research design from respondents' perspectives

The potential respondents to any marketing research investigation play a vital role in deciding which research design will actually work in practice. A subject of study may be complex and need time for respondents to reflect upon and put words to the questions posed. Certain methods are more likely to build up a rapport and trust, in these circumstances, putting the respondent in the right frame of mind, and getting them to respond in a full and honest manner. Figure 3.1 is a framework that serves to remind how respondents may be accessed, and what kinds of response may be generated.

In Figure 3.1 the box under the heading 'Layers of response from respondents' represents how respondents may react to questions posed to them. In the first layer of '*Spontaneous, Reasoned, Conventional*' are questions that respondents can express a
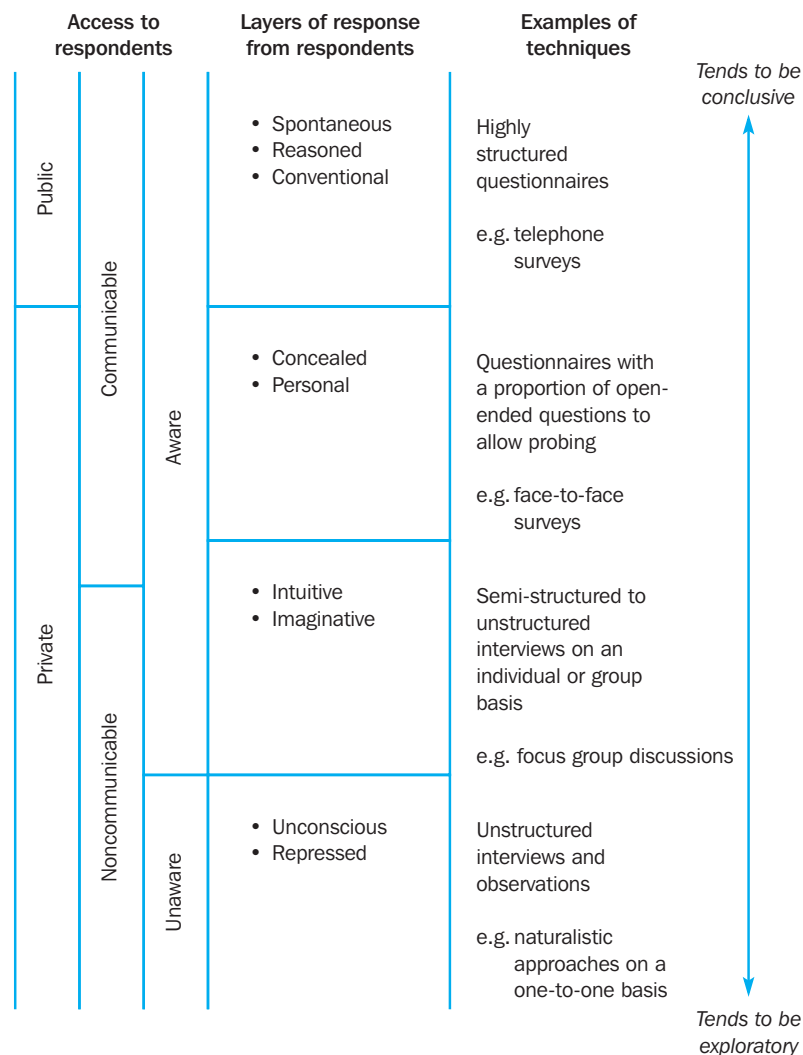
Figure 3.1
Responses to interviewing[3]

view about quickly, that are simple for them to reflect upon, relating to common every-day occurrences that are at the forefront of their mind. In such circumstances, simple structured questioning (or self-reporting) in a standardised manner is possible. Further, the same procedure can be conducted in a consistent manner to a whole array of 'types' of respondent such as age groups, social class and intellectual levels. For example, if questions were posed on which newspapers someone reads, it is a reasonable assumption that respondents are aware of the title(s), the title(s) can be communicated and the topic is not a sensitive issue. In these circumstances, where answers to questions on reading habits are relatively easy to access and respond to, highly structured questionnaires are appropriate. Clearly, in such situations, quantitative techniques are applicable that allow very detailed descriptions or experiments to be made.

Progressing down, at the second level are questions that are more personal and more sensitive, such as the use of personal hygiene products or, in business-to-business research, commercially sensitive information. Again, structured questionnaires can measure the relevant issues but an amount of rapport may be needed to induce respondents to trust the interviewer and reveal their 'more personal' attitudes and behaviour.

At the third level are questions that require respondents to be creative. For example, if respondents were to be asked about their attitudes and behaviour towards eating yoghurt, this could be done in a very structured manner. Questions could be set to determine when it was eaten, favourite flavours and brands, where it was bought, how much was spent, etc. The same can be said of alcohol consumption, though this could well be a sensitive issue for many respondents. Now imagine a new product idea that mixes yoghurt and alcohol. What combinations of alcohol and yoghurt would work, and what types of consumer would be attracted to it? Would it be a dessert liqueur such as Baileys Irish Cream or frozen yoghurt to compete with the Häagen Dazs luxury ice creams? Would champagne, advocaat, whisky or beer be the best alcoholic ingredient? Should any fruits be added? Individually? Forest fruits? Tropical fruits? How would it be packaged? What name would best suit it? What price level would it sell at? On what occasions would it be consumed?

Answering these questions demands a great amount of creativity and imagination. It demands that respondents reflect upon ideas, can play with ideas and words and dig deep to draw out ideas in a relaxed manner. Structured questionnaires cannot do this; such a scenario would work best with the use of focus groups.

At the fourth level may be questions that respondents may not be able to conceptualise, never mind be willing to express what they feel about a particular issue. An example of such an issue may be trying to understand the childhood influences of family and friends on an individual's perception and loyalty to brands of washing-up liquid. Another example may be understanding the image consumers have of themselves and an image they wish to portray by spending €20,000 on a Rolex wristwatch. Respondents do not normally have to think through such issues or articulate them, until a marketing researcher comes along! In such circumstances, the characteristics of the individual determine what is the best way to probe and elicit appropriate responses. Nothing is standardised or consistent in these circumstances, the researcher having to shape the questions, probes and observations as they see fit in each interview.

As well as understanding how respondents may react to particular issues, researchers should also understand how the context or environment may affect respondents. As an example in the first level of Figure 3.1, respondents may be more relaxed and feel in control if they can answer the set questions about their newspaper reading habits on the Internet rather than on the street. In the example at the start of this chapter that explored the hopes and dreams of 'emergent drinkers', techniques were used at levels 3 and 4 of Figure 3.1. The context of the interviews was in 'leading-

edge bars'. This context could have helped the target respondents to relax, to develop a better rapport with interviewers and other respondents, and to think more about the issues and express their feelings more clearly. If the interviews were conducted over the Internet, the same levels of relaxation and rapport may not work. If the interviews were targeted at older respondents, they may have felt very self-conscious in 'leading-edge bars', which may restrict their responses. Researchers therefore must understand characteristics of respondents, how they react to particular issues and how they react in different contexts or environments. These factors are illustrated in Figure 3.2, which acts as a reminder of the understanding of respondents that researchers must develop, in order to choose and apply the best research technique.

Exploratory research
A research design characterised by a flexible and evolving approach to understand marketing phenomena that are inherently difficult to measure.

Conclusive research
A research design characterised by the measurement of clearly defined marketing phenomena.

Research design classification

Research designs may be broadly classified as exploratory or conclusive (see Figure 3.3). The differences between **exploratory research** and **conclusive research** are summarised in Table 3.1.
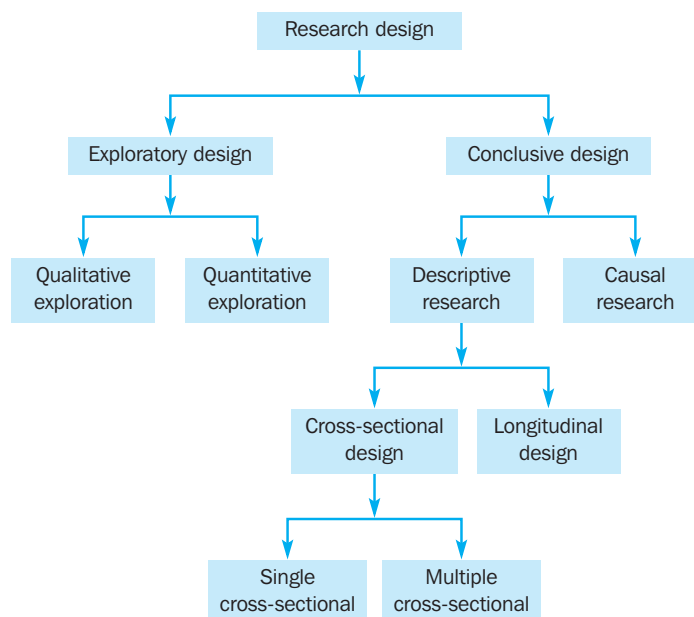
Figure 3.3
A classification of marketing research designs

Table 3.1 Differences between exploratory and conclusive research

	Exploratory	*Conclusive*
Objectives	To provide insights and understanding of the nature of marketing phenomena To understand	To test specific hypotheses and examine relationships To measure
Characteristics	Information needed may be loosely defined Research process is flexible, unstructured and may evolve Samples are small Data analysis can be qualitative or quantitative	Information needed is clearly defined Research process is formal and structured Sample is large and aims to be representative Data analysis is quantitative
Findings/results	Can be used in their own right May feed into conclusive research May illuminate specific conclusive findings	Can be used in their own right May feed into exploratory research May set a context to exploratory findings
Methods	Expert surveys Pilot surveys Secondary data Qualitative interviews Unstructured observations Quantitative exploratory multivariate methods	Surveys Secondary data Databases Panels Structured observations Experiments

The primary objective of exploratory research is to provide insights into and an understanding of marketing phenomena. It is used in instances where the subject of the study cannot be measured in a quantitative manner or where the process of measurement cannot realistically represent particular qualities. For example, if a researcher was trying to understand what 'atmosphere' meant in a restaurant, exploratory research may help to establish all the appropriate variables and how they connected together. What role did music play? What type of music? How loud? What types of furniture? What colours and textures? What types of lighting? What architectural features? This list could go on of what 'atmosphere' may mean in the context of a restaurant experience for particular types of consumer. 'Atmosphere' may not be measurable from the respondent's perspective. From the perspective of the creative director in an advertising agency, quantitative measurements of the individual components of 'atmosphere' may not create the holistic feel of a restaurant in a manner they can relate to.

Exploratory research may also be used in cases where you must define the problem more precisely, identify relevant courses of action, or gain additional insights before going on to confirm findings using a conclusive design. In such circumstances, the information needed will be loosely defined at an exploratory stage, using research questions rather than specific hypotheses or actual measurements. The research process that is adopted is characterised as being flexible, loosely structured and, in some circumstances, evolutionary in nature.

In an example of a flexible, loosely structured and evolutionary approach, consider conducting personal interviews with industry experts. The sample, selected to generate maximum insight, is small and non-representative. However, the emphasis in the sampling procedure is focused upon 'quality' individuals who are willing to open up, use their imagination, be creative and reveal perhaps sensitive thoughts and behav-

iour. 'Quality' also may emerge from their level of expertise; for example, there may only be a small population of chief executives in airline companies in Europe. If a small sample of, say, six chief executives from the largest and fastest developing airlines allowed access to a marketing researcher and revealed their attitudes and behaviour, insights may be gained that no conclusive study could achieve. By being flexible in the issues to discuss, loosely structured in how probes and additional issues emerge, and evolutionary in the nature of who to talk to and the best context in which to gain their confidence and get them to express what they really feel, an exploratory design can be very beneficial.

There is an exception to exploratory designs being built around qualitative techniques. There are examples of quantitative findings being used for exploratory purposes. For example, within a survey that examines specific research questions and hypotheses lies the opportunity to examine additional connections between questions that had not been initially considered. Simple correlations through to multivariate techniques that explore potential connections between questions may be conducted; this process is known as data mining (examined in more detail in Chapter 5). In essence, data mining searches for significant connections or patterns in a dataset that a researcher or decision-maker may be unaware of.

To summarise, exploratory research is meaningful in any situation where the researcher does not have enough understanding to proceed with the research project. Exploratory research is characterised by flexibility and versatility with respect to the methods, because formal research protocols and procedures are not employed. It rarely involves structured questionnaires, large samples and probability sampling plans. Rather, researchers are alert to new ideas and insights as they proceed. Once a new idea or insight is discovered, they may redirect their exploration in that direction. That new direction is pursued until its possibilities are exhausted or another direction is found. For this reason, the focus of the investigation may shift constantly as new insights are discovered. Thus, the creativity and ingenuity of the researcher play a major role in exploratory research. Exploratory research can be used for any of the purposes listed in Table 3.2.

Table 3.2 A summary of the uses of exploratory research designs

1	To obtain some background information where absolutely nothing is known about the problem area.
2	To define problem areas fully and to formulate hypotheses for further investigation and/or quantification.
3	Concept identification and exploration in the development of new product or forms of marketing communications.
4	During a preliminary screening process such as in new product development, in order to reduce a large number of possible projects to a smaller number of probable ones.
5	To identify relevant or salient behaviour patterns, beliefs, opinions, attitudes, motivations, etc. and to develop structures of these constructs.
6	To develop an understanding of the structure of beliefs and attitudes in order to aid the interpretation of data structures in multivariate data analyses.
7	To explore the reasons that lie behind the statistical differences between groups that may emerge from secondary data or surveys.
8	To explore sensitive or personally embarrassing issues from the respondents' and/or the interviewer's perspective.
9	To explore issues that respondents may hold deeply, that are difficult for them to rationalise and they may find difficult to articulate.
10	To 'data-mine' or explore quantitative data to reveal hitherto unknown connections between different measured variables.

The objective of conclusive research is to describe specific phenomena, to test specific hypotheses and examine specific relationships. This requires that the information needed is clearly specified.[4] Conclusive research is typically more formal and structured than exploratory research. It is based on large, representative samples, and the data obtained are subjected to quantitative analysis. Conclusive research can be used for any of the purposes listed in Table 3.3.

Table 3.3 A summary of the uses of conclusive research designs

1	To describe the characteristics of relevant groups, such as consumers, salespeople, organisations, or market areas.
2	To estimate the percentage in a specified population exhibiting a certain form of behaviour.
3	To count the frequency of events, especially in the patterns of consumer behaviour.
4	To measure marketing phenomena to represent larger populations or target markets.
5	To be able to integrate findings from different sources in a consistent manner, especially in the use of marketing information systems and decision support systems.
6	To determine the perceptions of product or service characteristics.
7	To compare findings over time that allow changes in the phenomena to be measured.
8	To measure marketing phenomena in a consistent and universal manner.
9	To determine the degree to which marketing variables are associated.
10	To make specific predictions.

As shown in Figure 3.3, conclusive research designs may be either descriptive or causal, and descriptive research designs may be either cross-sectional or longitudinal. Each of these classifications is discussed further, beginning with descriptive research.

Descriptive research

Descriptive research
A type of conclusive research that has as its major objective the description of something, usually market characteristics or functions.

As the name implies, the major objective of **descriptive research** is to describe something, usually market characteristics or functions.[5] A major difference between exploratory and descriptive research is that descriptive research is characterised by the prior formulation of specific research questions and hypotheses. Thus, the information needed is clearly defined. As a result, descriptive research is pre-planned and structured. It is typically based on large representative samples. A descriptive research design specifies the methods for selecting the sources of information and for collecting data from those sources.

Examples of descriptive studies in marketing research are as follows:

- Market studies describing the size of the market, buying power of the consumers, availability of distributors, and consumer profiles
- Market share studies determining the proportion of total sales received by a company and its competitors
- Sales analysis studies describing sales by geographic region, product line, type of the account and size of the account
- Image studies determining consumer perceptions of the firm and its products
- Product usage studies describing consumption patterns
- Distribution studies determining traffic flow patterns and the number and location of distributors

■ Pricing studies describing the range and frequency of price changes and probable consumer response to proposed price changes

■ Advertising studies describing media consumption habits and audience profiles for specific television programmes and magazines.

These examples demonstrate the range and diversity of descriptive research studies. Descriptive research can be further classified into cross-sectional and longitudinal research (Figure 3.3).

Cross-sectional designs

Cross-sectional design
A type of research design involving the collection of information from any given sample of population elements only once.

The cross-sectional study is the most frequently used descriptive design in marketing research. **Cross-sectional designs** involve the collection of information from any given sample of population elements only once. They may be either single cross-sectional or multiple cross-sectional (Figure 3.3). In **single cross-sectional designs**, only one sample of respondents is drawn from the target population, and information is obtained from this sample only once. These designs are also called sample survey research designs. In **multiple cross-sectional designs**, there are two or more samples of respondents, and information from each sample is obtained only once. Often, information from different samples is obtained at different times. The following examples illustrate single and multiple cross-sectional designs respectively.

Single cross-sectional design
A cross-sectional design in which one sample of respondents is drawn from the target population and information is obtained from this sample once.

Multiple cross-sectional design
A cross-sectional design in which there are two or more samples of respondents, and information from each sample is obtained only once.

e x a m p l e

Designing coupons from cross-sections[6]

A cross-sectional study based on a single survey was conducted to determine the effectiveness of sales promotion coupons in stimulating sales, as well as to assess coupon user and non-user profiles. The data were collected from 8,000 households. The results showed that 31% of all coupon-redeeming households accounted for 72% of all redemptions. Demographically, heavy coupon redeemers were large households with children and annual incomes exceeding €30,000, with female heads of household aged 35 to 54 who worked part-time. Light users of coupons were smaller households with female heads who were younger and worked full-time. Such information was useful to consumer products firms like Procter & Gamble that rely heavily on coupon promotion, as it enabled them to target their promotions to heavy coupon redeemers.■

e x a m p l e

Chase and Grabbits multiply like rabbits[7]

Eating behaviour trends were examined in a marketing research project commissioned by the Pillsbury Company. This project involved data from food diaries collected over three time waves. Each wave had a different sample of 1,000 households for a total sample size of 3,000 in the multiple cross-sectional design. Based on an analysis of eating patterns, the market was divided into five segments: Chase and Grabbits, Functional Feeders, Down Home Stokers, Careful Cooks and Happy Cookers. The changes in composition of these segments were examined over time. For example, the Chase and Grabbits experienced the biggest increase over the 15-year period (+136%). Currently, this group represents 26% of the total sample. Their desire for more convenience also increased over time. Says one Chase and Grabbit, 'Someday all you'll have to do is take a pill and it'll give you everything you need.' This information enabled the Pillsbury Company to target different products for different segments. For example, the Chase and Grabbit represented a prime segment for prepared foods and TV dinners. ■

The survey of coupon use, a single cross-sectional design, involved only one group of respondents who provided information only once. On the other hand, the Pillsbury study involved three different samples, each measured only once, with the measures obtained five years apart. Hence, the latter study illustrates a multiple cross-sectional design. A type of multiple cross-sectional design of special interest is cohort analysis.

Cohort analysis consists of a series of surveys conducted at appropriate time intervals, where the cohort serves as the basic unit of analysis. A cohort is a group of respondents who experience the same event within the same time interval.[8] For example, a birth (or age) cohort is a group of people who were born during the same time interval, such as 1951–60. The term cohort analysis refers to any study in which there are measures of some characteristics of one or more cohorts at two or more points in time.

It is unlikely that any of the individuals studied at time 1 will also be in the sample at time 2. For example, the age cohort of people between 8 and 19 years was selected, and their soft drink consumption was examined every 10 years for 30 years. In other words, every 10 years a different sample of respondents was drawn from the population of those who were then between 8 and 19 years old. This sample was drawn independently of any previous sample drawn in this study from the population of 8 to 19 years. Obviously, people who were selected once were unlikely to be included again in the same age cohort (8 to 19 years), as these people would be much older at the time of subsequent sampling. This study showed that this cohort had increased consumption of soft drinks over time. Similar findings were obtained for other age cohorts (20–29, 30–39, 40–49, and 50+). Further, the consumption of each cohort did not decrease as the cohort aged. These results are presented in Table 3.4 in which the consumption of the various age cohorts over time can be determined by reading down the diagonal. These findings contradict the common belief that the consumption of soft drinks will decline with the greying of Western economies. This common but erroneous belief has been based on single cross-sectional studies. Note that if any column of Table 3.4 is viewed in isolation (as a single cross-sectional study) the consumption of soft drinks declines with age, thus fostering the erroneous belief.[9]

Table 3.4 **Consumption of soft drinks by various age cohorts**

(Percentage consuming on a typical day)

Age	1950	1960	1970	1980	
8–19	53	63	73	81	
20–29	45	61	76	76	C8
30–39	34	47	68	71	C7
40–49	23	41	59	68	C6
50+	18	29	50	52	C5
		C1	C2	C3	C4

C1: cohort born prior to 1900   C4: cohort born 1921–30   C7: cohort born 1951–60
C2: cohort born 1901–10   C5: cohort born 1931–40   C8: cohort born 1961–70
C3: cohort born 1911–20   C6: cohort born 1941–50

Cohort analysis is also used to predict changes in voter opinions during a political campaign. Well-known marketing researchers such as MORI or Gallup, who specialise in political opinion research, periodically question cohorts of voters (people with similar voting patterns during a given interval) about their voting preferences to predict election results. Thus, cohort analysis is an important cross-sectional design. The other type of descriptive design is longitudinal design.

Longitudinal designs

Longitudinal design
A type of research design involving a fixed sample of population elements measured repeatedly. The sample remains the same over time, thus providing a series of pictures that, when viewed together, portray a vivid illustration of the situation and the changes that are taking place.

In **longitudinal designs**, a fixed sample (or samples) of population elements is measured repeatedly. A longitudinal design differs from a cross-sectional design in that the sample or samples remain the same over time. In other words, the same people are

studied over time. In contrast to the typical cross-sectional design, which gives a snapshot of the variables of interest at a single point in time, a longitudinal study provides a series of 'pictures'. These 'pictures' give an in-depth view of the situation and the changes that take place over time. For example, the question 'how did the German people rate the performance of President Helmut Kohl immediately after unification of West and East Germany?' would be addressed using a cross-sectional design. A longitudinal design, however, would be used to address the question 'how did the German people change their view of Kohl's performance during his presidency?'

Panel
A sample of respondents who have agreed to provide information at specified intervals over an extended period.

Often, the term **panel** is used interchangeably with the term longitudinal design. A panel consists of a sample of respondents, generally households, who have agreed to provide information at specified intervals over an extended period. Panels are maintained by syndicated firms, and panel members are compensated for their participation with gifts, coupons, information or cash.[10]

Data obtained from panels may provide information on market shares that are based on an extended period of time. Such data may also allow the researcher to examine changes in market share over time.[11]

Relative advantages and disadvantages of longitudinal and cross-sectional designs

The relative advantages and disadvantages of longitudinal versus cross-sectional designs are summarised in Table 3.5. A major advantage of longitudinal design over the cross-sectional design is the ability to detect change as a result of repeated measurement of the same variables on the same sample.

Table 3.5 **Relative advantages and disadvantages of longitudinal and cross-sectional designs**

Evaluation criteria	Cross-sectional design	Longitudinal design
Detecting change	–	+
Large amount of data collection	–	+
Accuracy	–	+
Representative sampling	+	–
Response bias	+	–

Note: A + indicates a relative advantage over the other design, whereas a – indicates a relative disadvantage.

Tables 3.6 and 3.7 demonstrate how cross-sectional data can mislead researchers about changes over time. The cross-sectional data reported in Table 3.6 reveal that the purchases of Brands A, B and C remain the same in periods 1 and 2. In each survey, 20% of the respondents purchase Brand A, 30% Brand B, and 50% Brand C. The longitudinal data presented in Table 3.7 show that substantial change, in the form of brand switching, occurred in the study period. For example, only 50% (100/200) of the respondents who purchased Brand A in period 1 also purchased it in period 2. The corresponding repeat

Table 3.6 **Cross-sectional data may not show change**

Brand purchased	Time period	
	Period 1 Survey	Period 2 Survey
Total surveyed	1,000	1,000
Brand A	200	200
Brand B	300	300
Brand C	500	500

Table 3.7 **Longitudinal data may show substantial change**

Brand purchased in Period 1	Brand purchased in Period 2			
	Brand A	Brand B	Brand C	Total
Total surveyed	200	300	500	1,000
Brand A	100	50	50	200
Brand B	25	100	175	300
Brand C	75	150	275	500

purchase figures for Brands B and C are, respectively, 33.3% (100/300) and 55% (275/500). Hence, during this interval Brand C experienced the greatest loyalty and Brand B the least. Table 3.7 provides valuable information on brand loyalty and brand switching (such a table is called a turnover table or a brand-switching matrix).[12]

Longitudinal data enable researchers to examine changes in the behaviour of individual units and to link behavioural changes to marketing variables, such as changes in advertising, packaging, pricing and distribution. Since the same units are measured repeatedly, variations caused by changes in the sample are eliminated and even small variations become apparent.[13]

Another advantage of panels is that relatively large amounts of data can be collected. Because panel members are usually compensated for their participation, they are willing to participate in lengthy and demanding interviews. Yet another advantage is that panel data can be more accurate than cross-sectional data.[14] A typical cross-sectional survey requires the respondent to recall past purchases and behaviour; these data can be inaccurate because of memory lapses. Panel data, which rely on continuous recording of purchases in a diary, place less reliance on the respondent's memory. A comparison of panel and cross-sectional survey estimates of retail sales indicates that panel data give more accurate estimates.[15]

The main disadvantage of panels is that they may not be representative. Non-representativeness may arise because of:

1 *Refusal to cooperate.* Many individuals or households do not wish to be bothered with the panel operation and refuse to participate. Consumer panels requiring members to keep a record of purchases have a cooperation rate of 60% or less.
2 *Mortality.* Panel members who agree to participate may subsequently drop out because they move away or lose interest. Mortality rates can be as high as 20% per year.[16]
3 *Payment.* Payment may cause certain types of people to be attracted, making the group unrepresentative of the population.

Another disadvantage of panels is response bias. New panel members are often biased in their initial responses. They tend to increase the behaviour being measured, such as food purchasing. This bias decreases as the respondent overcomes the novelty of being on the panel, so it can be reduced by initially excluding the data of new members. Seasoned panel members may also give biased responses because they believe they are experts or they want to look good or give the 'right' answer. Bias also results from boredom, fatigue and incomplete diary entries.[17]

Causal research

Causal research
A type of conclusive research where the major objective is to obtain evidence regarding cause-and-effect (causal) relationships.

Causal research is used to obtain evidence of cause-and-effect (causal) relationships. Marketing managers continually make decisions based on assumed causal relationships. These assumptions may not be justifiable, and the validity of the causal relationships should be examined via formal research.[18] For example, the common

assumption that a decrease in price will lead to increased sales and market share does not hold in certain competitive environments. Causal research is appropriate for the following purposes:

1 To understand which variables are the cause (independent variables) and which variables are the effect (dependent variables) of marketing phenomena.
2 To determine the nature of the relationship between the causal variables and the effect to be predicted.
3 To test hypotheses.

Like descriptive research, causal research requires a planned and structured design. Although descriptive research can determine the degree of association between variables, it is not appropriate for examining causal relationships. Such an examination requires a causal design, in which the causal or independent variables are manipulated in a relatively controlled environment. Such an environment is one in which the other variables that may affect the dependent variable are controlled or checked as much as possible. The effect of this manipulation on one or more dependent variables is then measured to infer causality. The main method of causal research is experimentation.[19]

Due to the complexity and importance of this subject, Chapter 11 has been devoted to causal designs and experimental research.

Relationships between exploratory, descriptive and causal research

We have described exploratory, descriptive and causal research as major classifications of research designs, but the distinctions among these classifications are not absolute. A given marketing research project may involve more than one type of research design and thus serve several purposes.[20] Which combination of research designs to employ depends on the nature of the problem. We offer the following general guidelines for choosing research designs.

1 When little is known about the problem situation, it is desirable to begin with exploratory research. Exploratory research is appropriate for the following:
   (a) When the nature of the topic under study cannot be measured in a structured, quantifiable manner.
   (b) When the problem needs to be defined more precisely.
   (c) When alternative courses of action need to be identified.
   (d) When research questions or hypotheses need to be developed.
   (e) When key variables need to be isolated and classified as dependent or independent.
2 Exploratory research may be an initial step in a research design. It may be followed by descriptive or causal research. For example, hypotheses developed via exploratory research can be statistically tested using descriptive or causal research.
3 It is not necessary to begin every research design with exploratory research. It depends on the precision with which the problem has been defined and the researcher's degree of certainty about the approach to the problem. A research design could well begin with descriptive or causal research. To illustrate, a consumer satisfaction survey that is conducted annually need not begin with or include an exploratory phase.
4 Although exploratory research is generally the initial step, it need not be. Exploratory research may follow descriptive or causal research. For example, descriptive or causal research results in findings that are hard for managers to interpret. Exploratory research may provide more insights to help understand these findings.

The relationships between exploratory, descriptive and causal research are further illustrated by the following example. The example starts with a description of the environmental context of a marketing problem, shows the related marketing decision problem and the related marketing research problem, and then evaluates potential research designs that could work.

How would you like your alligator cooked, madam?

Environmental context

Supermarket X has seen a continual decline in the sales of beef over the past five years. Over this period, a number of factors have emerged which collectively may have eroded consumer confidence in the product and changed their attitudes and behaviour. The factors may be summarised as follows.

■ *Animal health scares* related to diseases in cattle throughout Europe, including BSE and foot-and-mouth disease.
■ *Health awareness* in terms of relationship between diet, nutrition and health has become a topic that is more widely debated – especially the role and benefits of red meats.
■ The *organic food* movement has significantly grown over this period, raising awareness in consumers of food production and what they deem to be 'quality' food.
■ *Animal welfare*, in terms of how animals are treated on farms, on their journey to the abattoir and in the slaughter process, has become a topic that is more widely debated.

Marketing decision problem and related marketing research problem

The marketing decision-makers in Supermarket X could perceive these issues as threats or opportunities and react in a number of ways. The following are two examples of a multitude of directions they could choose. The **marketing decision problem** could be '*should a new product be introduced'?,* i.e. beef is in decline, therefore determine which product is felt to be the 'replacement', setting a **marketing research problem** of '*to determine consumer preferences and purchase intentions for the proposed new product'.* The **marketing decision problem** could alternatively be '*should the advertising campaign be changed to allay consumer fears'?,* i.e. having a belief that 'ex-consumers' of beef can be tempted back, if only the negative connotations related to the product be addressed, setting a **marketing research problem** of '*to understand the nature of consumer fears and to test the impact of various advertising formats in terms of changing consumer attitudes'.*

Research design

The marketing research problem '*to determine consumer preferences and purchase intentions for the proposed new product'* could be tackled with a descriptive research design, further classified as a cross-sectional design. The second description of a marketing research problem, '*to understand the nature of consumer fears and to test the impact of various advertising formats in terms of changing consumer attitudes',* has two components, so it may have a more complex research design. The first component, to understand the nature of fears and ways of overcoming those fears, could be tackled with an exploratory research design. This could be followed by a causal research design where the impacts of different advertising formats are tested out.

The implication of the above connections between a marketing research problem and a research design may be that there is a simple, single design that will answer all the questions that the decision-maker faces; this is not always the case. There may be a number of research techniques that need to run simultaneously or cannot be effectively applied unless other techniques precede or succeed them. As one learns of the nature of different research techniques and their benefits and limitations, it becomes clear that techniques can connect together. By combining different research techniques, greater power in understanding and measuring consumers may be achieved, ultimately giving greater support to decision-makers.

The following research design illustrates this point by pulling together a number of research techniques to create a research design. This design should not be seen as the 'ideal' solution; it is but one of many research designs that could be used to tackle the marketing research problem.
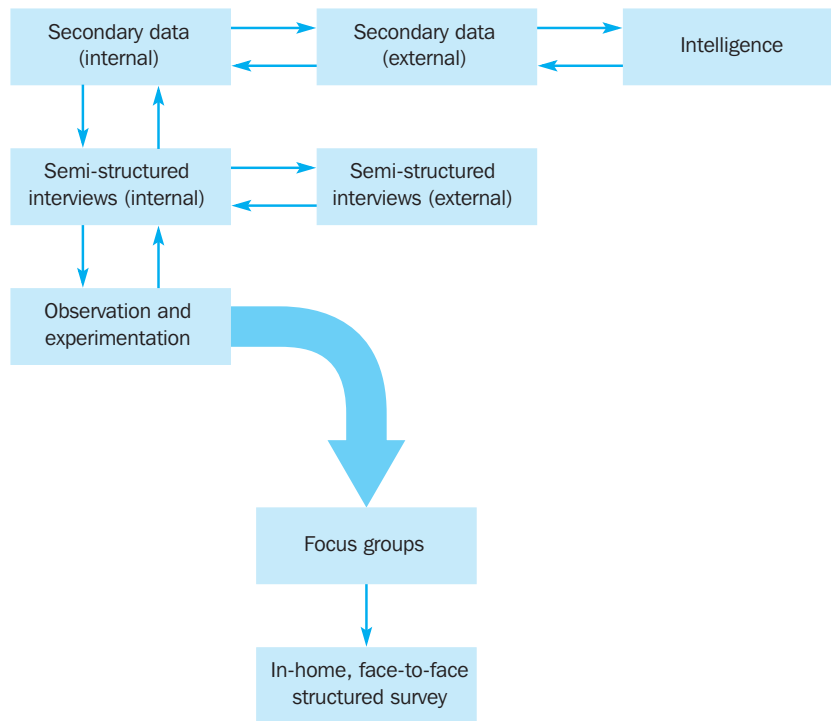
**Figure 3.4
Should a new product
be introduced to
replace beef? –
research design**

The research design starts with **internally generated secondary data** where, depending upon how data have been collected and stored, a huge array of issues could be examined. For example, sales trends of beef could be analysed to determine whether there are differences in the decline based upon beef cuts, e.g. minced beef vs. steaks, or geography, e.g. sales patterns in different stores. If different sales promotions have been tried out, the correlations with increased sales could be examined. If different alternatives such as ostrich meat, kangaroo meat or even soya-based cuts have been sold, how have they performed?

In conjunction with this stage, the researcher could collect and analyse **externally generated secondary data**. Again, a huge array of issues could be examined. For example, production levels of different alternatives to beef could be examined, looking in detail at the levels of growth or decline in new products and established products and where these products come from. In this stage, statistics that set out the wider forces that shape the market can be quantified.

In conjunction with this stage, the researcher could collect and examine different forms of **intelligence** (covered in detail in Chapter 4). For example, the writings of economists, health experts and even chefs from leading newspapers and magazines from all over the world could be evaluated. This source, being far more qualitative, may help in the interpretation of some of the analyses of internal and external secondary data. It may spark off new ideas of data to track down and new connections between data sources. Intelligence may help to locate experts who may have access to secondary data or be willing to talk in more detail.

With a very rich descriptive base of secondary data and intelligence, the researcher could conduct exploratory **semi-structured interviews** with managers within Supermarket X. These qualitative interviews could help to interpret many of the secondary data and intelligence analyses. They could help explain, for example, the reasons behind varying levels of success in advertising and sales promotion activities in regenerating beef sales. Why different alternatives were chosen and how they were promoted, positioned in the store and priced could be examined.

Concurrently, **semi-structured interviews** could be conducted with individuals who work outside Supermarket X and subsequently may be more difficult to gain access to. These interviews could be directed at experts with knowledge of health issues, culinary trends and lifestyle changes. Again, on a qualitative basis, the future of different alternatives to beef

could be examined and the comparative benefits and limitations to beef evaluated. Ideas could be generated that could be discussed and developed in interviews with managers in Supermarket X, making the semi-structured interviews concurrent and interrelated in nature.

The exploratory stage of semi-structured interviewing could generate ideas about the nature of alternatives to beef, the types of consumer they should be targeted to and the ways in which these alternatives may be promoted. Having an array of stores, researchers could set up an **experiment** to try to establish what may be influencing purchases of beef and alternatives. If consumers using the supermarket have loyalty or store cards, an **observation** of their patterns of purchase may also be made.

Having established an understanding of the variables that may affect the choice of beef and alternatives, the researcher may seek a more in-depth understanding of the most significant variables. By setting up **focus groups** with representatives from, e.g., young single persons, single parents, young families and elderly couples, chosen issues can be explored in a most creative manner. Discussion may ensue, recipes can be tasted, packages and forms of promotions can be commented upon, all resulting in a much stronger understanding of consumer preferences and purchase intentions for the proposed new product.

Finally, the researcher may wish to test out the ideas developed from the focus groups in a conclusive manner. A **face-to-face survey** may be conducted in target respondents' homes, using laptops to conduct the interviews, recording the responses but also showing video and audio recordings that may help to convey the nature and style of beef alternatives. By this stage the researcher would have a very clear understanding of the issues that are relevant to the marketing decision-makers and to the target respondents and ultimate consumers. Conclusive data that determine consumer preferences and purchase intentions for the proposed new product would be established. ■

This example can be criticised for taking too long to undertake, being too expensive and perhaps applying too many techniques that do not offer sufficient additional understanding. Such criticism cannot really be addressed without knowing the value that decision-makers may get from this decision support, compared with how much they would have to pay for it. For this illustration it does not matter, in that the intention was to show that different research techniques can support each other and can work concurrently. Decision-makers can receive interim reports and feed back their ideas to give more focus to the issues and types of respondent in subsequent stages. The example also illustrates that researchers can be very creative in their choice of techniques that combine to make up a research design.

Given that the design presented uses techniques that could be termed exploratory, descriptive and causal, the question this raises is 'how may we describe the overall research design?' The final research technique used was **conclusive**, **descriptive and single cross-sectional**, and this encapsulates the overall design. In deciding what encapsulates the overall research design, one examines the ultimate aim of an investigation, and in this case it was to describe in a conclusive manner.

An application of marketing research that utilises the whole spectrum of research designs illustrated above is in product testing. See the Companion Website for Professional Perspective 17 by Gavin Emsden of Nestlé. Gavin's article 'The acid test' describes the array of techniques used by Nestlé in product testing, i.e. the continual improvement of existing products and the development of new products. Another application that similarly utilises the whole spectrum of research designs is in marketing communications. See Professional Perspective 18 'SUMMO 2000: outdoor research on the move' by Lex van Meurs, Marcel van dar Kooi and Siebe Geert de Boer. They present the elements and findings of their research design to measure outdoor advertising reach in the Netherlands.

Potential sources of error in research designs

Several potential sources of error can affect a research design. A good research design attempts to control the various sources of error. Although these errors are discussed in detail in subsequent chapters, it is pertinent at this stage to give brief descriptions.

Where the focus of a study is a quantitative measurement, the **total error** is the variation between the true mean value in the population of the variable of interest and the observed mean value obtained in the marketing research project. As shown in Figure 3.5, total error is composed of random sampling error and non-sampling error.

Random sampling error

Random sampling error occurs because the particular sample selected is an imperfect representation of the population of interest. Random sampling error is the variation between the true mean value for the population and the true mean value for the original sample. Random sampling error is discussed further in Chapters 14 and 15.

Non-sampling error

Non-sampling errors can be attributed to sources other than sampling, and may be random or non-random. They result from a variety of reasons, including errors in problem definition, approach, scales, questionnaire design, interviewing methods, and data preparation and analysis. Non-sampling errors consist of non-response errors and response errors.

A **non-response error** arises when some of the respondents included in the sample do not respond. The primary causes of non-response are refusals and not-at-homes (see Chapter 15). Non-response will cause the net or resulting sample to be different in size or composition from the original sample. Non-response error is defined as the

Total error
The variation between the true mean value in the population of the variable of interest and the observed mean value obtained in the marketing research project.

Random sampling error
The error because the particular sample selected is an imperfect representation of the population of interest. It may be defined as the variation between the true mean value for the sample and the true mean value of the population.

Non-sampling error
An error that can be attributed to sources other than sampling and that can be random or non-random.

Non-response error
A type of non-sampling error that occurs when some of the respondents included in the sample do not respond. This error may be defined as the variation between the true mean value of the variable in the original sample and the true mean value in the net sample.

Figure 3.5
Potential sources of error in research designs

variation between the true mean value of the variable in the original sample and the true mean value in the net sample.

Response error arises when respondents give inaccurate answers or their answers are mis-recorded or mis-analysed. Response error is defined as the variation between the true mean value of the variable in the net sample and the observed mean value obtained in the marketing research project. Response errors can be made by researchers, interviewers or respondents.[21]

Errors made by the researcher include surrogate information, measurement, population definition, sampling frame and data analysis errors.

Response error
A type of non-sampling error arising from respondents who do respond but who give inaccurate answers or whose answers are mis-recorded or mis-analysed. It may be defined as a variation between the true mean value of the variable in the net sample and the observed mean value obtained in the market research project.

- *Surrogate information error* may be defined as the variation between the information needed for the marketing research problem and the information sought by the researcher. For example, instead of obtaining information on consumer choice of a new brand (needed for the marketing research problem), the researcher obtains information on consumer preferences because the choice process cannot be easily observed.
- *Measurement error* may be defined as the variation between the information sought and information generated by the measurement process employed by the researcher. While seeking to measure consumer preferences, the researcher employs a scale that measures perceptions rather than preferences.
- *Population definition error* may be defined as the variation between the actual population relevant to the problem at hand and the population as defined by the researcher. The problem of appropriately defining the population may be far from trivial, as illustrated by the case of affluent households. Their number and characteristics varied depending on the definition, underscoring the need to avoid population definition error. Depending upon the way the population of affluent households was defined, the results of this study would have varied markedly.

The fifth way of defining an affluent household: 'households that can afford to employ staff to sweep up their loose change'.

How affluent is affluent?

The population of the affluent households was defined in four different ways in a study:

1 Households with income of €50,000 or more.
2 The top 20% of households, as measured by income.
3 Households with net worth over €250,000.
4 Households with discretionary income to spend being 30% higher than that of comparable households. ■

- *Sampling frame error* may be defined as the variation between the population defined by the researcher and the population as implied by the sampling frame (list) used. For example, the telephone directory used to generate a list of telephone numbers does not accurately represent the population of potential consumers due to unlisted, disconnected and new numbers in service.
- *Data analysis error* encompasses errors that occur while raw data from questionnaires are transformed into research findings. For example, an inappropriate statistical procedure is used, resulting in incorrect interpretation and findings.

Response errors made by the interviewer include respondent selection, questioning, recording and cheating errors.

- *Respondent selection error* occurs when interviewers select respondents other than those specified by the sampling design or in a manner inconsistent with the sampling design. For example, in a readership survey, a non-reader is selected for the interview but classified as a reader of *The European* in the 15–19-year-old category in order to meet a difficult quota requirement.
- *Questioning error* denotes errors made in asking questions of the respondents or in not probing, when more information is needed. For example, while asking questions an interviewer does not use the exact wording given in the questionnaire.
- *Recording error* arises due to errors in hearing, interpreting and recording the answers given by the respondents. For example, a respondent indicates a neutral response (undecided) but the interviewer misinterprets that to mean a positive response (would buy the new brand).
- *Cheating error* arises when the interviewer fabricates answers to a part or the whole of the interview. For example, an interviewer does not ask the sensitive questions related to a respondent's debt but later fills in the answers based on personal assessment.

Response errors made by the respondent comprise inability and unwillingness errors.

- *Inability error* results from the respondent's inability to provide accurate answers. Respondents may provide inaccurate answers because of unfamiliarity, fatigue, boredom, faulty recall, question format, question content and other factors. For example, a respondent cannot recall the brand of toothpaste purchased four weeks ago.
- *Unwillingness error* arises from the respondent's unwillingness to provide accurate information. Respondents may intentionally misreport their answers because of a desire to provide socially acceptable answers, to avoid embarrassment, or to please the interviewer. For example, to impress the interviewer, a respondent intentionally says that they read *The Economist* magazine.

These sources of error are discussed in more detail in subsequent chapters; what is important here is that there are many sources of error. In formulating a research design, the researcher should attempt to minimise the total error, not just a particular source. This admonition is warranted by the general tendency among naive researchers to control sampling error with large samples. Increasing the sample size

does decrease sampling error, but it may also increase non-sampling error, for example by increasing interviewer errors. Non-sampling error is likely to be more problematic than sampling error. Sampling error can be calculated, whereas many forms of non-sampling error defy estimation. Moreover, non-sampling error has been found to be the major contributor to total error, whereas random sampling error is relatively small in magnitude.[22] The point is that total error is important. A particular type of error is important only in that it contributes to total error.

Sometimes, researchers deliberately increase a particular type of error to decrease the total error by reducing other errors. For example, suppose that a mail survey is being conducted to determine consumer preferences for purchasing fashion clothing from department stores. A large sample size has been selected to reduce sampling error. A response rate of 30% may be expected. Given the limited budget for the project, the selection of a large sample size does not allow for follow-up mailings. Past experience, however, indicates that the response rate could be increased to 45% with one follow-up mailing and to 55% with two follow-up mailings. Given the subject of the survey, non-respondents are likely to differ from respondents in many features. Hence, it may be desirable to reduce the sample size to make money available for follow-up mailings. While decreasing the sample size will increase random sampling error, the two follow-up mailings will more than offset this loss by decreasing non-response error.

International marketing research

While conducting international marketing research, it is important to realise that, given environmental differences, the research design appropriate for one country may not be suitable in another. Consider the problem of determining household attitudes towards major appliances in Holland and Saudi Arabia. While conducting exploratory research in Holland, it is appropriate to conduct focus groups jointly with male and female heads of households. It would be inappropriate to conduct such focus groups in Saudi Arabia, however. Given the traditional culture, wives are unlikely to participate freely in the presence of their husbands. It would be more useful to conduct one-on-one in-depth interviews with both male and female heads of households being included in the sample. An understanding of environmental influences can affect the choice and application of individual research techniques. Ultimately, the rules of using different techniques to build up an understanding of consumers remains the same, regardless of geographic boundaries, as illustrated in the following example.

e x a m p l e | **Chinese sport a favourable attitude towards sporting goods**[23]

A marketing research study was performed to gain insight and understanding into the behaviour of consumers in China in relation to sports and sporting goods. The research design included both exploratory and conclusive research. Exploratory research was necessary to gain an understanding of the social, cultural, economic and legal environment of China. The researcher also utilised both secondary data and one-to-one depth interviews with retail outlet managers to gather this information.

For the conclusive research, a single cross-sectional design was used. A survey was developed and 4,000 questionnaires were distributed to respondents in 10 Chinese cities. Of these, 54% were completed and returned. A single cross-sectional design was appropriate because the researcher wanted to gain an understanding of the current state of consumers' attitudes, motivations and behaviours in relation to sporting goods.

As well as findings such as the strong brand recognition of Adidas, Nike, Asics and Reebok, specific social, cultural and environmental trends were also identified. For instance, China is becoming more fitness-conscious. China is also becoming a more casual society, which is leading to the use of sporting goods for facilitation of the casual lifestyle. Also, the Chinese tend to shop a lot on Sundays and the Chinese wife tends to make family purchasing decisions. ■

In many countries, particularly developing countries, consumer panels have not been developed, making it difficult to conduct descriptive longitudinal research. Likewise, in many countries the marketing support infrastructure – that is, retailing, wholesaling, advertising and promotional infrastructure – is lacking, making it infeasible to implement a causal design involving a field experiment. In formulating a research design, considerable effort is required to ensure the equivalence and comparability of secondary and primary data obtained from different countries. In the context of collecting primary data, qualitative research, survey methods, scaling techniques, questionnaire design and sampling considerations are particularly important. These topics are discussed in more detail in subsequent chapters.

Ethics in marketing research

During the research design stage, not only are the concerns of the researcher and the client involved, but the rights of the respondents also must be respected. Although normally there is no direct contact between the respondents and the other stakeholders (client and researcher) during the research design phase, this is the stage when decisions with ethical ramifications, such as using hidden video or audio tape recorders, are made.

The basic question of the type of research design which should be adopted (i.e. descriptive or causal, cross-sectional or longitudinal) has ethical overtones. For example, when studying brand switching in toothpaste purchases, a longitudinal design is the only actual way to assess changes in an individual respondent's brand

Researchers cannot just approach children and ask them what is in their shopping basket.

choice. A research firm that has not conducted many longitudinal studies may try to justify the use of a cross-sectional design. Is this ethical?

Researchers must ensure that the research design utilised will provide the information needed to address the marketing research problem that has been identified. The client should have the integrity not to misrepresent the project and should describe the constraints under which the researcher must operate and not make unreasonable demands. Longitudinal research takes time. Descriptive research might require interviewing customers. If time is an issue, or if customer contact has to be restricted, the client should make these constraints known at the start of the project.

Equally important, the responsibilities to the respondents must not be overlooked. The researcher should design the study so as not to violate the respondents' right to safety, right to privacy, or right to choose. Furthermore, the client must not abuse power to jeopardise the anonymity of the respondents.

As well as their general code of conduct, ESOMAR produces a series of guidelines that are specific to particular research techniques (Internet research), types of respondent (children) and types of industry (pharmaceutical).

To see how ESOMAR guides the ethical practices of the marketing research industry in Europe, visit **www.esomar.nl/codes_and_guidelines.html**.

To see how marketing research associations in individual countries throughout the world guide the ethical practices of domestic marketing research, visit **www.esomar.nl/assocs/mr_associations.html**.

Internet and computer applications

The Internet can facilitate the implementation of different types of research designs.

Exploratory research

If an exploratory research design is to be utilised, forums, chat rooms or newsgroups can be used to discuss a particular topic to great depth. Files can be exchanged that can include moving images and sounds, allowing questions and probes to be built around this material. Formal focus groups may be conducted with experts or individuals representing target groups, all on a global basis if needed. Illustrations of qualitative interviews using the Internet are presented in Chapters 6 to 8.

Conclusive research

Many descriptive studies utilise secondary data in defining the nature of a problem, as a technique in its own right and as a means to develop sampling plans. The use of the Internet for these purposes is discussed in Chapter 4. In primary data collection, the Internet can be used for surveys (discussed in Chapter 10) and in panels (Chapters 4 and 10). The use of the Internet for causal research designs is discussed in Chapter 11.

In addition to Internet applications, computers can also help to control total error. By using computers, researchers can see how the various sources of error will affect the results and what levels of errors may be acceptable. It is relatively easy to estimate random sampling error when probability sampling schemes are used. Estimating the impact of various non-sampling errors, however, is much more problematic. Simulation can be conducted to determine how the distributions and levels of various non-sampling errors will affect final results.[24]

Summary

A research design is a framework or blueprint for conducting the marketing research project. It specifies the details of how the project should be conducted in order to fulfil set research objectives. Research designs may be broadly classified as exploratory or conclusive. The primary purpose of exploratory research is to develop understanding and provide insights. Conclusive research is conducted to measure and describe phenomena, test specific hypotheses and examine specific relationships. Conclusive research may be either descriptive or causal. The findings from both exploratory and conclusive research can be used as input into managerial decision-making.

The major objective of descriptive research is to describe market characteristics or functions. Descriptive research can be classified into cross-sectional and longitudinal research. Cross-sectional designs involve the collection of information from a sample of population elements at a single point in time. These designs can be further classified as single cross-sectional or multiple cross-sectional designs. In contrast, in longitudinal designs repeated measurements are taken on a fixed sample. Causal research is designed for the primary purpose of obtaining evidence about cause-and-effect (causal) relationships.

Many research designs combine techniques that can be classified as exploratory, descriptive and causal. In cases where there is an array of interrelating techniques, the researcher should examine the ultimate aim of an investigation, and decide what encapsulates the overall research design, i.e. a desire to explore, describe or experiment.

A research design consists of six components. Error can be associated with any of these components. The total error is composed of random sampling error and non-sampling error. Non-sampling error consists of non-response and response errors. Response error encompasses errors made by researchers, interviewers and respondents. In formulating a research design when conducting international marketing research, considerable effort is required to ensure the equivalence and comparability of secondary and primary data obtained from different countries. In terms of ethical issues, the researchers must ensure that the research design used will provide the information sought and that the information sought is the information needed by the client. The client should have the integrity not to misrepresent the project and should describe the situation within which the researcher must operate and must not make unreasonable demands. Every precaution should be taken to ensure the respondents' or subjects' right to safety, right to privacy, or right to choose.

Questions

1 Define research design in your own words.

2 What expectations do marketing decision-makers have of research designs?

3 How does the subject of enquiry as seen by potential research respondents affect research design?

4 How does formulating a research design differ from developing an approach to a problem?

5 Differentiate between exploratory and conclusive research.

6 What are the major purposes for which exploratory research is conducted?

7 Describe how quantitative techniques may be used in exploratory research.

8 What are the major purposes for which descriptive research is conducted?

9 Discuss the advantages and disadvantages of panels.

10 Compare and contrast cross-sectional and longitudinal designs.

11 Describe cohort analysis. Why is it of special interest?

12 What is a causal research design? What is its purpose?

13 What is the relationship between exploratory, descriptive and causal research?

14 What potential sources of error can affect a research design?

15 Why is it important to minimise total error rather than any particular source of error?

Notes

1 Acreman, S. and Pegram, B., 'Getting to know you', *Research* (November 1999), 36–41.

2 Boyd, H.W. Jr, Westfall, R. and Stasch, S.F., *Marketing Research, Text and Cases*, 7th edn (Homewood, IL: Irwin, 1989), 49.

3 Adapted from Cooper, P. and Braithwaite, A., 'Qualitative technology - new perspectives on measurement and meaning through qualitative research', *Market Research Society Conference, 2nd pre-conference workshop*, 1979.

4 Lee, H., Lindquist, J.D. and Acito, F., 'Managers' evaluation of research design and its impact on the use of research: an experimental approach', *Journal of Business Research* 39(3) (July 1997), 231–40; Wilson, R.D., 'Research design: qualitative and quantitative approaches', *Journal of Marketing Research* 33(2) (May 1996), 252–5.

5 For an example of descriptive research, see Song, S.M. and Perry, M.E., 'The determinants of Japanese new product success', *Journal of Marketing Research* 34 (February 1997), 64–76.

6 'Coupon use low among young, working women', *Marketing News* (10 April 1987), 24.

7 'Eating behaviour trends revealed in Pillsbury study', *Quirk's Marketing Research Review* (June–July 1988), 14–15, 39, 44.

8 Misra, R. and Panigrahi, B., 'Changes in attitudes toward women: a cohort analysis', *International Journal of Sociology & Social Policy* 15(6) (1995), 1–20; Glenn, N.D., *Cohort Analysis* (Beverly Hills, CA: Sage Publications, 1981).

9 Rentz, J.O., Reynolds, F.D. and Stout, R.G., 'Analyzing changing consumption patterns with cohort analysis', *Journal of Marketing Research* 20 (February 1983), 12–20. See also Rentz, J.O. and Reynolds, F.D., 'Forecasting the effects of an aging population on product consumption: an age-period-cohort framework', *Journal of Marketing Research* (August 1991), 355–60.

10 For example, see Elrod, T. and Keane, M.P., 'A factor-analytic probit model for representing the market structure in panel data', *Journal of Marketing Research* 32 (February 1995), 1–16.

11 For recent applications of panel data, see Sivakumar, K. and Raj, S.P., 'Quality tier competition: how price change influences brand choice and category choice', *Journal of Marketing* 61(3) (July 1997), 71–84. For a basic treatment see Markus, G.B., *Analyzing Panel Data* (Beverly Hills, CA: Sage Publications, 1979).

12 Table 3.7 can also be viewed as a transition matrix. It depicts the brand-buying changes from period to period. Knowing the proportion of consumers who switch allows for early prediction of the ultimate success of a new product or change in market strategy. See Sudman, S. and Ferber, R., *Consumer Panels* (Chicago, IL: American Marketing Association, 1979), 19–27.

13 Johnston, M.W., Parasuraman, A., Futrell, C.M. and Black, W.C., 'A longitudinal assessment of the impact of selected organizational influences on salespeople's organizational commitment during early employment', *Journal of Marketing Research* (August 1990), 333–44.

14 Sudman, S. and Ferber, R., 'A comparison of alternative procedures for collecting consumer expenditure data for frequently purchased products', and Wright, R.A., Beisel, R.H., Oliver, J.D. and Gerzowski, M.C., 'The use of a multiple entry diary in a panel study on health care', both in Ferber, R. (ed.), *Readings in Survey Research* (Chicago, IL: American Marketing Association, 1978), 487–502, 503-12; Wind, Y. and Lerner, D., 'On the measurement of purchase data: surveys versus purchase diaries', *Journal of Marketing Research* 16 (February 1979), 39–47; and McKenzie, J., 'The accuracy of telephone call data by diary methods', *Journal of Marketing Research* 20 (November 1983), 417–27.

15 Coupe, R.T. and Onudu, N.M., 'Evaluating the impact of CASE: an empirical comparison of retrospective and cross-sectional survey approaches', *European Journal of Information Systems* 6(1) (March 1997), 15–24. Sudman, S. and Ferber, R., *Consumer Panels* (Chicago, IL: American Marketing Association, 1979), 19–27.

16 Van Den Berg, D.G., Lindeboom, M. and Ridder, G., 'Attribution in longitudinal panel data and the empirical analysis of dynamic labour market behaviour', *Journal of Applied Econometrics* 9(4) (October–December 1994), 421–35; Winer, R.S., 'Attrition bias in econometric models estimated with panel data', *Journal of Marketing Research* 20 (May 1983), 177–86.

17 Maytas, L. and Sevestre, P. (eds), *The Econometrics of Panel Data, A Handbook of the Theory with Applications* (Norwell, MA: Kluwer Academic Publishers, 1996).

18 Hulland, J., Ho, Y. and Lam, S., 'Use of causal models in marketing research: a review', *International Journal of Research in Marketing* 13(2) (April 1996), 181–97. See also Cox, K.K. and

Enis, B.M., *Experimentation for Marketing Decisions* (Scranton, PA: International Textbook, 1969), 5.

19 For an application of causal research see Unnava, R.H., Bumkrant, R.E. and Erevelles, S., 'Effects of presentation order and communication modality on recall and attitude', *Journal of Consumer Research* (21 December 1994), 481–90.

20 Cooper, P.J., Diamond, I. and High, S., 'Choosing and using contraceptives: integrating qualitative and quantitative research methods in family planning', *Journal of the Market Research Society* 35 (October 1993), 325–40.

21 Dutka, S. and Frankel, L.R., 'Measuring response error', *Journal of Advertising Research* 37(1) (January/February 1997), 33–9.

22 Rollere, M.R., 'Control is elusive in research design', *Marketing News* 31(19) (15 September 1997), 17; Corlett, T. , 'Sampling errors in practice', *Journal of the Market Research Society* 38(4) (April 1997), 307–18.

23 Geng, L. *et al.*, 'Sports marketing strategy: a consumer behaviour case analysis in China', *Multinational Business Review* (Spring 1997), 147–54; Gallup Organisation, '1997 Survey: The People's Republic of China. Consumers Attitudes & Lifestyle Trends', Website www.gallup.com/poll/special/china.html.

24 Leung, J.W.K. and Lai, K.K., 'A structured methodology to build discrete-event simulation models', *Asia Pacific Journal of Operations Research* 14(1) (May 1997), 19-37; Malhotra, N.K., 'An approach to the measurement of consumer preferences using limited information', *Journal of Marketing Research* 23 (February 1986), 33-40; Malhotra, N.K., 'Analyzing marketing research data with incomplete information on the dependent variable', *Journal of Marketing Research* 24 (February 1987), 74–84.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Secondary data collection and analysis

Objectives

After reading this chapter, you should be able to:

1 define the nature and scope of secondary data and distinguish secondary data from primary data;

2 analyse the advantages and disadvantages of secondary data and their uses in the various steps of the marketing research process;

3 evaluate secondary data using the criteria of specifications, error, currency, objectives, nature and dependability;

4 describe in detail the different sources of secondary data, focusing upon external sources in the form of published materials, and syndicated services;

5 discuss in detail the syndicated sources of secondary data, including household and consumer data obtained via surveys, mail diary panels and electronic scanner services, as well as institutional data related to retailers, wholesalers and industrial or service firms;

6 explain the need to use multiple sources of secondary data and describe single-source data;

7 identify and evaluate the sources of secondary data useful in international marketing research;

8 understand the ethical issues involved in the use of secondary data.

The act of sourcing, evaluating and analysing secondary data can realise great insights for decision-makers. It is also vital to successful problem diagnosis, sample planning and collection of primary data.

Overview

The collection and analysis of secondary data help to define the marketing research problem and develop an approach. In addition, before collecting primary data, the researcher should locate and analyse relevant secondary data. Thus, secondary data can be an essential component of a successful research design. Secondary data can help in sample designs and in the details of primary research methods. In some projects, research may be largely confined to the analysis of secondary data because some routine problems may be addressed based only on secondary data. In addition, given the huge explosion of secondary data sources available, sufficient data may be accessed to solve a particular marketing research problem.

This chapter discusses the distinction between primary data, secondary data and marketing intelligence. The advantages and disadvantages of secondary data are considered, and criteria for evaluating secondary data are presented, along with a classification of secondary data. Internal secondary data are described, and major sources of external secondary data – such as published materials, online and offline databases, and syndicated services – are also discussed The sources of secondary data useful in international marketing research are discussed. Several ethical issues that arise in the use of secondary data are identified.

To begin with, we present an example that illustrates the nature of secondary data, how it may be evaluated, and its relationship to primary data collection.

e x a m p l e

Flying high on secondary data[1]

Money magazine recently published the results of a study conducted to uncover the airline characteristics that consumers consider most important. In order of importance, these characteristics were safety, price, baggage handling, on-time performance, customer service, ease of reservations and ticketing, comfort, frequent flyer schemes and food.

If Air France was considering conducting a marketing research study to identify characteristics of its service that should be improved, this article might be a useful source of secondary data. Before using the data, Air France should evaluate them according to several criteria.

First, the research design used to collect the data should be examined. This *Money* magazine article includes a section that details the research design used in the study. *Money* used a face-to-face survey of 1,017 'frequent flyers'. The results of the survey had a margin of error of 3%. Air France would have to decide whether 'frequent flyers' in the USA could be generalised to the population they wish to understand, whether 1,017 was a sufficient sample size for their purposes and whether a margin of error of 3% was acceptable. In addition, Air France should evaluate what type of response or non-response errors may have occurred in the data collection or analysis process.

The currency of the data and objective of the study would be important to Air France in deciding whether to utilise this article as a source of secondary data. Air France would also need to look at the nature and dependability of the data. For example, they would need to examine how the nine choice criteria were defined. If the criterion price was measured in terms of fare per kilometre, is this a meaningful and acceptable definition to decision-makers at Air France? With regard to dependability, Air France would need to evaluate the reputation of *Money* magazine and of ICR, the research company hired by *Money* to undertake the survey. They would also need to recognise the fact that *Money* used secondary data in its study; how dependable are the sources they used?

The *Money* magazine article might be useful as a starting place for a marketing research project for Air France. It could be helpful in formulating the nature of decision-making problems and associated research objectives. There may be limitations in regard to reliability, dependability or even how generalisable it may be to Air France's target consumers. Many lessons and ideas may be generated from this article that may lead to other secondary data sources and in the design of a well-focused primary data collection. ■

Defining primary data, secondary data and marketing intelligence

Primary data
Data originated by the researcher specifically to address the research problem.

Secondary data
Data collected for some purpose other than the problem at hand.

Marketing intelligence
Qualified observations of events and developments in the marketing environment.

Primary data are originated by a researcher for the specific purpose of addressing the problem at hand. They are individually tailored for the decision-makers of organisations that pay for well-focused and exclusive support. Compared with readily available data from a variety of sources, this tailoring means higher costs and a longer time frame in collecting and analysing the data.

Secondary data are data that have already been collected for purposes other than the problem at hand. At face value this definition seems straightforward, especially when contrasted to the definition of primary data. However, many researchers confuse the term, or quite rightly see some overlap with marketing intelligence.

Marketing intelligence can be defined as 'qualified observations of events and developments in the marketing environment'. The use of the word 'observations' is presented in a wide sense to include a variety of types of data, broadly concerned with 'environmental scanning'.[2] In essence, though, marketing intelligence is based upon data that in many instances have been collected for purposes other than the problem at hand. To clarify this overlap in definitions, Table 4.1 compares secondary data with marketing intelligence through a variety of characteristics.

Table 4.1 A comparison of secondary data and marketing intelligence[3]

Characteristic	Secondary data	Marketing intelligence
Structure	Specifications and research design tend to be apparent	Can be poorly structured; no universal conventions of reporting
Availability	Tend to have regular updates	Irregular availability
Sources	Generated in-house and from organisations with research prowess	Generated in-house and from unofficial sources
Data type	Tend to be quantitative; many issues need qualitative interpretation	Tends to be qualitative; many issues difficult to quantify
Source credibility	Tend to be from reputable and trustworthy research sources	Questionable credibility; can be generated from a broad spectrum of credibility
Terms of reference	Tend to have clear definitions of what is being measured	Ambiguous definitions; difficult to compare over different studies
Analysis	Mostly conventional quantitative techniques	Opinion based, interpretative
Ethics	In-company data gathering may be covered by data protection acts; externally generated data may be covered by research codes of conduct, e.g. ESOMAR	Some techniques may be seen as industrial espionage – though there is an ethical code produced by the Society of Competitive Intelligence Professionals

Note in the above comparisons the repeated use of the word 'tend'. The boundaries between the two are not absolutely rigid. Consider the example at the start of this chapter, an article published in *Money* magazine. The journalist may have collected, analysed and presented quantitative data to support their qualitative interpretation of the future developments of a market. The data they use and present may come from credible sources and be correctly analysed, but what about their choice of data to support their argument? Other sources of data that may contradict their view may be ignored. The data they present can be seen as a secondary data source and interpreted in its own right by a researcher. The interpretation and argument of the journalist can

be seen as intelligence and have some credibility. In its entirety, such an article has elements of both secondary data and marketing intelligence, and it may be impossible to pull them apart as mutually exclusive components.

As will become apparent in this chapter, there are clear criteria for evaluating the accuracy of secondary data, which tend to be of a quantitative nature. Marketing intelligence is more difficult to evaluate but this does not mean that it has less value to decision-makers or researchers. Certain marketing phenomena cannot be formally measured; researchers may not be able to gain access to conduct research, or the rapid unfolding of events means that it is impracticable to conduct research. The following example illustrates the importance of intelligence to many companies.

example

Behind enemy lines [4]

Robin Kirkby, Director of European Consulting for intelligence specialist Fuld & Company, says there are three principal factors driving investment in intelligence.

'The Internet, globalisation and higher expectations from customers are all putting companies under more pressure to differentiate themselves from the competition. It's frustrating that intelligence gets associated with spying; it's actually a highly ethical activity, focused on underlying competitive dynamics and planning future change.'

According to research by The Futures Group (TFG), 60% of companies have an organised system for collecting competitive intelligence, while 82% of companies with revenues over €10bn make systematic use of it. TFG ranked the leading eight users of competitor intelligence as:

1   Microsoft
2   Motorola
3   IBM
4   Procter & Gamble
5= General Electric
5= Hewlett-Packard
7= Coca-Cola
7= Intel ■

Such widespread use of intelligence in major organisations means it has some role to play in supporting decision-makers, but it has many limitations, which are apparent in Table 4.1. In the development of better-founded information support, credible support can come from the creative collection and evaluation of secondary data. This requires researchers to connect and validate different data sources, ultimately leading to decision-maker support in its own right and support of more focused primary data collection. As this chapter and Chapter 5 unfold, examples of different types of secondary data will emerge and the applications of secondary data will become apparent.

Advantages and uses of secondary data

Secondary data offer several advantages over primary data. Secondary data are easily accessible, relatively inexpensive and quickly obtained. Some secondary data, such as those provided by the National Censuses, are available on topics where it would not be feasible for a firm to collect primary data. Although it is rare for secondary data to provide all the answers to a non-routine research problem, such data can be useful in a variety of ways.[5] Secondary data can help you:

1 Diagnose the research problem
2 Develop an approach to the problem
3 Develop a sampling plan
4 Formulate an appropriate research design (for example, by identifying the key variables to measure or understand)

5 Answer certain research questions and test some hypotheses

6 Interpret primary data with more insight

7 Validate qualitative research findings.

Given these advantages and uses of secondary data, we state the following general rule:

> *Examination of available secondary data is a prerequisite to the collection of primary data. Start with secondary data. Proceed to primary data only when the secondary data sources have been exhausted or yield marginal returns.*

The rich dividends obtained by following this rule are illustrated in the example at the start of this chapter. It shows that the collection and analysis of even one relevant secondary data source can provide valuable insights. The decision-maker and researcher can use the ideas generated in secondary data as a very strong foundation to primary data design and collection. However, the researcher should be cautious in using secondary data, because they have some limitations and disadvantages.

Disadvantages of secondary data

Because secondary data have been collected for purposes other than the problem at hand, their usefulness to the current problem may be limited in several important ways, including relevance and accuracy. The objectives, nature and methods used to collect the secondary data may not be appropriate to the present situation. Also, secondary data may be lacking in accuracy or may not be completely current or dependable. Before using secondary data, it is important to evaluate them according to a series of factors.[6] These factors are discussed in more detail in the following section.

Criteria for evaluating secondary data

The quality of secondary data should be routinely evaluated, using the criteria presented in Table 4.2 and discussion in the following sections.[7]

Specifications and research design

The specifications or the research design used to collect the data should be critically examined to identify possible sources of bias. Such design considerations include size and nature of the sample, response rate and quality, questionnaire design and administration, procedures used for fieldwork, and data analysis and reporting procedures. These checks provide information on the reliability and validity (these concepts will be further developed in Chapter 13) of the data and help determine whether they can be generalised to the problem at hand. The reliability and validity can be further ascertained by an examination of the error, currency, objectives, nature and dependability associated with the secondary data.

Error and accuracy

The researcher must determine whether the data are accurate enough for the purposes of the present study. Secondary data can have a number of sources of error or inaccuracy, including errors in the approach, research design, sampling, data collection, analysis, and reporting stages of the project. Moreover, it is difficult to evaluate the accuracy of secondary data because the researcher did not participate in the research. One approach is to find multiple sources of data if possible, and compare them using standard statistical procedures.

Table 4.2 Criteria for evaluating secondary data

Criteria	Issues	Remarks
Specifications and research design	■ Data collection method ■ Response rate ■ Population definition ■ Sampling method ■ Sample size ■ Questionnaire design ■ Fieldwork ■ Data analysis	Data should be reliable, valid and generalisable to the problem at hand.
Error and accuracy	Examine errors in ■ Approach ■ Research design ■ Sampling ■ Data collection ■ Data analysis ■ Reporting	Assess accuracy by comparing data from different sources.
Currency	Time lag between collection and publication. Frequency of updates	Census data are periodically updated by syndicated firms.
Objective	Why were the data collected	The objective will determine the relevance of data.
Nature	■ Definition of key variables ■ Units of measurement ■ Categories used ■ Relationships examined	Reconfigure the data to increase their usefulness, if possible.
Dependability	Source: ■ Expertise ■ Credibility ■ Reputation ■ Trustworthiness	Preference should be afforded to an original rather than an acquired source.

example

Number crunch[8]

In December 1997, the Audit Bureau of Circulations (ABC) met UK newspaper publishers and major media buyers from the Institute of Practitioners in Advertising. The meeting aimed to thrash out a formula that could restore ABC's credibility as a trading currency.

Most observers agreed that the ABC's troubles were a direct result of squabbling between media owners. As circulations have continued to slide, the press barons have fought to hold on to their market share through price cuts, promotions and enhanced editorial packages. This has introduced an unprecedented volatility into their sales figures. Not content with trumpeting their own gains, some companies have sought to show up the deficiencies in their rivals' sales figures.

Figures under fire

The argument is best understood through a straightforward example of what is at stake. Let us take the October 1997 ABC figure for *The Times*, which was 814,899. That figure was a monthly circulation average which, prior to the recent dispute, would have been the only official benchmark that agencies used as a negotiating point with press owners (though they turn to data from the National Readership Survey and the Target Group Index to argue their case). At the heart of the dispute has been how that monthly figure is comprised. For example, were all the issues sold at the full price or were some given away cheaply as part of a subscription or promotional offer? Were any sold or given in bulk to an airliner or retailer, and if so how many? What about papers sold to Eire or Spain? Were they included in the total and if so how could that be justified as a piece of credible advertising data?

Another hot issue was the reliability of the monthly figure. Advertisers were dissatisfied with a number that they believe fails to reflect the reality of what they were buying. Director of press buying at The Media Centre, Tim Armes, says 'we'd like to know what each paper sells

daily and we'd like to know week to week fluctuations. The papers all boast about Saturday but keep quiet about Tuesday and Thursday. If one day is dramatically higher than the average, you don't have to be a brain surgeon to realise the others are lower.' ■

As this example indicates, the accuracy of secondary data can vary, particularly if they relate to phenomena that are subject to change. Moreover, data obtained from different sources may not agree. In these cases, the researcher should verify the accuracy of secondary data by conducting pilot studies or by other appropriate methods. Often, by exercising creativity this can be done without much expense or effort.

Currency: when the data were collected

Secondary data may not be current and the time lag between data collection and publication may be long, as is the case with much census data. Moreover, the data may not be updated frequently enough for the purpose of the problem at hand. Decision-makers require current data; therefore, the value of secondary data is diminished as they become dated. For instance, although the Census of Population data are comprehensive, they may not be applicable to major cities in which the population has changed rapidly during the last two years. Likewise, in the GlobalCash Project, the lists of the largest companies in Europe, used to decide who should be surveyed, have to be updated to reflect changes that take place in the two years between each study.

Objective: the purpose for which the data were collected

Data are invariably collected with some objective in mind, and a fundamental question to ask is why the data were collected in the first place. The objective for collecting data will ultimately determine the purpose for which that information is relevant and useful. Data collected with a specific objective in mind may not be appropriate in another situation. In the example at the start of this chapter, the sample surveyed by *Money* magazine was made up of 'frequent flyers'. The objective of the study was 'to uncover the airline characteristics consumers consider most important'. Air France, however, may wish to target 'business class' flyers and 'to uncover perceptions related to trade-offs made in customer service–price–safety'. Even though there may be identical questions used in both studies, the target respondents may be different, the rationale for the study presented to respondents will be different, and ultimately the 'state of mind' respondents may be in when they come to comparable questions will be different. The *Money* survey was conducted for entirely different objectives from those Air France have for their study. The findings from the *Money* survey may not directly support decision-making at Air France, though they may help to define who Air France should talk to and what questions they should put to them.

Nature: the content of the data

The nature, or content, of the data should be examined with special attention to the definition of key variables, the units of measurement, the categories used and the relationships examined. If the key variables have not been defined or are defined in a manner inconsistent with the researcher's definition, then the usefulness of the data is limited. Consider, for example, secondary data on consumer preferences for TV programmes. To use this information, it is important to know how preference for programmes was defined. Was it defined in terms of the programme watched most often, the one considered most needed, most enjoyable, most informative, or the programme of greatest service to the community?

Likewise, secondary data may be measured in units that may not be appropriate for the current problem. For example, income may be measured by individual, family, household or spending unit and could be gross or net after taxes and deductions. Income may be classified into categories that are different from research needs. If the

researcher is interested in high-income consumers with gross annual household incomes of over €120,000, secondary data with income categories of less than €20,000, €20,001–€50,000, €50,001–€75,000 and more than €75,000 will not be of use. Determining the measurement of variables such as income may be a complex task, requiring the wording of the definition of income to be precise. Finally, the relationships examined should be taken into account in evaluating the nature of data. If, for example, actual behaviour is of interest, then data inferring behaviour from self-reported attitudinal information may have limited usefulness. Sometimes it is possible to reconfigure the available data – for example, to convert the units of measurement – so that the resulting data are more useful to the problem at hand.

Dependability: how dependable are the data?

An overall indication of the dependability of data may be obtained by examining the expertise, credibility, reputation and trustworthiness of the source. This information can be obtained by checking with others who have used the information provided by the source. Data published to promote sales, to advance specific interests, or to carry on propaganda should be viewed with suspicion. The same may be said of data published anonymously or in a form that attempts to hide the details of the data collection research design and process. It is also pertinent to examine whether the secondary data came from an original source, one that generated the data, or an acquired source, one that procured the data from an original source. Generally, secondary data should be secured from an original rather than an acquired source. There are at least two reasons for this rule: first, an original source is the one that specifies the details of the data collection research design, and second, an original source is likely to be more accurate and complete than a surrogate source.

Classification of secondary data

Internal data
Data available within the organisation for whom the research is being conducted.

Figure 4.1 presents a classification of secondary data. Secondary data may be classified as either internal or external. **Internal data** are those generated within the organisation for which the research is being conducted. This information may be available in a ready-to-use format, such as information routinely supplied by the management decision
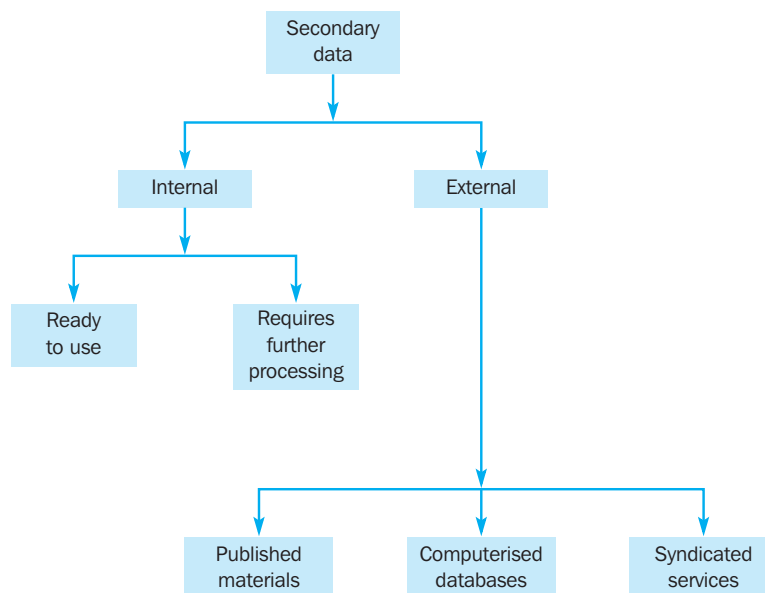
Figure 4.1
A classification of secondary data

support system. On the other hand, these data may exist within the organisation but may require considerable processing before they are useful to the researcher. For example, a variety of information can be found on sales invoices. Yet this information may not be easily accessed; further processing may be required to extract it. Secondary data generated from internal sources will be examined in more detail in Chapter 5. **External data**, on the other hand, are those generated by sources outside the organisation. These data may exist in the form of published material, online databases, or information made available by syndicated services. Externally generated secondary data may be more difficult to access, more expensive, and more difficult to evaluate for its accuracy, in comparison with internal secondary data. These factors mean that, before collecting external secondary data, it is useful to analyse readily available internal secondary data.

External data
Data that originate outside the organisation.

Published external secondary sources

Sources of published external secondary data include local authorities, regional and national governments, the EC, non-profit organisations (e.g. Chambers of Commerce), trade associations and professional organisations, commercial publishers, investment brokerage firms, and professional marketing research firms.[9] In fact, such a quantity of data is available that the researcher can be overwhelmed. Therefore, it is important to classify published sources (see Figure 4.2). Published external sources may be broadly classified as general business data or government data. General business sources comprise guides, directories, indexes and statistical data. Government sources may be broadly categorised as census data and other publications. These data types are discussed further with specific sources used as examples.

General business sources

Businesses publish a lot of information in the form of books, periodicals, journals, newspapers, magazines, reports and trade literature. This information can be located by using guides, directories and indexes. Sources are also available to identify statistical data.

Guides. Guides are an excellent source of standard or recurring information. A guide may help identify other important sources of directories, trade associations and trade publications. Guides are one of the first sources a researcher should consult. The following example illustrates the use of the Electronic Buyer's Guide.
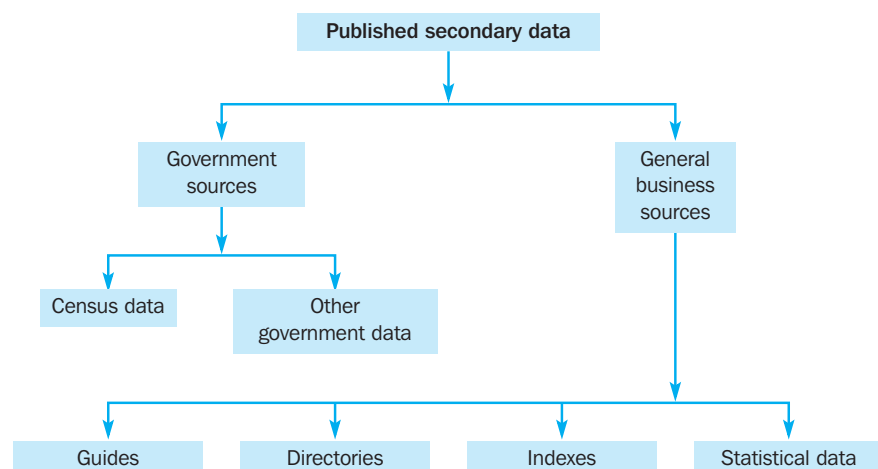
Figure 4.2
A classification of published secondary sources

Keeping up to data[10]

The market for business information is growing and the shift towards electronic sources continues. The choice of new media for directory publishers includes offline methods, such as CD-ROM, and online, such as the Internet. In addition, while online systems account for most information sales, the use of CDs is growing faster.

Access to up-to-date information is a big attraction. There is a perception that on-screen information is more up-to-date than a book. However, many CD-ROM directories are not updated any more frequently than their paper equivalents. So, while information on the Internet can be kept up-to-date, CD-ROMs come into their own when storing large amounts of information.

Miller Freeman launched the Electronic Buyer's Guide in 1996 and by October 1997 it was on the Internet (see **www.mfi.com**). There is some information available free of charge; if more information is wanted then there are charges. Alternatively, anyone who buys the guide on CD-ROM is given a password that gives them free access to the Internet directory. ■

Directories. Directories are helpful for identifying individuals or organisations that collect specific data. An example of a directory that you can examine on the Internet is the Central and Eastern European Business Directory. This interactive site provides current information on businesses and organisations in 24 central and eastern European countries (**www.ceebd.co.uk**). Another example is Europages, a reference business directory in Europe that classifies 500,000 companies in 30 European countries. Again, this can be accessed through the Internet and is available in English, French, German, Italian and Spanish versions (**www.europages.com**).

Indexes. It is possible to locate information on a particular topic in several different publications by using an index and abstracts. Indexes and abstracts, therefore, can increase the efficiency of the search process. Several indexes and abstracts are available for both academic and business sources. Examples of newspaper indexes include the *Financial Times Index* (**www.news.ft.com**), *Le Monde Index* (**www.le-monde.fr**), and the Japanese Business News online, *The Nikkei Weekly* (**www.nikkei.co.jp**). These indexes allow researchers to identify sources of particular topics, industries and individuals.

An example of a marketing index is the *Marketing Surveys Index* published by Euromonitor (**www.euromonitor.com**). This is a most comprehensive and up-to-date directory of business research on European and world markets. The index contains details of published research reports, a brief summary of its contents, a keyword index to the markets and products covered, and bibliographic details of the report.

A particularly useful abstract for marketing researchers is the *Market Research Abstracts* published by the Market Research Society in Britain (**www.mrs.org.uk**). Major European and American journals that relate to marketing research are reviewed and an abstract of each article is presented. The abstract is published twice a year. It is divided into sections that cover survey techniques; statistics, models and forecasting; attitude and behaviour research; psychographics, personality and social psychology; advertising and media research; applications of research; industrial market research; market research and general applications; and new product development. Such an abstract allows the researcher to quickly identify and evaluate the worth of journal papers that are relevant to their particular study.

Non-government statistical data. Published statistical data are of great interest to researchers. Graphic and statistical analyses can be performed on these data to draw important insights. Examples of non-governmental statistical data include trade associations such as the Swedish Tourism Trade Association (**www.sverigeturism.se**). The Swedish Information 'Smorgasbord' is a large single source of information in English on Sweden, Swedish provinces, nature, culture, lifestyle, society and industry. Another

example is Euromonitor (**www.euromonitor.com**), which publishes monthly market research journals covering subjects under the headings of Market Research Europe, Market Research GB, Market Research International and Retail Monitor International.

The United Nations provides an example of an organisation with a Statistics Division that provides a wide range of statistical outputs on a global basis (**www.un.org/depts/unsd**). The Statistics Division produces printed publications of statistics and statistical methods in the fields of international merchandise trade, national accounts, demography and population, social indicators, gender, industry, energy, environment, human settlements and disability. The Division also produces general statistical compendiums including the *Statistical Yearbook* and *World Statistics Pocketbook*. Many of the Division's databases in these fields are available as electronic publications in the form of CD-ROM, diskette and magnetic tape and on the Internet.

Government sources

European governments and the EU also produce large amounts of secondary data. Each European country has its own statistical office which produces lists of the publications available (and the costs involved). Examples of national statistical offices include the Centraal Bureau voor de Statistiek Nederlands (**www.cbs.nl**), Danmarks Statistik (**www.dst.dk**), the Federal Statistical Office of Germany (**www.statistik-bund.de**), the French Institut National de la Statistique et des Études Economiques (**www.insee.fr**), and the British Office for National Statistics (**www.statistics.gov.uk**). All of these offices have Internet links that allow you to quickly examine the array of publications that they produce. Their publications may be divided into census data and other publications.

Census data. Most European countries produce either catalogues or newsletters that describe the array of census publications available and the plans for any forthcoming census. In Britain, for example, *Census News* (**www.statistics.gov.uk/ census2001/cennews.asp**) is a newsletter that contains the latest information about the 2001 census and previous censuses and is available four to six times a year. Census Marketing in Britain can supply unpublished data from the 1961, 1971, 1981 and 1991 censuses in the form of Small Area Statistics (SAS). SAS are available for standard census areas within England and Wales, such as counties, local government districts, London boroughs, wards, civil parishes and enumeration districts. Maps can also be purchased to complement the data.

Census data can be kept in electronic formats, allowing it to be analysed and presented in a variety of formats at a detailed geographical level. Given the long periods between national censuses and the amount of change that can occur in these periods, other data sources are used to maintain an up-to-date picture of specific regions.

As well as general population censuses, national statistical offices produce an array of industrial censuses. These may include industrial production, housing, construction, agriculture, restaurants and hotels, and financial services, for example **www.statistics.gov.uk/statbase/mainmenu.asp**.

Other government publications. In addition to the census, national statistical offices collect and publish a great deal of statistical data. Examining the Department of Statistics and Research in Cyprus as an example (**www.pio.gov.cy**), major industrial categories such as agriculture, construction, retailing and tourism are classified, with a whole array of available statistics. More generally, demographic, health, household income and expenditure, and labour statistical reports are also available.

Examples of reports from the British Office for National Statistics include *Family Spending* and *Economic Trends* (**www.statistics.gov.uk/statbase/publication.asp**). *Family Spending* provides a snapshot of household spending in the UK, explaining in detail how consumers spend their money. Expenditure patterns are broken down by age, economic

status, income and geography. *Economic Trends* provides monthly macroeconomic statistics of economic trends which include key data such as national accounts, gross domestic product, disposable income, balance of payments, trade in goods, prices, labour market information, industrial output, consumer sales credit and interest rates, as well as extensive commentary to put all this information into context.

In the EC, statistics are collected and published by the Statistical Office of the European Community (SOEC) in a series called Eurostat (**www.europa.eu.int/com/euro stat**). Tables normally contain figures for individual member states of the EU plus totals for all countries. Eurostat divides its publications into themes, which are:

- *Theme 1* – General statistics
- *Theme 2* – Economy and finance
- *Theme 3* – Population and social conditions
- *Theme 4* – Energy and industry
- *Theme 5* – Agriculture, forestry and fisheries
- *Theme 6* – External trade
- *Theme 7* – Distributive trades, services and transport
- *Theme 8* – Environment
- *Theme 9* – Research and development.

It also produces general titles which include *Eurostat Yearbook* (annual), *Basic Statistics* (annual), *Europe in Figures* (annual), *Key Figures* (monthly) and *Eurostatistics* (monthly).

To examine any of the national statistics offices in Europe, visit **www.cso.ie/ links/eurolinks.html** and follow a link to the country of your choice. To examine other international statistics organisations visit **www.cso.ie/links/interlinks.html**.

Computerised databases

Most published information is also available in the form of computerised databases. Computerised databases contain information that has been made available in computer readable form for electronic distribution. From the 1980s through to today, the number of databases, as well as the vendors providing these services, has grown enormously. Computerised databases offer a number of advantages over printed data, including:[11]

1 The data are current and up-to-date, as publishers and data compilers are now using computers as the primary production technology.
2 The search process is more comprehensive, quicker and simpler. Online vendors provide ready access to hundreds of databases. Moreover, this information can be accessed instantaneously, and the search process is simplified as the vendors provide uniform search protocols and commands for accessing the database.
3 The cost of accessing these is relatively low, because of the accuracy of searching for the right data, and the speed of location and transfer of data.
4 It is convenient to access these data using a personal computer fitted with an appropriate communication device, such as a modem or a communication network.

While computerised database information can be helpful, it is vast and can be confusing. Thus a classification of computerised databases is helpful.

Classification of computerised databases

Computerised databases may be classified as online, Internet or offline as shown in Figure 4.3. Online databases consist of a central data bank that is accessed with a computer (or dumb terminal) via a telecommunications network. Internet databases can be accessed, searched and analysed on the Internet. It is also possible to download data from the Internet and store it in the computer or an auxiliary storage device.[12]
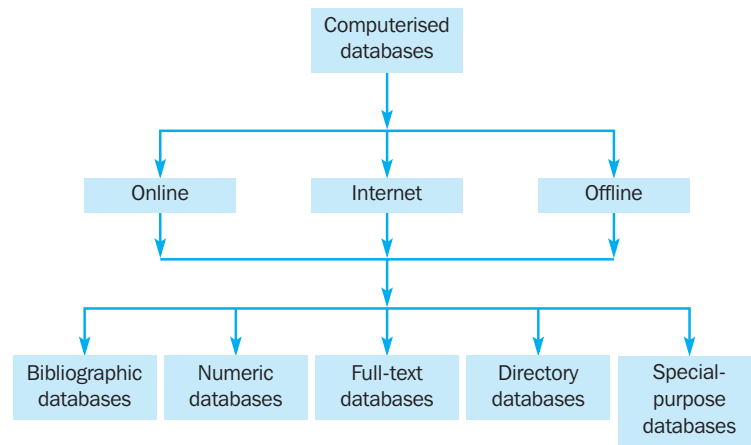
**Figure 4.3
A classification of
computerised
databases**

Online databases
Databases, stored in
computers, that require a
telecommunications network
to access.

Internet databases
Databases that can be
accessed, searched and
analysed on the Internet. It is
also possible to download
data from the Internet and
store it on the computer or an
auxiliary device.

Offline databases
Databases that are available
on diskette or CD-ROM.

Bibliographic databases
Databases composed of
citations to articles in journals,
magazines, newspapers,
marketing research studies,
technical reports, government
documents, and the like. They
often provide summaries or
abstracts of the material cited.

Numeric databases
Databases containing
numerical and statistical
information that may be
important sources of
secondary data.

Full-text databases
Databases that contain the
complete text of secondary
source documents comprising
the database.

Directory databases
Databases that provide
information on individuals,
organisations and services.

**Special-purpose
databases**
Databases that contain
information of a specific
nature, e.g. data on a
specialised industry.

Offline databases make the information available on diskettes and CD-ROM disks. Thus, offline databases can be accessed at the user's location without the use of an external telecommunications network.[13]

Online, **Internet** and **offline databases** may be further classified as bibliographic, numeric, full text, directory or special-purpose databases. **Bibliographic databases** are composed of citations to articles in journals, magazines, newspapers, marketing research studies, technical reports, government documents and the like.[14] They often provide summaries or abstracts of the material cited. The earlier example of *Market Research Abstracts* (**www.imriresearch.com/index.html**) is an example of a bibliographic database. Another example is the *Aslib Index to Theses* (**www.theses.com**); this bibliographic database lists theses at masters and doctoral level and research degrees, including abstracts from Britain and Ireland.

Numeric databases contain numerical and statistical information. For example, some numeric databases provide time series data about the economy and specific industries. The earlier examples of census-based numeric databases using data over a series of censuses provide an example of a numeric database.

Full-text databases contain the complete text of the sources of the database. Examples include *World Advertising Research Center WARC* (**www.warc.com**), *Searchbank: European ASAP* (**www.infotrac.london.galegroup.com**) and *FT Discovery* (**www.lexis-nexis.com**). *WARC* is a supplier of intelligence to the global marketing, advertising, media and research communities. *Searchbank* has over 100 full-text journals on subjects including business, economics, current affairs and new technologies. It includes a spectrum of journals from professional trade publications through to refereed academic journals. *FT Discovery* presents the *Global News* link that gives full access to the Financial Times, World Reporter, Europe Intelligence Wire and Asia Intelligence Wire. It is possible to search on specific countries, sectors and publications.

Directory databases provide information on individuals, organisations and services. *European Interactive Directories* (**www.euroyellowpages.com**) is an example of a directory that has channels based upon EU community activities, country channels and thematic channels such as suppliers, wholesalers and shopping centres. Another example worth examining is the *ESOMAR* directory (**www.esomar.nl**) which provides details of member organisations throughout the world as well as many other publications of value to marketing researchers based in Europe.

Finally, there are **special-purpose databases**. For example, the Non-Governmental Organisation NGO directory (**www.rec.org/REC/Databases/NGODirectory/ NGOfind.html**) helps to track down information about environmental organisations working in central and eastern Europe. It has contact information for over 2,700 organisations from over 15 central and eastern European countries.

In addition, virtually all libraries of major universities maintain special-purpose databases of research activities that reflect the distinct specialisms of that university. Beyond the internally generated, special-purpose databases, university libraries and reference libraries maintain computerised databases with instructions relating to what may be accessed and how it may be accessed. Another library source worth examining for computerised sources is the European Commission's 'Libraries' site (**www.europa.eu.int**). The site, which is multilingual, is distributed by the EUROPA server. EUROPA is the portal site of the European Union. It provides up-to-date coverage of European affairs and essential information on European integration. Users can access Websites of each of the EU institutions.

Syndicated sources of secondary data

Syndicated sources (services)
Information services offered by marketing research organisations that provide information from a common database to different firms that subscribe to their services.

In addition to published data or data available in the form of computerised databases, syndicated sources constitute the other major source of external secondary data. **Syndicated sources**, also referred to as **syndicated services**, are companies that collect and sell common pools of data designed to serve information needs shared by a number of clients. These data are not collected with a focus on a specific marketing problem, but the data and reports supplied to client companies can be personalised to fit specific needs. For example, reports could be organised based on the clients' sales territories or product lines. Using syndicated services is frequently less expensive than commissioning tailored primary data collection. Figure 4.4 presents a classification of syndicated sources. Syndicated sources can be classified based on the unit of measure-
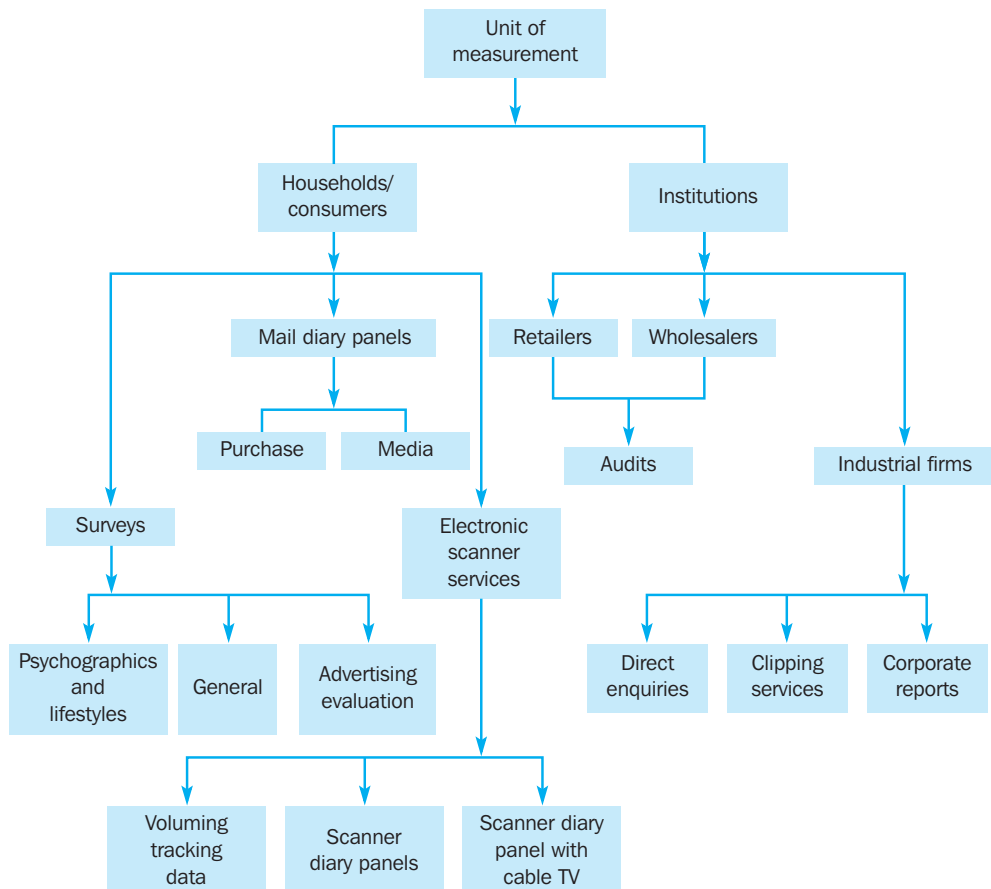
Figure 4.4
A classification of syndicated services

ment (households and consumers or institutions). Household and consumer data may be obtained from surveys, diary panels or electronic scanner services. Information obtained through surveys consists of values and lifestyles, advertising evaluation, or general information related to preferences, purchase, consumption and other aspects of behaviour. Diary panels emphasise information on purchases or media consumption. Electronic scanner services might provide scanner data only, scanner data linked to diary panels, or scanner data linked to diary panels and (cable) TV. When institutions are the unit of measurement, the data may be obtained from retailers, wholesalers or industrial firms. An overview of the various syndicated sources is given in Table 4.3. Each of these sources will be discussed.

Table 4.3 Overview of syndicated services

Type	Characteristics	Advantages	Disadvantages	Uses
Surveys	Surveys conducted at regular intervals	Most flexible way of obtaining data; information on underlying motives	Interviewer errors; respondent errors	Market segmentation, advertising theme selection and advertising effectiveness
Mail diary panels	Households provide specific information regularly over an extended period of time; respondents asked to record specific behaviour as it occurs	Recorded purchase behaviour can be linked to the demographic/ psychographic characteristics	Lack of representativeness; response bias; maturation	Forecasting sales, market share and trends; establishing consumer profiles, brand loyalty and switching; evaluating test markets, advertising and distribution
Diary media panels	Electronic devices automatically recording behaviour, supplemented by a diary	Same as mail diary panel	Same as mail diary panel	Establishing advertising rates; selecting media programme or air time; establishing viewer profiles
Scanner volume tracking data	Household purchases are recorded through electronic scanners in supermarkets	Data reflect actual purchases; timely data, less expensive	Data may not be representative; errors in recording purchases; difficult to link purchases to elements of marketing mix other than price	Price tracking, modelling effectiveness of in-store modelling
Scanner diary panels with cable TV	Scanner panels of households that subscribe to cable TV	Data reflect actual purchases; sample control; ability to link panel data to household characteristics	Data may not be representative; quality of data limited	Promotional mix analyses, copy testing, new product testing, positioning
Audit services	Verification of product movement by examining physical records or performing inventory analysis	Relatively precise information at retail and wholesale levels	Coverage may be incomplete; matching of data on competitive activity may be difficult	Measurement of consumer sales and market share; competitive activity; analysing distribution patterns; tracking of new products
Industrial firms	Data banks on industrial establishments created through direct enquiries of companies, clipping services and corporate reports	Important source of information in industrial firms; particularly useful in initial phases of the projects	Data are lacking in terms of content, quantity and quality	Determining market potential by geographic area, defining sales territories, allocating advertising budget

Syndicated data from households

Surveys
Interviews with a large number of people using a questionnaire.

Omnibus survey
A distinctive form of survey that serves the needs of a syndicate group. The omnibus survey targets particular types of respondents such as those in specific locations, e.g. Luxembourg residents, or customers of particular types of product, e.g. business air travellers. With that target group of respondents, a core set of questions can be asked with other questions added as syndicate members wish.

Psychographics
Quantified profiles of individuals based upon lifestyle characteristics.

Lifestyles
Distinctive patterns of living described by the activities people engage in, the interests they hold of themselves and the world around them.

Surveys

Various syndicated services regularly conduct **surveys** and **omnibus surveys**. In general, these surveys involve interviews with a large number of respondents using a pre-designed questionnaire. The distinction of the omnibus survey (example at **www.mori.com/charity/omnibus.shtml**) is that it targets particular types of respondents such as those in certain geographic locations, e.g. Luxembourg residents, or consumers of particular types of products, e.g. business air travellers. With that target group of respondents, a core set of questions can be asked with other questions added as syndicate members wish. Other syndicate members can 'jump on the omnibus' and buy the answers to all the questionnaire responses or to specific questions of their choice. Surveys and omnibus surveys may be broadly classified based on their content as psychographics and lifestyles, advertising evaluation, or general surveys.

Psychographics and lifestyles. **Psychographics** refer to the psychological profiles of individuals and to psychologically based measures of lifestyle. **Lifestyles** refer to the distinctive modes of living of a society or some of its segments. Together, these measures are generally referred to as activities, interests and opinions.

Advertising evaluation. The purpose of advertising evaluation surveys is to assess the effectiveness of advertising using print and broadcast media. Television commercials are evaluated using either the recruited audience method or the in-home viewing method. In the former method, respondents are recruited and brought to a central viewing facility, such as a theatre or mobile viewing laboratory. The respondents view the commercials and provide data regarding knowledge, attitudes and preferences related to the product being advertised and the commercial itself. See example at **www.bmrb.co.uk/researchingadvertising/addevelopment.htm**.

In the in-home viewing method, consumers evaluate commercials at home in their normal viewing environment. New commercials can be pre-tested at the network level or in local markets. A survey of viewers is then conducted to assess the effectiveness of the commercials.

General surveys. Surveys are also conducted for a variety of other purposes, including examination of purchase and consumption behaviour. Because a variety of data can be obtained, survey data have numerous uses. They can be used for market segmentation, as with psychographic and lifestyle data, and for establishing consumer profiles. Surveys are also useful for determining product image, measurement and positioning, and conducting price perception analysis. Other notable uses include advertising theme selection and evaluation of advertising effectiveness.

Mail diary panels

Often, survey data can be complemented with data obtained from diary panels (see example at **www.tnsofres.com/about/americas/argentina/argentina.cfm**). Panels were discussed in Chapter 3 in the context of longitudinal research designs. Diary panels are samples of respondents who provide specified information at regular intervals over an extended period of time. These respondents may be organisations, households or individuals, although household diary panels are most common. The distinguishing feature of diary panels is that the respondents record specific behaviours as they occur, in a diary. Typically, the diary is returned to the research organisation every one to four weeks. Panel members are compensated for their

participation with gifts, coupons, information or cash. Based on the content of information recorded, diary panels can be classified as diary purchase panels or diary media panels.

In diary media panels, electronic devices automatically record viewing behaviour, thus supplementing a diary. **Diary media panels** yield information helpful for establishing advertising rates by radio and TV networks, selecting appropriate programming, and profiling viewer or listener subgroups. Advertisers, media planners and buyers find panel information particularly useful. **Diary purchase panels** provide information useful for forecasting sales, estimating market shares, assessing brand loyalty and brand-switching behaviour, establishing profiles of specific user groups, measuring promotional effectiveness, and conducting controlled store tests.

Compared with sample surveys (see Chapter 10), diary panels offer certain distinct advantages.[15] Panels can provide longitudinal data (data can be obtained from the same respondents repeatedly). People who are willing to serve on panels may provide more and higher-quality data than sample respondents. In diary purchase panels, information is recorded at the time of purchase, eliminating recall errors.[16] Information recorded by electronic devices is accurate because it eliminates human errors.

The disadvantages of diary panels include lack of representativeness, maturation and response biases. They may under-represent certain groups such as minorities and those with low education levels. This problem is further compounded by refusal to respond and attrition of panel members. Over time, maturation sets in, and the panel members must be replaced. Response biases may occur, since simply being on the panel may alter behaviour. Because purchase or media data are entered by hand, recording errors are also possible (see Chapter 3).

Electronic scanner services

The following example illustrates the nature and scope of electronic scanner services as undertaken by A.C. Nielsen, who conduct consumer panel services in 18 countries around the world.

> **Diary media panels**
> A data gathering technique composed of samples of respondents whose television viewing behaviour is automatically recorded by electronic devices, supplementing the purchase information recorded in a diary.

> **Diary purchase panels**
> A data gathering technique in which respondents record their purchases in a diary.

example

A.C. Nielsen – the business economist at work[17]

Most of A.C. Nielsen's revenue comes from selling information on *fast-moving consumer goods* (FMCG). This information is compiled from either scanner data obtained from thousands of supermarkets, pharmacists and department stores, or from a 40,000-household panel who electronically record every aspect of every goods purchase they made using that item's bar code.

They go beyond their standard forms of analysing and presenting existing data to more creative interpretations that attempt to tell the marketer something he or she does not know. Often, creativity is valued more than sophisticated econometric techniques, as illustrated in the following two examples.

Inventory and sales data are available on food store retail sales, so an inventory–sales ratio can be calculated. This ratio had been increasing for several years up to 1992 when it began to shrink. Yet it is generally accepted that the size of the average new store is increasing, implying the need for additional inventories relative to sales. Coincidentally, more sophisticated inventory management techniques have become especially prevalent in food stores in recent years. Apparently, the ratio is being driven more by improved inventory management than by the opening of new, larger stores. A discrete cause and effect cannot be proven, but a linkage between the ratio and better inventory control mechanisms is highly probable.

Consumer spending patterns in Asia are very difficult to analyse, given the paucity of data available and the lack of data comparability across countries or spending components. By calculating consumer spending on food as a share of total consumer spending, and then ordering the results by per capita GDP, a relationship becomes obvious. Poorer countries

spend proportionately more of their resources on food, around 50%. For middle-income countries, this ratio slides from around 40% to 20%, and then holds steady near 20% for the developed economies. The implication is that, as a country obtains a middle-income status, because of this declining ratio, food sales will not grow as fast as other categories of consumer spending. ■

Although information provided by surveys and diary panels is useful, electronic scanner services are becoming increasingly popular. The role of scanned data as a foundation to developing sophisticated consumer databases is developed in Chapter 5. In this chapter we examine scanned data as a distinct source of syndicated data. **Scanner data** reflect some of the latest technological developments in the marketing research industry. They are collected by passing merchandise over a laser scanner that optically reads the bar-coded description (Universal Product Code, or UPC) printed on the merchandise. This code is then linked to the current price held in the computer memory and used to prepare a sales slip. Information printed on the sales slip includes descriptions as well as prices of all items purchased. Checkout scanners, now used in many retail stores, are revolutionising packaged goods marketing research.

Three types of scanner data are available: **volume tracking data**, **scanner diary panels**, and **scanner diary panels with cable TV**. Volume tracking data provide information on purchases by brand, size, price and flavour or formulation, based on sales data collected from the checkout scanner tapes. This information is collected nationally from a sample of supermarkets with electronic scanners. In scanner diary panels, each household member is given an ID card that looks like a credit card. Panel members present the ID card at the checkout counter each time they shop. The checker keys in the ID numbers and each item of that customer's order. The information is stored by day of week and time of day.[18]

An even more advanced use of scanning, scanner diary panels with cable TV, combines diary panels with new technologies growing out of the cable TV industry. Households on these panels subscribe to one of the cable or TV systems in their market. By means of a cable TV 'split', the researcher targets different commercials into the homes of the panel members. For example, half the households may see test commercial A during the 6:00 pm newscast while the other half see test commercial B. These panels allow researchers to conduct fairly controlled experiments in a relatively natural environment.[19]

Uses of scanner data. Scanner data are useful for a variety of purposes.[20] National volume tracking data can be used for tracking sales, prices and distribution, for modelling, and for analysing early warning signals. Scanner diary panels with cable TV can be used for testing new products, repositioning products, analysing promotional mix, and making advertising decisions, including budget, copy and media, and pricing. These panels provide marketing researchers with a unique controlled environment for the manipulation of marketing variables.

Advantages and disadvantages of scanner data. Scanner data have an obvious advantage over surveys and diary panels; they reflect purchasing behaviour that is not subject to interviewing, recording, memory or expert biases. The record of purchases obtained by scanners is complete and unbiased by price sensitivity, because the panellist is not required to be particularly conscious of price levels and changes. Another advantage is that in-store variables such as pricing, promotions and displays are part of the dataset. The data are also likely to be current and can be obtained quickly. Finally, scanner panels with cable TV provide a highly controlled testing environment.

Scanner data
Data obtained by passing merchandise over a laser scanner that reads the UPC code from the packages.

Volume tracking data
Scanner data that provide information on purchases by brand, size, price, and flavour or formulation.

Scanner diary panels
Scanner data where panel members are identified by an ID card, allowing information about each panel member's purchases to be stored with respect to the individual shopper.

Scanner diary panels with cable TV
The combination of a scanner diary panel with manipulations of the advertising that is being broadcast by cable television companies.

A major weakness of scanner data is lack of representativeness. National volume tracking data may not be generalisable to the total population, because only large supermarkets have scanners. In addition, certain types of outlets such as food warehouses, pharmacists and mass merchandisers are excluded. Likewise, scanners have limited geographical dispersion and coverage.

The quality of scanner data may be limited by several factors. Not all products may be scanned. For example, to avoid lifting a heavy item, a sales assistant may use the register to ring it up. If an item does not scan on the first try, the assistant may key in the price and ignore the bar code. Sometimes a consumer purchases many flavours of the same item, but the assistant scans only one package and then rings in the number of purchases. Thus, the transaction is inaccurately recorded. With respect to scanner panels, the available technology permits the monitoring of only one TV set per household. Hence, there is a built-in bias if the household has more than one TV set. In addition, the system provides information on TV sets in use rather than actual viewing behaviour. Although scanner data provide behavioural and sales information, they do not provide information on underlying attitudes and preferences and the reasons for specific choices.

Syndicated data from institutions

Retailer and wholesaler audits

Audit
A data collection process derived from physical records or performing inventory analysis. Data are collected personally by the researcher, or by representatives of the researcher, and are based on counts usually of physical objects rather than people.

As Figure 4.4 shows, syndicated data are available from retailers and wholesalers as well as industrial firms (see the example of the work of German marketing research company GfK at **www.gfkms.com**). The most popular means of obtaining data from retailers and wholesalers is an audit. An **audit** is a formal examination and verification of product movement carried out by examining physical records or analysing inventory. Retailers and wholesalers who participate in the audit receive basic reports and cash payments from the audit service. Audit data focus on the products or services sold through the outlets or the characteristics of the outlets themselves.

e x a m p l e

Retail auditing for retailing information[21]

Introduced in the USA in 1933, A.C. Nielsen pioneered food and drug indices to measure and understand the performance and dynamics of product sales. The Retail Measurement Service of A.C. Nielsen uses store audit data on product movement, market share, distribution, price and other market-sensitive information in over 80 countries across six continents. Using in-store scanning of product codes and store visits by auditors, sample and census information is gathered across the food, household, health and beauty, durables, confectionery and beverage industries. ■

Retail audit data can be useful to consumer product firms. For example, if Colgate-Palmolive is contemplating the introduction of a new toothpaste brand, a retail audit can help determine the size of the total market and distribution of sales by type of outlet and by different regions.

Wholesale audit services, the counterpart of retail audits, monitor warehouse withdrawals. Participating operators, who include supermarket chains, wholesalers and frozen-food warehouses, typically account for over 80% of the volume in the area.

The uses of retail and wholesale audit data include:

1 determining the size of the total market and the distribution of sales by type of outlet, region, or city;
2 assessing brand shares and competitive activity;

Electronic scanning can capture the products used by a consumer far more accurately than manual methods.

3 identifying shelf space allocation and inventory problems;
4 analysing distribution problems;
5 developing sales potentials and forecasts;
6 developing and monitoring promotional allocations based on sales volume.

Audits provide relatively accurate information on the movement of many different products at the wholesale and retail levels. Furthermore, this information can be broken down by a number of important variables, such as brand, type of outlet and size of market. Audits have limited coverage, however; not all markets or operators are included. In addition, audit information may not be timely or current, particularly compared with scanner data. Typically, there is a two-month gap between the completion of the audit cycle and the publication of reports. Another disadvantage is that, unlike scanner data, audit data cannot be linked to consumer characteristics. In fact, there may even be a problem in relating audit data to advertising expenditures and other marketing efforts. Some of these limitations are overcome in computerised audit panels.

Industrial firms

These provide syndicated data about industrial firms, businesses and other institutions. Syndicated data are collected by making direct enquiries to organisations, from clipping services and through the analysis of corporate reports (see example at **www.journalismnet.com/choose/clippings.htm**). The range and sources of syndicated data available to industrial goods firms are more limited than those available to consumer goods firms.

Industrial firm information is useful for sales management decisions, including identifying prospects, defining territories, setting quotas, and measuring market potential by geographic areas. It can also aid in advertising decisions such as targeting prospects, allocating advertising budgets, selecting media, and measuring advertising effectiveness. This kind of information is useful for segmenting the market and for designing custom products and services for important segments.

International marketing research

A wide variety of secondary data are available for international marketing research. As with domestic research, the problem is not lack of data but the plethora of information available. Evaluation of secondary data is even more critical for international than for domestic projects. Different sources report different values for a given statistic, such as the gross domestic product (GDP), because of differences in the way the unit is defined. Measurement units may not be equivalent across countries. In France, for example, many workers are paid a thirteenth monthly salary each year as an automatic bonus, resulting in a measurement construct that is different from those in other countries.[22] The accuracy of secondary data may also vary from country to country. Data from highly industrialised countries in Europe are likely to be more accurate than those from developing nations. Business and income statistics are affected by the taxation structure and the extent of tax evasion. Population censuses may vary in frequency and year in which the data are collected. In Britain, for example, the census is conducted every 10 years, whereas in the People's Republic of China there was a 29-year gap between the censuses of 1953 and 1982. This situation, however, is changing quickly. Several syndicated firms are developing huge sources of international secondary data, as illustrated in the following example.

example

Los Medios y Mercados de Latinoamerica (LMML)[23]

Started in 1994 by Audits & Surveys worldwide, *Los Medios y Mercados de Latinoamerica* (The Markets and Media of Latin America) is the largest multinational survey of media and consumer habits that is conducted in Latin America to provide managers with important information for their marketing strategies. The study, which is repeated every year, aims at tracking the development of media and consumer habits in Latin America.

A recent multinational survey was conducted in 18 Latin American countries, including Argentina, Brazil, Colombia, Mexico and Venezuela. It sampled 6,634 respondents between the ages of 12 and 64. Its probability sample, representing urban as well as rural Latin America, can be projected to 280 million people or 79 million households.

The research design for this survey involved two steps. First, the personal interview technique was used to measure the variety of media, including newspapers, multinational and local magazines, television and radio. Then a 25-page self-administered booklet was passed to the respondent to measure their product consumption and usage in over 100 categories and 800 brands. Demographic data gathered about the respondents included country/region, age, sex, employment status, occupation, education, household size, annual household income, car ownership, household goods owned and services taken.

Companies can easily use these data, as the survey results are provided in a set of 14 printed volumes, and also in computer database formats including an SPSS format. ■

Ethics in marketing research

Possible ethical dilemmas exist when using internal or external secondary data. Some ethical issues that are pertinent include:

- The unnecessary collection of primary data when the problem can be addressed based on secondary data alone
- The use of only secondary data when primary data are needed
- The use of secondary data that are not applicable
- The use of secondary data that have been gathered through morally questionable means
- Compromising the anonymity of customer details held on databases.

As was discussed in Chapter 2, the unnecessary collection of expensive primary data when the research problem can be addressed using only secondary data is unethical. In this case, the researcher is using a more expensive method that is less appropriate. Similarly, the exclusive reliance on secondary data when the research problem requires primary data collection could raise ethical concerns. This is particularly true if the researcher is charging a fixed fee for the project and the research design was not specified in advance. Here again, the researcher's profit goes up, but at the expense of the client.

The researcher is ethically obliged to ensure the relevance and usefulness of secondary data to the problem at hand. The secondary data should be evaluated by the criteria discussed earlier in this chapter. Only data judged to be appropriate should be used.

Internet and computer applications

The World Wide Web as an online source of secondary data

The World Wide Web is a vital source of secondary data and intelligence for the marketing researcher. The speed of the Internet can aid rapid problem diagnosis and data collection at various stages of the research process. Given the global nature of the technology, the Internet is a vital tool for the international marketing researcher. Searching the Web is facilitated by using generalist search engines such as Google, Yahoo! or AltaVista, which require a few keywords to get hundreds of sites related to one subject. One can go directly to the Websites of traditional suppliers of secondary data from government or business sources. Many of those sites also have inside search engines that sort data from the supplier's internal database. Information on the Web is of great value as generally it is current, though care must be taken to note when Web pages have been updated. It should be noted that not all secondary data on the Web is free. The Web may reveal the existence of data on a particular subject or industry, but remember the costs involved in conducting quality research. Hence the Web may be used to give an awareness and a 'taste' of secondary data but it does not necessarily mean 'free' data. See the Companion Website for Professional Perspective 9 by Trevor Fenwick. Trevor's article 'How online retrieval is devaluing research' is a reminder of the quality issues of secondary data gathered from the Internet.

➤➤➤
See Professional Perspective 9.

Internal secondary data

Large organisations have intranets, which greatly facilitate the search for access to secondary data. The Coca-Cola Company, for example, has developed powerful intranet applications that enable Coca-Cola managers worldwide to search for past and present research studies and a wide variety of marketing-related information on the basis of keywords. Once located, the information can be accessed online. Even sensitive and restricted information can be accessed by obtaining permission electronically. Visit www.intranets.com for a fuller description of intranet technology, evaluations of software and an interactive tour.

External secondary data

Information can be obtained by visiting various business-related sites that provide sales leads and mailing lists, business profiles and credit ratings. Various newspapers, magazines and journals can be accessed on the Web with excellent indexing facilities to locate particular subjects, companies and individuals. Government data for the European Community and for individual countries through to regional and city councils can be accessed via the Web, though the quality and quantity of data available through government sources can vary enormously.

Syndicated sources of information

For syndicated sources of information one can visit the home pages of the various marketing research companies and providers of syndicated information. The A.C. Nielsen home page at www.acnielsen.com is a good example. This site provides links to various manufacturers and to various countries such as Britain, Canada and Spain. Other good sources of syndicated data include Mintel, which can be reached at www.mintel.co.uk, Taylor Nelson Sofres at www.tnsofres.com/research/prodfinder.cfm (it is worth working through the great array of marketing research products they offer), the German marketing research company GfK (English version) at www.gfk.de/english/presse/broschueren/produkte/BehaviorScan_e.pdf, and a French version of BehaviorScan at www.marketingscan.fr/produits/ behaviorscan/produits_behavior.html.

International secondary data

The Internet has emerged as the most extensive source of secondary information. The utility of the Internet for the marketing researcher is further enhanced due to the easy accessibility and retrieval of information and the ability to cross-validate information from a variety of sources. Most of the Internet links highlighted in this chapter allow for secondary data to be gathered from countries all over the world. Whilst necessarily we have a focus on European matters, it is well worth visiting www.quirks.com. The Quirks.com Website is a most thorough source for information on marketing research, including case studies of successful research projects and a comprehensive list of directories.

As a final illustration of the power of the Internet for secondary data collection and analyses, go to the Dun and Bradstreet home page at www.dbn.com, and directly access the services and syndicated information for the country of your choice.

Summary

In contrast to primary data, which originate with the researcher for the specific purpose of the problem at hand, secondary data and intelligence are data originally collected for other purposes. Secondary data can be obtained quickly and are relatively inexpensive. They have limitations, and should be carefully evaluated to determine their appropriateness for the problem at hand. The evaluation criteria consist of specifications, error, currency, objectivity, nature and dependability.

A wealth of information exists in the organisation for which the research is being conducted. This information constitutes internal secondary data. External data are generated by sources outside the organisation. These data exist in the form of published (printed) material, online and offline databases, or information made available by syndicated services. Published external sources may be broadly classified as general business data or government data. General business sources comprise guides, directories, indexes and statistical data. Government sources may be broadly categorised as census data and other data. Computerised databases may be online or offline. Both online and offline databases may be further classified as bibliographic, numeric, full-text, directory or specialised databases.

Syndicated sources are companies that collect and sell common pools of data designed to serve a number of clients. Syndicated sources can be classified based on the unit of measurement (households and consumers or institutions). Household and consumer data may be obtained via surveys, diary purchase or media panels, or electronic scanner services. When institutions are the unit of measurement, the data may be obtained from retailers, wholesalers or industrial units. It is desirable to combine information obtained from different secondary sources.

Several specialised sources of secondary data are useful for conducting international marketing research. The evaluation of secondary data becomes even more critical, however, because the usefulness and accuracy of these data can vary widely. Ethical dilemmas that can arise include the unnecessary collection of primary data, the use of only secondary data when primary data are needed, the use of secondary data that are not applicable, and the use of secondary data that have been gathered through morally questionable means.

Questions

1 What are the differences between primary data, secondary data and marketing intelligence?

2 What are the relative advantages and disadvantages of secondary data?

3 At what stages of the marketing research process can secondary data be used?

4 Why is it important to locate and analyse secondary data before progressing to primary data?

5 How may secondary data be used to validate qualitative research findings?

6 What is the difference between internal and external secondary data?

7 How can intranet technology help in the location and dissemination of secondary data?

8 By what criteria may secondary data be evaluated?

9 What criteria would you look for when examining the design and specifications of secondary data? Why is it important to examine these criteria?

10 To what extent should you use a secondary data source if you cannot see any explicit objectives attached to that research?

11 If you had two sources of secondary data for a project, the first being dependable but out of date, the second not dependable but up to date, which would you prefer?

12 Describe, with examples, the main types of government and business secondary data sources.

13 List and describe the main types of syndicated sources of secondary data.

14 Explain what a diary panel is. What are the advantages and disadvantages of traditional and scanner diary panels?

15 What is an audit? Describe the uses, advantages and disadvantages of audits.

Notes

1 Keating, P., 'The best airlines to fly today', *Money* (November 1997), 118–28.

2 Aguilar, F.J., *Scanning the Business Environment* (London: Macmillan, 1967).

3 Adapted from Brownlie, D., 'Environmental scanning', in Baker, M.J., *The Marketing Book*, 3rd edn, (Oxford: Butterworth-Heinemann), 158.

4 Curtis, J., 'Behind enemy lines', *Marketing* (24 May 2001), 28–9.

5 For a recent application of secondary data, see Czinkota, M.R. and Ronkainen, I.A., 'Market research for your export operations: Part I – Using secondary sources of research', *International Trade Forum* 3 (1994), 22–33; Jain, D., Mahajan, V. and Muller, E., 'Innovation diffusion in the presence of supply restrictions', *Marketing Science* (Winter 1991) 83–90.

6 Jacob, H., *Using Published Data: Errors and Remedies* (Beverly Hills, CA: Sage Publications, 1984).

7 Stewart, D.W., *Secondary Research: Information Sources and Methods* (Beverly Hills, CA: Sage Publications, 1984), 23–33.

8 Fry, A., 'Number crunch', *Marketing* (4 December 1997), 29.

9 Fries, J.R., 'Library support for industrial marketing research', *Industrial Marketing Management* 11 (February 1982), 47–51.

10 Miller, R., 'Keeping up to data', *Marketing* (30 January 1997), 35.

11 Post, C., 'Marketing data marts help companies stay ahead of the curve and in front of the competition', *Direct Marketing* 59 (April 1997), 37–40.

12 Notess, G.R., 'Searching the hidden Internet', *Database* 20 (June/July 1997), 37–40.

13 Quint, B., 'Assume the position, take the consequences', *Information Today* 13 (June 1996), 11–13.

14 Notess, G.R., 'The Internet as an on-line service: bibliographic databases on the net', *Database* 19 (August/September 1996), 92–5.

15 Parfitt, J.H. and Collins, B.J.K., 'Use of consumer panels for brand share predictions', *Journal of the Market Research Society* 38(4) (October 1996), 341–67; Ramaswamy, V. and Desarbo, W.S., 'SCULPTURE: a new methodology for deriving and analyzing hierarchical product-market structures from panel

data', *Journal of Marketing Research* 27 (November 1990), 418–27.

16 Sudman, S., 'On the accuracy of recording of consumer panels I', *Journal of Marketing Research* (August 1964), 69–83; Sudman, S. and Bradburn, M., 'Response effects in surveys', *Journal of Marketing Research* (May 1964), 14–20; Sudman, S. and Ferber, R., *Consumer Panels* (Chicago, IL: American Marketing Association, 1978); Sudman, S., *On the Accuracy of Recording of Consumer Panels II*, Learning Manual (New York: Neal-Schumen Publishers, 1981).

17 Handler, D., 'The business economist at work: linking economics to market research: A.C. Nielsen', *Business Economics* 31 (October 1996), 51.

18 Andrew, R.L. and Srinivasan, T.C., 'Studying consideration effects in empirical choice models using scanner panel data', *Journal of Marketing Research* 32 (February 1995), 30-41; and Bucklin, R.E., Gupta, S. and Han, S., 'A brand's eye view of response segmentation in consumer brand choice behaviour', *Journal of Marketing Research* 32 (February 1995), 66–74.

19 It is possible to combine store-level scanner data with scanner panel data to do an integrated analysis. See, for example, Russell, G.J. and Kamakura, W.A., 'Understanding brand competition using micro and macro scanner data', *Journal of Marketing Research* 31 (May 1994), 289–303.

20 Examples of scanner data applications include Fader, P.S. and McAlister, L., 'An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data', *Journal of Marketing Research* 27 (August 1990), 322–32; Waarts, E., Carree, M. and Wierenga, B., 'Full-information maximum likelihood estimation of brand positioning maps using supermarket scanning data', *Journal of Marketing Research* 28 (November 1991), 483–90.

21 Based on information obtained from the suppliers.

22 Douglas, S.P. and Craig, C.S., *International Marketing Research* (Upper Saddle River, NJ: Prentice Hall, 1983).

23 Rydholm, J., 'A united effort', *Quirk's Marketing Research Review* (October 1996).

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Internal secondary data and the use of databases

Objectives

After reading this chapter, you should be able to:

1 describe the nature and purpose of internal sources of secondary data;

2 describe how different technological developments have increased the array of internally generated secondary data;

3 understand how databases are developing into powerful means to understand consumer behaviour through 'electronic observation';

4 understand how databases support traditional forms of marketing research to build up behavioural and attitudinal 'pictures' of target markets;

5 understand how geodemographic information systems can help in integrating data sources and in the graphical display of findings in a non-statistical manner;

6 describe how the link-up of different databases and survey data can be developed through the use of datawarehouses and be analysed through data mining techniques;

7 understand international data capture issues;

8 understand the ethical problems of having individual consumer data held on databases.

Different departments have dealings with customers and typically hold data on them. In order to develop an understanding of these customers, it is essential that these data are integrated.

Overview

Marketing research as a function does not support marketing decision-making in isolation. As discussed in Chapter 1, it may be seen as part of a broader marketing information system that supports strategic decision-making. Many information technology advances have been made in recent years that have fundamentally changed the way that marketing decisions are supported. For example, significant developments in database technology have meant that scanning systems in retail stores, loyalty card data, store panel data and survey data can be fused together to present very clear and up-to-date 'pictures' of consumers. As well as giving direct support to the marketer, these systems give more focus to marketing research activity and direct support to many stages of research.

This chapter describes how internal secondary data and databases have developed to make major impacts upon how decision-makers are supported. The data collected and analysed through database marketing can be seen as secondary data sources. As with all good secondary data sources, they have a major impact upon the conduct and direction of primary data collection, analyses and interpretation. There are also many ethical issues related to utilising internally generated customer data and the use of databases.

We introduce our discussion with an example of the use of databases. Databases generated within companies or bought in from specialist sources are primarily viewed as a tool to generate direct sales and target promotion activities. However, internally generated customer databases are a secondary data source of value to marketing researchers. Technological developments in the collection, analysis and presentation of data present great opportunities to researchers. There are also potential conflicts with the philosophy of anonymity in marketing research. This example illustrates how a customer database can reveal behavioural characteristics of customers, valuable information to decision-makers in its own right, but for the researcher an excellent foundation to a research design in deciding who to research and what issues to focus upon.

example

Making data go further

In the Netherlands, the use of Geodemographic Information Systems (GIS) is very sophisticated due to the range of data available from official sources such as utilities and the post office. Center Parcs, the chain of holiday villages owned by the brewery Scottish and Newcastle, used a GIS to target potential visitors to its new Dutch centre.

Nicole Kessels, Center Parcs' direct marketing manager (Netherlands) says: 'a segmentation system linked to our own database gives us a valuable insight into who our customers are, how they behave and where we can find potential new customers.' ■

Internal secondary data

Chapter 4 described the nature, purpose and value of secondary data to marketing researchers. A vital source of secondary data comes from within organisations that commission marketing research, namely internal secondary data. These data are generally seen as being 'operational data', i.e. data that represent the daily activities and transactions of a business. Daily transactions may be held in different departments such as sales, accounts or human resources and stored in different manners. The use of operational data has presented opportunities to researchers for as long as businesses have been recording their daily transactions. Even in the days of transactions being recorded manually, it has been a task for marketing researchers to track down different sources of data and analyse them. Locating and analysing internal sources of

secondary data should be the starting point in any marketing research project. The main reasons are that, as these data have already been collected, there are no additional data collection costs, there should be no access problems (though individual managers may make access difficult for personal or political reasons) and the quality of the data should be easier to establish (in comparison with externally generated data).

Most organisations have a wealth of in-house information even if they are not marketing or customer focused, so some data may be readily available. For example, imagine a timber merchant that sells wood to builders and cabinetmakers. It creates invoices for all its sales. Its accounts department handles this process and maintains the data that it generates. Yet, there exists much consumer behaviour data in these invoices. They could be analysed by:

- What products customers buy
- Which customers buy the most products
- Which customers repeat purchases
- Which customers appear only when there are special offers
- Where these customers are located
- How these customers pay – by cash or credit
- Which customers are the most profitable
- Seasonal patterns of purchasing behaviour by product types and customer types.

There may also be data that relate to promotions activities such as spending on advertising, trade fairs, sponsorship deals or personal selling. The researcher could look for details of spending in these areas and seek correlations with any of the analyses of customer behaviour. The task facing the marketing researcher is to search for such data, conduct analyses and present these to decision-makers to interpret. Such a process may focus the thoughts of decision-makers, by realising the potential that lies in these data. With this focus, other types of **operational data** and their value may be realised, managers in other parts of the organisation may release data that they guarded, and connections to sources of intelligence may be generated. Here lies the basis of generating clearly focused primary data collection, and effective marketing research.

Operational data
Data generated about an organisation's customers, through day-to-day transactions.

More marketing decision-makers have realised the benefits of analysing customer data. This realisation and technological developments in collecting, analysing and presenting customer data have given birth to a concept known as Customer Relationship Management (CRM). The main challenge of implementing CRM is to integrate customer data from the post, telephone, personal visits and the Internet into a central data pool so that it holds all transactions and contacts with each customer allied with the ability to update the data constantly and to access it immediately whenever necessary.[1] An illustration of the use and effect of CRM is presented in the following example.

example

Getting to know clients lifts profits [2]

The principles of CRM are simple. Businesses gather accurate information about customers and prospects. Having identified the customers or segments that account for the highest profits, they devise marketing strategies that differentiate between different groups. Greater resources are focused on higher value customers. Every opportunity is used to amass additional information about each client to personalise sales messages and build a closer relationship.

When Mercury Asset Management began to experiment with CRM, they were able to demonstrate that 59% of their profits came from 1% of their customers. Over six months they moved from having a standard type of literature for all their customers and prospects to no fewer than 7,700 types of literature. Digital printing made this personalisation relatively simple. The first stage of this personalisation process was the compilation of comprehensive information about customers. ■

Customer database
A database that details characteristics of customers and prospects that can include names and addresses, geographic, demographic and buying behaviour data.

Many companies see the benefits of compiling comprehensive information about their customers and invest great amounts in developing and maintaining a **customer database**. The customer database for many companies is used to drive all marketing strategies. Customer data can be created by companies from past records, promotional devices such as competitions or direct response advertising. The database is used to stimulate marketing activities, and the response from these activities is fed back to improve and update it. Database marketing is a circular activity where every iteration improves the total value of the database.[3] So, when consumers 'hook up' to an online company, through their PC, their TV or even their mobile phone, they help to develop the customer database. They supply personal details, their choice of products or services and their means of payment. From the knowledge gained from these transactions, new targeted offerings can be formulated, and the nature of the customers' response can be recorded. As the decision-maker learns more about their customers from transaction data, their awareness of gaps in their knowledge becomes more focused. Where those gaps cannot be filled with transaction data, the marketing researcher plays a vital role in the generation and interpretation of bespoke primary data. In the development of good research design, the customer database can be seen as a resource to the marketing researcher when conducting internal secondary data searches.

There is a whole array of different means to electronically capture customer transaction behaviour and even potential customers through their search for information to buy services and products. It is beyond the scope of this text to describe the array of CRM technologies, Internet trading and online business. We therefore will just concentrate on a concept introduced in the last chapter, concerning scanned data. From a basis of scanned data we illustrate how other data sources, including primary data from marketing research studies, can be integrated. This serves as a link to examine how decision-makers and researchers make sense of the masses of customer data that may be collected.

Scanning devices

Scanning device
Technology that reads the UPC code from merchandise by passing it over a laser scanner.

One of the most fundamental technological breakthroughs that has allowed the monitoring of product sales has been the bar code. With **scanning devices** to read bar codes has come the ability to quickly count and analyse sales. If a new product is launched, scanning data can monitor sales on a daily basis, breaking down the sales by advertising region and the type of outlet. The scanning device is an electronic means of observation. Consumers do not answer any questions, do not identify themselves; they merely enjoy the benefits of supermarket queues moving far more quickly compared with the days of checkout assistants manually entering the prices for individual goods in their baskets.

What product bar codes and scanning devices do not do is classify consumers. Classification is fundamental to marketing research techniques and ultimately marketing segmentation techniques. Is the new brand of yoghurt more popular with younger age groups compared with older groups? Have more Calvin Klein shirts been sold to male or female buyers? The following example illustrates a company that realised this limitation of not being able to classify consumers.

example

Know the clients, meet their needs [4]

Japanese convenience store chain 7-Eleven boosted its sales dramatically by following one very simple rule. Shop assistants were instructed to log two pieces of information at the point of sale: the purchasers' gender and their approximate age. Armed with this information, store managers changed their displays and were able to turn over stock several times a day. ∎

111

In the above example, 7-Eleven developed a crude means of linking patterns of sales to consumer types which improved decision-making and achieved better results. The next technological development develops the link from scanned purchasing data to characteristics of consumers. Compared with the 7-Eleven example, it allows a far more sophisticated means of understanding the characteristics of consumers and of linking their characteristics to their purchasing behaviour.

Relating customer data to scanning systems

e x a m p l e
GlobalCash Project

Loyal customers in the bank

In pan-European banks are many divisions that cover the spectrum from small domestic accounts through to large corporate accounts. In opening any account, banks ask many questions to allow their operations and transactions to work smoothly, to assess the creditworthiness of customers and to assess whether there are other services they could offer them. These questions may be asked on Websites, on the telephone or on a face-to-face basis.

Over time, a bank can build up a series of transaction records against a customer record. From this record, for example, banks can see who are their most profitable customers, who buys a range of connected services and who has been the most loyal over the years, to name but a few analyses.

They may find that certain customers have been banking with them for years, loyally saving a tiny sum paid over the counter each week. Such customers may cost a bank more to service than they can make on interest in the sum invested. A bank may look to means of making the transaction costs cheaper, offer other services that may increase profits, or even charge customers a fee to pay in savings over the counter! ■

The essence of the GlobalCash example is that the operational data used to open and service accounts links customer identification to product usage. Any promotional offers, competitive activity, new product offerings, telephone banking services or discounts, to name but a few marketing activities, can be analysed and related to classifications of customer. Banks and financial institutions are in the fortunate position of being able to ask many questions of potential customers before doing business with them. Such questions are primarily posed for security reasons but they also reveal much about customer characteristics. Every transaction or even enquiry can be linked to an individual customer. This allows banks and financial institutions the ability to analyse in great depth the relationships between types of customer and the patterns of their behaviour. Other types of business see this potential but do not have the power to identify the characteristics of each customer related to each transaction. The answer for many businesses has been to use loyalty card schemes.

Loyalty card
At face value, a sales promotion device used by supermarkets, pharmacists, department stores, petrol stations and even whole shopping centres and towns to encourage repeat purchases. For the marketing researcher, the loyalty card is a device that can link customer characteristics to actual product purchases.

The **loyalty card** is the device that supermarkets, pharmacists, department stores, petrol stations and even whole shopping centres and towns have developed in order to link customer characteristics to actual product purchases.

The loyalty card may be offered to customers as they make a purchase in a store. They normally complete an application form which may include their name and address, demographic details, household details, media usage, and even some lifestyle characteristics. Once the customer uses their loyalty card, the products they have purchased are scanned and a link can be made through the 'swiped' card to their characteristics that can then be related to the scanned data of their product purchases. In return, the customer earns 'points' for their total spend and may earn additional points for buying particular products. The points gained may be redeemed for cash, additional purchases or even goods and services in other retailers or restaurants.

From the marketing decision-makers' perspective, many benefits accrue from a loyalty card and product scanning system. The following list summarises the benefits to the marketer.

1 *Profiles of customers can be built up.* The types of individual that are being attracted to a store can be monitored. The returns and contributions made by particular types of customer can be measured. Profiles of the 'ideal' customer type can be built up, and plans developed to attract that type of customer.

2 *Products used and not used.* The types of product that are being bought or not bought can be monitored. From the customer profile, other types of product can be added to the range offered. Cross-selling of related products can be undertaken. Linked to the customer profile, actual customer behaviour can be understood more fully.

3 *Communications that have worked and not worked.* Merchandising displays, money-off coupons, three for the price of two, or a clip-out coupon from a local newspaper, for example, can be linked to individuals and products. The effectiveness of particular types of communication for particular types of consumer can be developed. Reassurance that the customer has made the right decision can be given where the size of purchase warrants it.

4 *Distribution methods can be tailored.* Certain customer types may prefer the convenience of a small store that they visit more than once a week for small purchases of 'staple' goods. Other customer types may shop once a month for the total household. Retailers can have different shop formats for different customers, may develop home delivery programmes or even develop Internet shopping systems.

The above four factors interact to allow marketing decision-makers to redefine their market(s) and the offerings they make to those markets. The iteration of target market definition and marketing mix tailored to those markets is at the heart of strategic marketing.

From the marketing researchers' perspective, many benefits also accrue from a loyalty card and product scanning system. The following list summarises the benefits to the marketing researcher:

1 *One big laboratory.* Experimental methods will be described in Chapter 11 but, in essence, the monitoring of customers, markets and interrelated marketing mix activities allows for many causal inferences to be established. For example, what is the effect, and upon whom, of raising the price of Häagen Dazs ice cream by 10%? What is the effect of inserting a cut-out coupon to give a discount on after-sun lotion, placed in *Cosmopolitan* magazine?

2 *Refining the marketing process.* With time series of responses to planned marketing activities, statistical models of consumer response can be built with associated probabilities of a particular outcome. Likewise, models of the consumer over their lifetime can be built. Again, statistical models can be built with associated probabilities of particular types of product being bought at different stages of a consumer's life.

3 *Developing a clear understanding of 'gaps' in the knowledge of consumers.* The scanner and loyalty card electronically observes behaviour but does not encapsulate attitudinal data. The nature and levels of satisfaction, what is perceived as good quality service, or what brand image is associated with a particular brand of vodka, are examples of attitudinal data. The use of the database helps to identify target populations to measure and the attitudinal data that need to be collected. In all there can be a much greater clarity in the nature of primary marketing research that tackles attitudinal issues.

4 *Linking behavioural and attitudinal data.* If attitudinal data are elicited from consumers, the data gathered can be analysed in its own right. It is possible, however,

to link the gathered data back to the behavioural data in the database. The term of 'fusing' the data from different sources is used. The key to the fusing lies in identifying individual respondents so that one large dataset is built up. The notion of fusing together databases and survey data from different sources is at the heart of building a strong understanding of consumers.

The above benefits show why many marketers and marketing researchers welcome the power of building an iterative customer database, through scanned product purchases and knowledge of customers who make those purchases. There are drawbacks, however, that focus on the nature of the 'loyalty card'. Loyalty card schemes may be viewed more as a sales promotion technique in much the same manner as giving trading stamps, a dividend or coupons to be redeemed after a period of saving rather than as a means to capture customer data. Compared with other sales promotion techniques, the loyalty card incurs huge operating costs.

There are few questions about the huge costs involved in developing and administering loyalty card schemes. For many retailers, however, the investment allows many marketing and marketing research benefits to be realised. It is only when viewed in the light of offering strategic decision-making power and a complement to an integrated marketing information system that such an investment makes sense. One means of integrating scanned data, with knowledge of customers, and of presenting the relationships and analyses in a spatial manner is through the use of geodemographic information systems.

Geodemographic data

One of the main elements of database power illustrated in the preceding section is the linking of different data sources from both scanner data and customer databases. The ability to create those links and to graphically display analyses has been achieved with the development of **geodemographic information systems** (GIS). At a base level, a GIS matches geographic information with demographic information, allowing analyses to be presented on **thematic maps**. This base can be built upon with data from customer databases, databases from other sources, and surveys. The combined data again can be presented on maps and in conventional statistical tables.

The geographic dimension is vital as a base to the system. Growing up and living in different geographical locations has an effect upon what we buy and the nature of our lifestyle. Look at the huge diversity of consumers around Europe! It is easy to see differences in consumers and their spending habits, between countries and regions within countries, between cities and towns and areas within a town, and even between different sides of a street. These differences emerge from a variety of factors. The following list summarises the main factors, using extreme examples in places. With closer analysis, more subtle differences can be seen which will be illustrated later in this chapter.

Geodemographic information system (GIS)
At a base level, a GIS matches geographic information with demographic information. This match allows subsequent data analyses to be presented on maps.

Thematic maps
Maps that solve marketing problems. They combine geography with demographic information and a company's sales data or other proprietary information and are generated by a computer.

1 *Physical geography and climate.* Consumers living in hot Mediterranean climates in villages close to the sea may have many different needs and wants compared with consumers in Scandinavian inner cities.
2 *Economic history, working opportunities.* Consumers who are primarily semi-skilled, working in a declining manufacturing sector, may have many different needs, wants and spending priorities compared with those in a region that attracts recent graduates to work in a burgeoning financial services sector.
3 *Political and legal differences.* Locations with a history of political and legal domination can affect the types of property and subsequently the types of people who

live there. The differences may be national, e.g. with policies that encourage state ownership of property, or tax breaks and discounts so that a rented property may be bought by its tenant. The differences may be regional, e.g. a local council may have structural plans to allow the building of new housing estates for families, on green-field sites on the outskirts of cities.

4 *Demographic make-up.* Regions made up of consumers living in predominantly retirement areas, such as seaside towns, will have many different requirements from regions that are heavily populated with single young people.

5 *Infrastructure links.* Infrastructure can include the means of travelling around an area as well as the nature and quality of leisure, sports and shopping facilities. Areas with different levels and quality of infrastructure attract different types of consumer. Families with two cars who can comfortably drive to facilities have different needs and wants when compared with individuals living alone who own a bicycle but not a car.

6 *Property types.* In different locations particular styles of property may dominate: flats rather than houses, multi-storey rather than low-rise, detached rather than terraced, bungalow rather than house. The type, size, quality and costs of property within an area attract different types of consumer.

Thus, differences can be seen between geographic locations that affect the lifestyle of residents, the array of products and services they buy, their ability to buy different types of products and services, and their hopes, fears and aspirations. The founding premise of a geodemographic information system is that the type of property a consumer lives in says much about their lifestyle and consumption patterns. Property type also encapsulates the other five factors that discriminate between consumers living in different geographic regions. For example, consumers living in small one-bedroom flats over shops in a city centre will tend to have very different lifestyles and consumption patterns from those of consumers living in large detached rural properties. Consumers in different property types have different propensities or probabilities of buying particular goods and services and of undertaking activities that make up their lifestyle. They also have different propensities to use and be exposed to different types of media.

From a marketing decision-making perspective, geography also plays a vital role. Knowing where one's consumers are located affects the means and costs of distribution. For example, should a retail outlet be built to gain the most returns? Which customers will have to pass our competitors in order to get to us? What features and facilities should the outlet have? The location of consumers also affects the means to communicate with them. Are consumers dispersed over a wide area or tightly clustered together? Do they read the same type of newspaper or magazine? Do they watch the same television programmes or films at the cinema?

A map therefore forms the foundation of a geodemographic information system – a map that can identify all properties in a country, all roads, shopping centres and major facilities in towns and cities. On top of a base map can be laid a range of statistical measures. They typically originate from a number of sources and have the common feature of being able to relate to a specific postcode or zip code. An example of such a system is one produced by Experian. They have developed systems for Australia, Belgium, Germany, Great Britain, Hong Kong, Ireland, Japan, New Zealand, Norway, South Africa, Spain, Sweden and the USA (see **www.experian.com**). Obviously the sources and details of data available in each of the above countries differ, as does the legislation that determines what can be stored and analysed on databases. Typically for each country, statistics can be gathered and used to develop individual geodemographic information systems based upon census data, postal address files, electoral registers, consumer credit data, directories of company directors, mail order purchase records, car registrations and data on access to retail outlets.

115

Geodemographic classification
This groups consumers together based on the types of neighbourhood in which they live. If a set of neighbourhoods are similar across a wide range of demographic measures, they may also offer similar potential across most products, brands, services and media.

From the data collected, the purpose is to classify consumers on a geodemographic basis. Experian define a **geodemographic classification** as follows:

Geodemographic classification groups consumers together based on the types of neighbourhood in which they live. If a set of neighbourhoods are similar across a wide range of demographic measures, they will also offer similar potential across most products, brands, services and media.

With the variables chosen for a particular country, i.e. the types of data that are available to build a geodemographic information system, cluster analyses are performed (Chapter 23 details the nature and purpose of cluster analysis). These analyses help to create consumer classifications, based upon the types of property they live in and the propensity of consumers to have certain lifestyles and behave in particular manners. The analyses ensure that each of the descriptions used is reasonably homogeneous in terms of demographic measurements and consumer behaviour. As well as being able to discriminate and describe distinctive groups of consumers, the analyses have to produce 'pictures' of consumers that are meaningful to marketing decision-makers. For Sweden, the resulting analyses have produced a classification of 10 main consumer types. Table 5.1 lists these types and the percentages of each type in the population. Each individual type can be further classified, Sweden having a total of 30 groups. For example, *Elite families* at 7% of the population can be further broken down into *Careerists in terraced houses* at 2.1% of the population, *Elite professionals* at 2.2% and *White collar metropolitan* at 2.6%.

Table 5.1 **Experian classification of the Swedish population**

Classification descriptor	% in Swedish population
Elite families	7.0
Middle income industrial	8.9
Low-middle income in flats	13.2
Well educated in metropolitan areas	13.7
Younger low income	5.5
Pensioners	8.7
Families with high incomes	3.8
High income in villas	4.9
Middle-aged families	10.7
Countryside	23.6

With a geodemographic information system, it is possible to pinpoint where the *Careerists in terraced houses* are located throughout Sweden, whether they are clustered in particular regions or cities, or whether they are dispersed. From such classifications and the data that can be added from other databases, models of consumer behaviour can be developed. The following example illustrates how Experian have developed a model to calculate the potential for retail outlets.

example

Site quality indicators [5]

Experian Goad has launched a system which shows the sort of power that geodemographic information systems can give to client companies, especially in the retail sector. Called Site Quality Indicators, it uses the accuracy of retail location maps to demonstrate the potential an outlet could have. This can be analysed by proximity to anchor stores, such as Marks & Spencer, how many minutes it is away from a car park, or by a variety of other variables. The model can be used to identify existing stores which are performing well or badly and then locations with a similar profile, or to identify why particular profiles are the way they are. The

system pulls in a lot of datasets, from the geographical to the customer specific. From this, models can be built for the particular retailer, which reflect the baselines within their own business. But that means users have to have data on their customers already, and they must be able to manipulate it in order to build their own Site Quality Indicators. ■

The above example also illustrates that the main power of the system comes from being able to add a customer database to the Experian database; this makes the 'picture' of consumers even clearer. Customers can be mapped out to see how far they live from a retail outlet or to see whether they pass a competitor's store to reach a retail outlet. The profile of customers that a company has can be compared with national, regional or city profiles. Data that are captured on customer databases can be mapped out. For example, the ABN AMRO bank can map out which customers have responded to an offer to take out a personal loan at a discounted rate, as well as building up a profile of those who respond. The following example illustrates how Experian's data are merged with customer databases.

Micromarketer goes Dutch

The Dutch Air Miles franchise LMN, whose shareholders include Shell and ABN AMRO bank, use Experian's Micromarketer product to segment existing customers. It has built a database of 2.3 million customers and segments them geodemographically in order to facilitate targeted mailings. LMN's customer data are supplemented by Experian's data before segmentation. Gerard Zandbergen, Mosaic Micromarketer manager of Experian Netherlands, says 'Data comes from several sources. One is the postal service. We also use market research bureaus and lists of private car owners.' Customers are segmented and the data fed into the GIS to provide a geographical element. Offers for visits to local theme parks and vouchers for local stores therefore go to people in specific regions. ■

In addition to customer behaviour being added to the geodemographic system, survey data can also be added. The key that would link the survey to the customer database may be either a named customer or a postcode. An example of the use of survey data may be in car retailing. A car retailer can map out who bought a new car from them. They may be able to profile and map out the types of individual who bought different types of car. The retailer may then profile and map out the buyers who return for servicing or to buy petrol or accessories. The manufacturer of the car sold and the retailer may conduct a satisfaction survey related to characteristics of the car and the service they received. The results of the survey can be analysed by the different Experian types and characteristics, and levels of customer satisfaction can be mapped out. Additional purchases of cars related to satisfaction or customer loyalty can be captured. It can be seen from the above that through the use of a geodemographic information system, profiles of target markets, measures of the success of marketing decisions and the means to model consumer behaviour can all be achieved.

Graphical representations can be made of customer behaviour, their attitudes and their levels of satisfaction. Using these data, the car retailer additionally has the potential to measure the propensity of potential customers in new locations to buy particular types of cars, petrol, accessories and so on.

Linking different types of data

The previous example illustrated how different types of data can be merged and mapped out to represent customer characteristics. One of the main applications of collecting customer data from different sources and linking it together would be to perform segmentation analyses. Examining the means by which target markets can be

segmented, it is clear to see that the five methods as illustrated in Figure 5.1 can be individually utilised or combined to build clearer 'pictures' or profiles of target consumers.

Figure 5.1 gives examples of where data may be obtained from, to help build up profiles of customers and markets. In the example of 'psychographics' or lifestyle measurements, data may be generated from electronic point of sale (EPOS) systems or surveys. In the case of the EPOS collection, the purchasing of particular types of products can indicate characteristics of a lifestyle. In a more direct manner, questions in a survey can help to build a profile of lifestyle behaviour. In its own right, 'lifestyle' can be a valid means of segmenting a market, perhaps positioning products and services to consumers who aspire to a particular lifestyle. However, being able to combine demographic measurements, broader behavioural characteristics and a knowledge of where these consumers live helps to build a 'picture' of consumers that facilitates strong marketing decision-making support.

Figure 5.1 indicates that as one moves from the demographic through to psychological characteristics the measurement process becomes more difficult. Putting aside the differences in techniques to capture 'demography', 'behaviour' or 'psychology', *what* is being captured becomes more difficult as one moves towards psychological variables. If one considers psychological variables that are vital to marketing which could be captured, examples such as satisfaction, loyalty, trust and quality are not as easy to capture as questions such as gender, age or where one lives. Chapter 12 will explore the concept of measurement in more depth, but at this stage consider what 'satisfaction' actually means, and then the problems of measuring that concept in a valid and consistent manner.

Conversely, as the measurements become more difficult to conduct, they add more to the 'picture' of consumer and market profiles. To say that a market is primarily female, aged between 25 and 40 and lives in a detached property with a mortgage, starts to build a 'picture' of target consumers. To add details of their media behaviour, the array of products and services they buy, characteristics of their lifestyle and their expectations helps to build up a rich and, for decision-makers, very useful 'picture' of target consumers.

Examining the variety of data sources that can be used in the interrelated variables that build market profiles, it is clear to see a role for traditional survey work, scanned data, customer data, externally generated secondary data and the use of loyalty cards. There is a clear interdependence among the different data sources with the increased sophistication of decision support systems that allow the 'fusing' of the data to be conducted.[6]
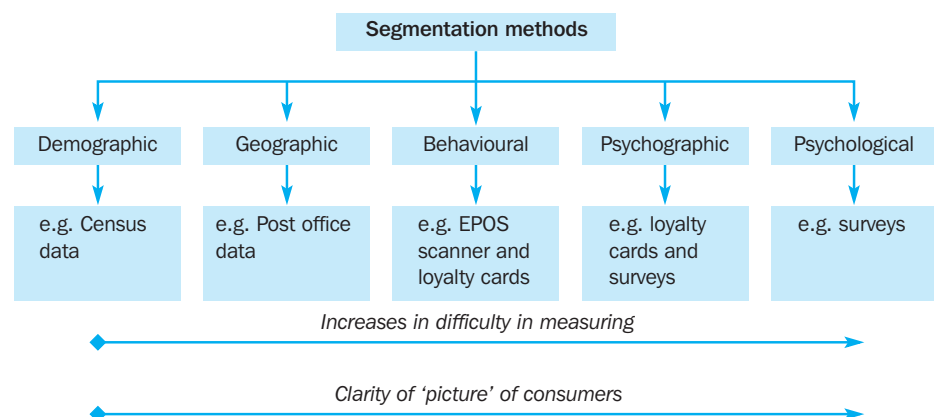
Figure 5.1
Methods of segmenting markets

Stages of development in using databases and survey data to build profiles of consumers and model marketing decisions

The last section discussed how different data could be combined to build strong 'pictures' of consumers. Reflecting upon the role of the marketing researcher in supporting the marketing decision-maker as detailed in Chapter 1, it is clear that the combination of survey data and databases plays a major role in fulfilling the following, helping to:

- describe the nature and scope of customer groups;
- understand the nature of forces that shape the needs of customer groups and the marketer's ability to satisfy those groups;
- test individual and interactive controllable marketing variables;
- monitor and reflect upon past successes and failures in marketing decisions.

The actual implementation of the decision support systems that allow the combination of data sources to be used in supporting decision-makers can take a great deal of time, expense and organisational learning. It is not the intention here to go through the planning, training and organisational issues in making the systems work, but to broadly summarise the stages that an organisation may go through in combining survey and database data. Figure 5.2 summarises the stages of integration; the following descriptions develop the summarised stages in more detail.

1 *Analyse existing consumer database.* These data could include the daily operational transactions or enquiries made to a company. As an internal secondary data source it is the cheapest and most readily available data – providing the organisation culture allows access and analysis to marketing researchers.

2 *Use supplied geodemographic profiles.* There are a growing number of geodemographic systems vendors who have been in operation for over 20 years (as well as the Experian system on **www.experian.com**, have a look at **www.caci.co.uk**). In this time they have been able to refine the data they collect and the analyses they produce to build consumer profiles. Companies can buy a base system 'off the shelf' from systems vendors, and add a variety of different databases.

3 *Combine existing consumer data with geodemographic profiles.* Using the mapping functions of the geodemographic system, existing customer data can be analysed using the profiles supplied with the system. Maps can be used to illustrate the catchment and types of customer and then to evaluate potential in new locations.

4 *Use other surveys (either own or from external sources) that build on geodemographic sources and customer database.* Surveys conducted by a company where either a customer identification or postcode are recorded can be added. Survey data can be analysed using the geographic profiles and analyses represented using maps.

5 *Use combined data sources to create own profiles of customers.* Companies gain experience from using the geodemographic profiles and adding their own data. Over a period of time they may see that the generalised definitions of consumers from the geodemographic system do not accurately represent their existing and target customers. With the benefit and use of their own data, they may take the raw data from the geodemographics vendor and produce their own classifications.

6 *Use of the datawarehouse.* Essentially, this is using many database sources to build one huge database that may be accessed, allowing data to be fused and analysed. This would be done to suit particular reporting requirements or specific queries from either marketing research or marketing managers. With the growth and significance of this development in decision support, the next section describes the datawarehouse in more detail.

| Stage 1: |
| Existing customer database |

↓

| Stage 2: |
| Geodemographic profiles |

↓

| Stage 3: |
| Combine customer data with geodemographic profiles |

↓

| Stage 4: |
| Add survey data |

↓

| Stage 5: |
| Create 'own' consumer profiles |

↓

| Stage 6: |
| Use of the datawarehouse |
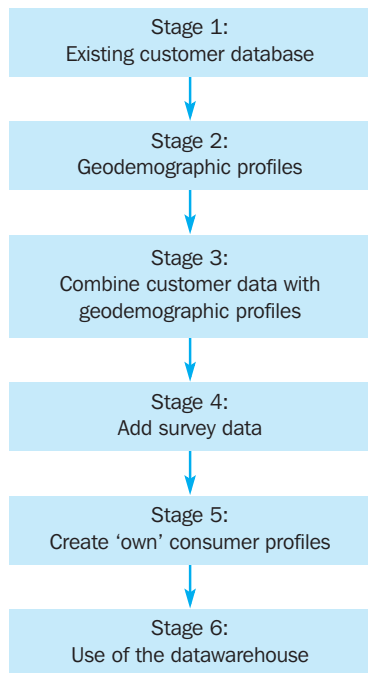
Figure 5.2
Stages of development in using databases and survey data to build profiles of consumers and model marketing decisions

Datawarehouse
This may be seen as a 'super-database', but more specifically it may be defined as a process of gathering disparate data from database and survey sources, and converting it into a consistent format that can aid business decision-making.

The datawarehouse

One of the most prolific users and innovators in **datawarehouses** is the banking industry. The following example illustrates the problems and opportunities for the banking industry of having many different departments with quite distinctive databases.

example

You say 'warehouse', I say 'database' . . .[7]

Three major changes are sweeping through bank marketing: banks are becoming even more customer driven, they are becoming increasingly information rich, and they are now dependent on constantly evolving computer technologies.

Being more customer driven results in a breaking down of the previously hermetically sealed functional areas of banking. Synergistic marketing and sales is now the name of the game. But as these invisible walls come tumbling down, banks are confronting the unintended result of departments with unique informational needs and personal computers: databases of highly valuable information that have no connection with each other. It is as if the tide has gone out, leaving tidal pools teeming with rich data separate from one another along a beach. In this case, the whole of the data really is greater than the sum of its parts. These pools of customer data are not just valuable in and of themselves. The greatest value is in the across-the-board juxtaposition of all the data pools with one another. That's where the confusing conceptual model of a datawarehouse comes in. 'Datawarehousing', then, is simply about the creation of a super-database. ■

The datawarehouse may be seen as a 'super-database', but more specifically it may be defined thus:[8]

A datawarehouse is as much a process of gathering disparate data, converting it into a consistent format that can aid business decision-making, as it is a configuration of software and hardware. Datawarehouses empower users by providing them with access to a whole array of information in an organisation, making it available for use in other applications.

From this definition, the datawarehouse can be described as having the following three qualities:

1 It is a collection of integrated databases designed to support managerial decision-making and problem solving.
2 It essentially becomes a giant database that can include survey data held in a database format.
3 It physically separates an organisation's operational data systems from its decision support systems.

At its most fundamental level, the datawarehouse has three components.

1 *Acquisition.* This includes all the programs, applications and various interfaces that extract data from existing databases. It continues with preparing the data and exporting it to the datawarehouse.
2 *Storage.* This is synonymous with any database. It simply involves a storage area to hold a vast amount of data from a variety of sources. The storage area is organised to make it easy to find and use the data. It will be updated from a variety of sources which could be through scanner and loyalty card data on customers, or through the use of intranet data as described in Chapter 4 when examining the compilation of competitor data.
3 *Access.* This encompasses both set reporting of predetermined events and the means of performing individual analyses, querying 'what-if' scenarios. The process of exploring the databases uses **data mining** techniques. As marketing researchers and decision-makers learn about markets and their effects upon those markets, the development of predictive models is facilitated.

Data mining
The process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition as well as statistical and mathematical techniques.

Data mining

Data mining is a process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories. The process uses pattern recognition as well as statistical and mathematical techniques.[9] Data mining should not be confused with datawarehousing. The datawarehouse could be termed a 'repository' or a place where large amounts of sometimes disparate sources of data are stored; data mining is a process that depends upon access to the data held in that repository.

Examples of what data mining aims to do are as follows.

■ Classify customers into specific categories that are meaningful to decision-makers
■ Identify potential target markets that possess the characteristics that decision-makers seek
■ Forecast sales or the use of services
■ Discover which types of products or services are purchased together
■ Discover patterns and trends over time, such as 'after graduation, students take a holiday', and be able to show the probabilities associated with different holiday types.

Data mining is a way of exploiting the data held by organisations to help discover and develop specific information or knowledge. As well as using proprietary software to perform pattern recognition, statistical and mathematical techniques, it can also be seen as a mental process undertaken by decision-makers. The decision-makers who interact with large datasets using data mining are generally not specialists in statistics, data analysis, datawarehousing and other data tools; they are information users, seeking support for their decision-making processes. With the aid of data mining

software, the decision-maker is encouraged to think in new ways and ask new questions. The decision-maker discovers more from the data and explores in new areas, integrating other sources of data, going through iterative processes to dig deeper. The exploration process involves the discovery of non-trivial relationships of dependence or associations, non-trivial clusters, factors or trends and an understanding of the managerial significance of these discoveries. A data mining process must be 'user-oriented' and that user is typically the decision-maker. The following examples illustrate the processes described above by showing how data mining has been used to explore supermarket shopping and catalogue shopping behaviour.

e x a m p l e

How the data mountain became a mountain of information[10]

Data visualisation is a data mining technique. As an example, a large tracking database covering the entire grocery purchasing of over 7,000 households on a daily basis for a year was searched for common links. Links were presented visually in what can be best described as a sort of 'spirograph' picture. As the data levels were drilled through, the links became more and more specific.

The process started with a 'blank sheet of paper' and quickly identified a strong link between a Sainsbury main shopper and a Tesco shopper. It quickly moved on to discover that these shoppers were strongly linked with grocery purchasing within Tesco and from there cereals were highlighted, then ready-to-eat cereals, then key branded staple lines (Cornflakes, Weetabix, Rice Krispies). The branded staple lines were discovered to be primary drivers for these shoppers.

Used in this way, data visualisation did not answer all their questions but, maybe more importantly, it ensured the right questions could be asked in subsequent studies. What was it about Tesco's branded cereals offering that attracted Sainsbury shoppers? What else were they buying on these trips? How can these shoppers be targeted to bring them back to Sainsburys? ■

e x a m p l e

Data mining at GUS[11]

Great Universal Stores (GUS), Britain's second largest general retailer, mails 14 million fashion catalogues per year and delivers about 395,000 orders daily to customers in the British Isles and Europe. Like other retailers, GUS knows that success depends on ordering the right items in the right quantities at the right time. If customers want items not in stock, they go elsewhere and the company quickly loses money and market share. To gather forecast demand data, GUS mails preview catalogues to 60,000 customers three months before each season begins, with these customers receiving a discount over a two-week period. Analysts use the resulting data to forecast the final demand for each catalogue item. For many years, GUS generated forecasts using multiple regression analyses and mainframe software. These estimates often proved to be too conservative, which resulted in shortages of popular items and unhappy customers.

To overcome this problem, GUS employed Clementine, SPSS's data mining software, that allowed them to perform more rapid and visual modelling of demand. According to Mathew Biddle, a Senior Project Analyst at GUS, 'On the mainframe, we had no graphical way of seeing patterns and trends. Clementine enabled us to explore new variables and quickly build new models. In the process, we generated further ideas.'

The analysis performed using Clementine was repeated and developed for several seasons, and comparisons were made to the existing multiple regression analyses on the mainframe. The data mining software produced superior forecasts, achieving a 20% improvement in assigning each forecast to a right risk level. The result of using data mining techniques has meant that popular items are more likely to be in stock, so customers are satisfied with GUS's service and less vulnerable to appeals from their competitors. ■

Further cases of data mining analytical techniques can be seen on **www.spss.com/ spssatwork/**. (Look out for the Clementine software.)

Databases and marketing research

There has been a phenomenal growth in the use of databases to support marketing decision-making. In larger organisations with many divisions or where mergers and acquisitions have taken place, the datawarehouse has facilitated the 'fusing' of data from many sources. Such developments are seen as a threat by many in the marketing research industry. However, many marketing research companies and marketing research departments within companies are embracing database techniques, utilising the synergistic benefits of matching database analyses with traditional survey data through data mining. To illustrate this point, consider the following quote from Greg Ward, Development Director for Taylor Nelson, the largest marketing research company in the UK.

The marketing research industry needs to acknowledge that databases are serious products and that both types of information have benefits. If you take the best of both – what we call information based marketing – you get something that is significantly more powerful. The 'them and us' situation does nobody any favours and the idea that the two disciplines bear no resemblance to each other is wrong.

As marketing decision-makers become more willing and able to interrogate databases and to creatively generate their own decision support, this does not mean the end of 'traditional' marketing research. As illustrated earlier when examining types of data that are used to build consumer and market profiles, psychological data play a vital role that is fulfilled by qualitative and quantitative marketing research. The marketing researcher needs to develop a greater awareness of both how data captured through traditional methods can be integrated with data held on databases *and* how the combined data creatively support decision-makers.

Databases, the development of datawarehouses and the use of data mining techniques are allowing a wider and shared use of data. The graphical formats of presenting data, especially using maps, break down many barriers in decision-makers who resist formal statistical analyses. They encourage managers to tailor output to meet their individual needs. The creativity that is the hallmark of marketing decision-making is supported by the creative collection and connections between data. Where there are gaps in decision-makers' knowledge, they can be more focused and precise in determining what marketing research support they need. Many marketing researchers are rising to meet this challenge.

The next example illustrates how many companies restructure themselves to integrate databases and traditional marketing research, seeing the disciplines as complementary, not competitive.[12]

example

Raiding the data bank [13]

Moves to overlap database marketing and marketing research are stepping up a gear, as companies like Boots, Royal Bank of Scotland and Sun Life of Canada restructure to integrate the disciplines more closely. Restructuring ranges from shifting the furniture so that teams sit together, to full integration of databases and research databases. Many companies are conducting the overlay themselves and redefining the role of marketing research in the process.

In the mid-1990s, Carola Southorn was group marketing services manager at travel and financial services company Sage and foresaw this trend. She spearheaded the development of guidelines for researchers handling databases, a milestone at the time. She says that 'the ethos on the two being very separate has been overtaken by events'.

It is argued that the urge to merge data is driven from the top. Marketing directors are concerned with the quality and relevance of market data, not its source. At the Royal Bank of Scotland both the head of marketing information and research, Maryan Broadbent, and her

database counterpart, Tim Crick, report to the Bank's Director of Retail Marketing, Ian Henderson. Maryan explains: 'It's no good me telling Ian what customers think, and Tim telling him how they behave. We need to know how attitudes and behaviour are related. Ian asks us not to give him independent views but to go away and give him a consolidated picture.' ■

International marketing research

Linking databases and survey data is transforming international marketing research. Within individual companies, customers may be analysed from operational data within a country, showing different patterns of behaviour between different regions or cities and relating that behaviour to their marketing activities. When a company operates across borders, country differences become just another geographical variable.

In deciding to operate or develop in a particular country, companies may buy a geodemographic information system (should one be developed for that country). A GIS may be used as a foundation to add to their operational data. From this point they may go through the stages of database development as laid out in Figure 5.2. The following example illustrates how a company looking at a particular country may use a GIS to great effect.

example

Making data go further

Expanding overseas may sound like a good idea, but ask any marketer who has attempted it and they will tell you how difficult it can be. Knowing the 'lie of the land' does not mean just a quick reconnaissance of the target market, it means detailed analysis. Geodemographical information systems (GIS) can offer a solution. Tesco, which is increasing its presence in Hungary and the Czech Republic by investing €1,100m in three shopping malls in Prague and Budapest, is one of the many companies using GIS to plan overseas projects. The retailer recently said it plans to build six megastores per year in Eastern and Central Europe. Several established map and data providers are meeting the demand created by large users, such as Tesco. Experian, Claritas and Equifax provide GIS packages and data for European markets and for parts of Asia, Australasia and South Africa. ■

Performing analyses within countries is proving most fruitful, provided a base GIS has been established. Problems start where there is no base GIS. In many countries there are great problems in tracking down and combining data sources that can be relied upon. Further, even if reliable data can be located, legislation may make the use of certain data types illegal. In many developing countries, the data needed to build a GIS are sparse. With the data that are available, much experimentation is needed to enable valid classifications that reflect consumer types which are useful to marketers and marketing researchers.

A further problem exists when making comparisons across countries using separate GIS systems for individual countries. Many GIS vendors have experimented and developed classification systems that allow comparisons between European countries. Building such classifications is difficult because of the inherent property differences between countries and differences in the data available for analysis.

The fusing together of different data sources – including customer data, geodemographic data and survey data – illustrates that the issues involved in understanding international markets are no different from those encountered with domestic markets.

Ethics in marketing research

Marketing researchers are confronted by problems posed by the wording of ESOMAR and individual country marketing research associations' codes of conduct.[14] The codes specify that the compilation of lists, registers or databanks of names and addresses for any non-research purpose shall in no way be associated directly or indirectly with marketing research. However, the examples detailed in this chapter show that supporting marketing decision-makers through databases and marketing research can be seen as part of a total information industry. Evidence of the many leading marketing research agencies involved in data collection and analyses through databases illustrates that databases need not be unethical. With due care it is possible to combine marketing research ethics and databases generated through database marketing. There are a growing number of companies that have used marketing research for many years that now combine the traditional role of marketing research manager with a wider role including database management. An essential part of this combined role lies in the management of customer databases, adding survey details to respondents' individual details, at either an individual or an aggregated level.

Given the phenomenal growth of databases in marketing and the support they offer to marketing decision-makers, they are here to stay. With well-planned 'traditional' marketing research integrated into database analyses, the strategic power of consumer and market analyses is phenomenal. If marketers abuse their knowledge of consumers, they stand to do great harm to their brands and corporate image. For example, in bank databases there are many opportunities for the cross-selling of products. Rather than welcoming the approach from another division of a bank, trying to sell insurance to an investment client, there can be a reaction against the approach, affecting the original business. Consumers are now more aware of how

For insurance companies, would 'practising your flying technique' constitute an unhealthy lifestyle?

valuable knowledge of their behaviour is and how it is used by marketers. They are willing to trade this knowledge for the kind of rewards that are gained from the use of their loyalty cards. Marketers are aware of the dangers of abusing the knowledge that their customers impart to them. However, there are issues of civil liberties that cannot be ignored. These are touched upon in the following example.

example

Loyalty for sale[15]

Provided that shoppers like the benefits and do not object to a system which records every bar of chocolate and bottle of gin purchased, no great harm will be done. However, there is a danger that, despite the safeguards of Data Protection Acts, this mass of information on consumers' habits could leak across the networks into unscrupulous hands. Issues of civil liberty would be raised if, for example, insurance companies could use the data to identify people whose purchases indicated an unhealthy lifestyle; or if the police could draw up a list of suspects by monitoring the purchase of specific items or unusual consumption patterns. ■

One of the benefits of the use of geodemographic systems is that in many cases the individual does not have to be identified; the postcode is a sufficient key to make a link between databases. This maintains the marketing research industry's maxim of respondent confidentiality.

Internet and computer applications

At the start of the 1990s, this chapter would not have existed in a marketing research text. The idea of conducting internal secondary data searches and analyses would have merited a paragraph or two as part of the process of developing primary data collection. Since then, the massive technological changes that have made the global use of the Internet commonplace, the increased storage space, speed and analysis capabilities of computers, and the increased sophistication of software collectively have made fundamental changes to the environment in which marketing researchers operate. The collection and analysis of customer data held within businesses have developed enormously and marketing researchers cannot ignore these developments. To see the impact of these developments, read Professional Perspective 8 on the Companion Website, which presents the case of Deutsche Bank by Trevor Merriden, who asks: 'How do you knot together the information of a vast organisation and make hundreds of millions of euros in the process?'

➤➤➤
See Professional Perspective 8.

In Chapter 1 we discussed how, for many years, marketing researchers have recognised the competition they face from an array of management consultants, but more recently competition has emerged from raw data providers such as call centres, direct marketing, database marketing and telebusinesses.[16] Much of the new competition has emerged from organisations that have utilised and developed Internet and computer applications and have been able to offer support to decision-makers, in faster, cheaper and more user-friendly formats, though not necessarily more rigorously or ethically. Leading marketing researchers have developed means of integrating many of the new formats for supporting decision-makers with traditional marketing research methods and have maintained their customary rigour and ethical behaviour. In order to get a feel of how decision-makers may be supported by some of the means discussed in this chapter, we recommend that you explore the following Websites. These Websites contain case studies of the applications of decision support.

When working through these cases, consider what 'gaps' may still exist in the knowledge of decision-makers that may be filled only by the use of traditional marketing research techniques.

Two major suppliers of software that help with customer relationship management mentioned at the start of this chapter are SAP **www.sap.com** and AIT **www.ait.co.uk**. They also show how operational data generated in organisations may be integrated.

The main geodemographic information systems to examine are produced by Experian on **www.experian.com**, which covers applications in Britain, France, Germany, Italy, the Netherlands, the USA, Asia Pacific countries and others. The Acorn system on **www.caci.co.uk** is primarily based on British applications, though you can see their international applications and alliances on **www.caci.com**.

A directory of datawarehouse, data mining and decision support resources, which includes software suppliers as well as consultants, can be found on **www.info-goal.com/dmc/dmcdwh.htm**. A French datawarehouse supplier with a case based on the Renault Company can be found on **www.decisionnel.fr**. Further explanation and illustration of the Clementine data mining software referred to in this chapter can be located on **www.spss.com**. Finally, visit **www.decisionscience.co.uk** to see a company that utilises data mining techniques with traditional qualitative and quantitative marketing research techniques.

Summary

The overall tone of this chapter has been to demonstrate that internally generated secondary data offer great opportunities not only for decision-makers but also for researchers. As with all good secondary data sources, they have a major impact upon the conduct and direction of primary data collection, analyses and interpretation.

Databases, including customer operational data, geodemographic data and survey data, are radically changing how marketing decision-making is being supported. There is much debate as to whether the use of databases is compatible with traditional techniques of marketing research. With the junk mail connotations of databases and compromises of respondent anonymity, many marketing researchers may seek to keep the marketing database at arm's length. However, handled with the professional acumen that marketing researchers have displayed for many years, the database presents great opportunities for the marketing researcher. In Europe, many of the leading marketing research agencies and research functions within companies have embraced the marketing database.

For the marketer, databases help to build profiles of consumers, linked to the products, communications and distribution methods those consumers favour. For the marketing researcher, databases can present the opportunity to experiment in 'one big laboratory', build models of consumer behaviour, develop an understanding of the gaps in knowledge of consumers and make links between behavioural and attitudinal data.

Much of the data that offers these benefits has been gained using data that capture customer buying behaviour. The use of the 'loyalty card' is one example. Different types of data, including scanner data, loyalty card data and survey data, may be combined using geodemographic information systems (GIS). Using base geographic and demographic data, existing customers can be analysed and mapped out.

Disparate database sources are pulled together through the use of datawarehouses. The datawarehouse integrates databases and survey data, allowing creative connec-

tions between data to be explored. Integrated databases and survey data can be explored using data mining techniques. These processes involve the use of proprietary software and inquisitive, creative decision-makers who search for trends and patterns in customer behaviour and attitudes. The development of datawarehouse and data mining expertise especially helps to cope with the problems of disparate databases and survey data from different countries.

The ethics of using databases provokes much debate in the marketing research industry. As many research practitioners grow more accustomed to using databases, marketing research guidelines and codes of practice are being developed to reflect the good practices that exist in many companies.

Questions

1. How may 'operational data' held in organisations help to build up an understanding of customer behaviour?

2. What is a customer database? Why may a marketing researcher wish to analyse the data held in a customer database?

3. What kinds of data can be gathered through electronic scanner devices?

4. Call in at a supermarket or store that operates a reward or loyalty card scheme that requires you to apply for membership. Pick up an application form and examine the nature of questions you are expected to answer. What marketing research use can be made of the data collected from this application form?

5. Describe the benefits to the *marketing decision-maker* of being able to capture data that identifies characteristics of consumers and their shopping behaviour in a store.

6. Describe the benefits to the *marketing researcher* of being able to capture data that identifies characteristics of consumers and their shopping behaviour in a store.

7. Why may the characteristics of consumers differ, based upon where they live?

8. What is a geodemographic classification of consumers?

9. How can the graphical representation of consumer characteristics using maps help marketing decision-making?

10. What benefits may be gained from fusing together customer characteristics held as internal secondary data, with a proprietary geodemographic information system held as external secondary data?

11. How does the compilation of different types of data help to build a strong 'picture' of consumer characteristics?

12. Describe the stages of development in using databases and survey data to build profiles of consumers and model marketing decisions.

13. What is a datawarehouse?

14. What is the difference between a datawarehouse and data mining?

15. What ethical problems exist with the use of databases that many traditional marketing researchers may find difficult to cope with?

Notes

1 Dawe, A., 'Integration is the thing', *The Times – Special Edition e-CRM* (11 April 2001), 2–3.

2 Sumner-Smith, D., 'Getting to know clients lifts profits', *The Sunday Times* (26 September 1999), 17.

3 Fletcher, K., 'The evolution and use of information technology in marketing', in Baker, M.J. (ed.), *The Marketing Book*, 3rd edn (Oxford: Butterworth-Heinemann, 1994), 352.

4 Hoare, S., 'Know the clients, meet their needs', *The Times – Special Edition e-CRM* (11 April 2001), 9.

5 Reed, D., 'Simply read', *Marketing Week* (2 October 1997).

6 A recent example of fusing research data with databases is presented in Leventhal, B., 'An approach to fusing market research with database marketing', *Journal of the Market Research Society* 39(4) (October 1997), 545–61.

7 Man, D., 'You say "warehouse", I say "database" . . .', *Bank Marketing* 29(4) (April 1997), 37.

8 Man, D., 'The faqs on datawarehousing', *Bank Marketing* 29(4) (April 1997), 38.

9 Jambu, M., 'Data mining for better knowledge', A case study for the telecommunications industry, ESOMAR, Power of Knowledge Conference, Berlin (September 1998).

10 Driver, L., 'How the data mountain became a mine of information', *ResearchPlus* (November 1996), 6.

11 www.spss.com/spssatwork/.

12 McElhatton, N., 'Raiding the data bank', *Research* (September 1999) 28–30. See also the arguments in Brand, C. and Jarvis, S., 'Mind games – the new psychology of research', *Market Research Society Conference*, 2000.

13 McElhatton, N., 'Raiding the data bank', *Research*, (September 1999), 28–30.

14 Hodgson, P., 'Databases: the time for decisions is nigh', *Research* (October 1993), 19.

15 'Loyalty for sale', *Financial Times* (18 September 1998).

16 Savage, M., 'Downstream danger', *Research* (May 2000), 25–7.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Qualitative research:
its nature and approaches

Objectives

After reading this chapter, you should be able to:

1 explain the difference between qualitative and quantitative research in terms of the objectives, sampling, data collection and analysis, and outcomes;

2 describe why qualitative research is used in marketing research;

3 understand the basic philosophical stances that underpin qualitative research;

4 understand the nature and application of ethnographic approaches;

5 understand how qualitative researchers develop theory through a grounded theory approach;

6 explain the potential of action research to qualitative marketing researchers;

7 discuss the considerations involved in collecting and analysing qualitative data collected from international markets;

8 understand the ethical issues involved in collecting and analysing qualitative data.

Qualitative research helps the marketer to understand the richness, depth and complexity of consumers.

Overview

Qualitative research forms a major role in supporting marketing decision-making, primarily as an exploratory design but also as a descriptive design. Researchers may undertake qualitative research to help define a research problem, to support quantitative, descriptive or causal research designs or as a design in its own right. Qualitative research is often used to generate hypotheses and identify variables that should be included in quantitative approaches. It may be used after or in conjunction with quantitative approaches where illumination of statistical findings is needed. In some cases qualitative research designs are adopted in isolation, after secondary data sources have been thoroughly evaluated or even in an iterative process with secondary data sources.

In this chapter, we discuss the differences between qualitative and quantitative research and the role of each in marketing research. We present reasons for adopting a qualitative approach to marketing research (Stage 2 of the marketing research process). These reasons are developed by examining the basic philosophical stances that underpin qualitative research. The concept of ethnographic techniques is presented, with illustrations of how such techniques support marketing decision-makers. The concept of grounded theory is presented, illustrating its roots, the steps involved and the dilemmas for researchers in attempting to be objective and sensitive to the expressions of respondents. Action research is an approach to conducting research that has been adopted in a wide variety of social and management research settings. Action research is developing in marketing research and offers great potential for consumers, decision-makers and researchers alike. The roots of action research are presented, together with the iterative stages involved and the concept of action research teams. The considerations involved in conducting qualitative research when researching international markets are discussed, especially in contrasting approaches between the US and Europe. Several ethical issues that arise in qualitative research are identified.

The first example illustrates a company whose marketing problems based on its products and competitive environment mean that it is well suited to support through qualitative research. The second example illustrates how companies are using techniques adapted from subject areas such as anthropology.

---

example

A research commitment more than skin deep[1]

L'Oréal is the largest supplier of toiletries and cosmetics in the world. The group tucks under its umbrella some of the best known brands and companies in the beauty business: cosmetics houses Lancôme, Vichy and Helena Rubenstein, and fragrance houses Guy Laroche, Cacharel and Ralph Lauren. Given the French penchant for qualitative research, and given the nature of the cosmetics industry, Anne Murray, Head of Research, was asked which type of research she favoured.

> *'We're not particularly pro-quantitative or qualitative. Nevertheless, I do think qualitative in our area is very important. There are many sensitive issues to cover – environmental concerns, animal testing, intimate personal products. And increasingly, we have given to us very technical propositions from the labs, and what is a technical breakthrough to a man in a white coat is not necessarily so to a consumer. So the research department has to be that interface between the technical side and the consumer.'* ■

---

example

Research takes an inventive approach[2]

In qualitative research, research agencies and companies are continually looking to find better ways to understand consumers' thought processes and motivations. This has led to a wealth of new research approaches, including techniques borrowed from anthropology, ethnography, sociology and psychology. Most new techniques have their basis in either observational research or

discussion groups. Advocates of the former believe the consumer tells you only half the story. They stress the importance of watching their behaviour, hence the use of anthropology and ethnography techniques.

Intel has a specialist team of 14 researchers, including ethnographers, anthropologists and psychologists, whose principal form of research is in the home. 'People don't tell you things because they don't think you'll be interested. By going into their homes you can see where and how they use their computers,' says Wendy March, Intel interaction designer of Intel Architecture. ■

These examples illustrate the rich insights into the underlying behaviour of consumers that can be obtained by using qualitative procedures.

Primary data: qualitative versus quantitative research

Qualitative research
An unstructured, primarily exploratory design based on small samples, intended to provide insight and understanding.

Quantitative research
Research techniques that seek to quantify data and, typically, apply some form of statistical analysis.

As explained in Chapter 4, primary data are originated by the researcher for the specific purpose of addressing the problem at hand. Primary data may be **qualitative** or **quantitative** in nature, as shown in Figure 6.1.

Dogmatic positions are often taken in favour of either qualitative or quantitative research by marketing researchers and decision-makers alike. The positions are founded upon which approach is perceived to give the most accurate understanding of consumers. The extreme stances on this issue mirror each other. Many quantitative researchers are apt to dismiss qualitative studies completely as giving no valid findings – indeed as being little better than journalistic accounts. They assert that qualitative researchers ignore representative sampling, with their findings based on a single case or only a few cases. Equally adamant are some qualitative researchers who firmly reject statistical and other quantitative methods as yielding shallow or completely misleading information. They believe that to understand cultural values and consumer behaviour requires interviewing or intensive field observation. Qualitative techniques they see as being the only methods of data collection sensitive enough to capture the nuances of consumer attitudes, motives and behaviour.[3]

Many qualitative researchers believe that intensive field observation is the only way to understand the depth and complexity of consumers.
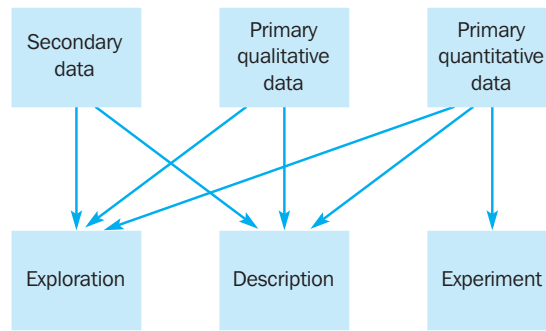
Figure 6.1
A classification of marketing research data

There are great differences between the quantitative and qualitative approaches to studying and understanding consumers. The arguments between qualitative and quantitative marketing researchers about their relative strengths and weaknesses are of real practical value. The nature of marketing decision-making encompasses a vast array of problems and types of decision-maker. This means that seeking a singular and uniform approach to supporting decision-makers by focusing on one approach is futile. Defending qualitative approaches for a particular marketing research problem through the positive benefits it bestows and explaining the negative alternatives of a quantitative approach is healthy – and vice-versa. Business and marketing decision-makers use both approaches and will continue to need both.[4]

The distinction between qualitative and quantitative research can be in the context of research designs as discussed in Chapter 3. There is a close parallel in the distinctions between 'exploratory and conclusive research' and 'qualitative and quantitative research'. There is a parallel, but the terms are not identical. There are circumstances where qualitative research can be used to present detailed descriptions that cannot be measured in a quantifiable manner, for example in describing characteristics and styles of music that may be used in an advertising campaign or in describing the interplay of how families go through the process of choosing, planning and buying a holiday.

Conversely, there may be circumstances where quantitative measurements are used to conclusively answer specific hypotheses or research questions using descriptive or experimental techniques. Beyond answering specific hypotheses or research questions, there may be sufficient data to allow data mining or an exploration of relationships between individual measurements to take place. The concept of data mining illustrated in Chapter 5 allows decision-makers to be supported through exploratory quantitative research.

The nature of qualitative research

Qualitative research encompasses a variety of methods that can be applied in a flexible manner, to enable respondents to reflect upon and express their views or to observe their behaviour. It seeks to encapsulate the behaviour, experiences and feelings of respondents in their own terms and context.

Qualitative research is based on at least two intellectual traditions.[5] The first and perhaps most important is the set of ideas and associated methods from the broad area of depth psychology.[6] This movement was concerned with the less conscious aspects of the human psyche. It led to a development of methods to gain access to individuals' subconscious and/or unconscious levels. So, while an individual may present a superficial explanation of events to themselves or to others, these methods sought to dig deeper and penetrate the superficial.

The second tradition is the set of ideas and associated methods from sociology, social psychology and social anthropology, and the disciplines of ethnography,

linguistics and semiology. The emphases here are upon holistic understanding of the world-view of people. The researcher is expected to 'enter' the hearts and minds of those they are researching, to develop an empathy with their experiences and feelings.

Both traditions have a concern with developing means of communication between the researcher and those being researched. There can be much interaction between the two broad traditions, which in pragmatic terms allows a wide and rich array of techniques and interpretations of collected data.

Qualitative research is a significant contributor to the market research industry, accounting for substantial expenditure, and is growing. In commercial terms, it is a billion-euro plus global industry. However, it is not just a matter of business value. Qualitative thinking has had a profound effect upon marketing and the market research industry as a whole.

Rationale for using qualitative research

It is not always possible, or desirable, to use structured quantitative techniques to obtain information from respondents or to observe them. Thus, there are several reasons to use qualitative techniques. These reasons, either individually or in any combination, explain why certain marketing researchers adopt a particular approach (Stage 2 of the marketing research process) to how they conduct research, analyse data and interpret their findings.

1 *Preferences and/or experience of the researcher.* Some researchers are more oriented and temperamentally suited to do this type of work. Just as some researchers enjoy the challenge of using statistical techniques, there are researchers who enjoy the challenges of qualitative techniques and the interpretation of diverse types of data. Such researchers have been trained in particular disciplines (e.g. anthropology) and philosophies (e.g. hermeneutics) that traditionally make use of qualitative research designs and techniques.

2 *Preferences and/or experience of the research user.* Some decision-makers are more oriented to receiving support in a qualitative manner. This orientation could come from their training but it could also be due to the type of marketing decisions they have to take. Decision-makers working in a creative environment of advertising copy or the development of brand 'personalities', for example, may have a greater preference for data that will feed such 'artistic' decisions. In the following example, consider how decision-makers would get to understand and represent the language used by teenagers. Consider also the implications for a brand if marketers do not fully understand the language and values of their target markets.

e x a m p l e	**First get the language right, then tell them a story**[7]

Teenagers immediately recognise a communication in their language and are very quick to judge whether advertisers have got it right. They see ads and either like them, reject them, ignore them or, in many cases, discuss them. Teenagers are so fluent in 'marketing speak' because marketing and advertising are perceived by them to be the kind of work which can be creative, interesting and acceptable. They discuss with one another the advertising which they perceive to be targeting them.

In the spring and summer of 1997, Pelgram Walters International conducted a study called Global Village. The Global Village hypothesis contended that teenagers around the world have a common language, which speaks to them in the filmed advertising medium. Part of the study consisted of focus group discussions of 12- to 18-year-olds.

Pepsi's Next Generation advertisement was criticised by more media-literate teenage markets (Britain, Germany and the US) for stereotyping teens and misunderstanding who they

are. The ad was a montage of very hip skateboarding teens, male teens wearing make-up, perhaps implying that Pepsi is for the next generation which looks thus. The main complaint was 'we don't look like that', the teens saying that they were not all the same as one another. By aligning the brand image with these extreme images, the commercial was less appealing to mainstream teen consumers. ■

3 *Sensitive information.* Respondents may be unwilling to answer or to give truthful answers to certain questions that invade their privacy, embarrass them, or have a negative impact on their ego or status. Questions that relate to sanitary products and contraception are examples of personally sensitive issues. In industrial marketing research, questions that relate to corporate performance and plans are examples of commercially sensitive issues. Techniques that build up an amount of rapport and trust, that allow gentle probing in a manner that suits individual respondents, can help researchers get close to respondents, and may allow sensitive data to be elicited.

4 *Subconscious feelings.* Respondents may be unable to provide accurate answers to questions that tap their subconscious. The values, emotional drives and motivations residing at the subconscious level are disguised from the outer world by rationalisation and other ego defences. For example, a person may have purchased an expensive sports car to overcome feelings of inferiority. But if asked 'Why did you purchase this sports car?' he may say 'I got a great deal', 'My old car was falling apart', or 'I need to impress my customers and clients.' The respondent does not have to put words to their deeper emotional drives until researchers approach them! In tapping into those deeper emotional drives, qualitative research can take a path that evolves and is right for the respondent.

5 *Complex phenomena.* The nature of what respondents are expected to describe may be difficult to capture with structured questions. For example, respondents may know what brands of wine they enjoy, what types of music they prefer or what images they regard as being prestigious. They may not be able to clearly explain why they have these feelings or where these feelings are coming from.

6 *The holistic dimension.* The object of taking a holistic outlook in qualitative research is to gain a comprehensive and complete picture of the whole context in which the phenomena of interest occur. It is an attempt to describe and understand as much as possible about the whole situation of interest. Each scene exists within a multi-layered and interrelated context and it may require multiple methods to ensure the researcher covers all angles. This orientation helps the researcher discover the interrelationships among the various components of the phenomenon under study. In evaluating different forms of consumer behaviour, the researcher seeks to understand the relationship of different contextual environments upon that behaviour. Setting behaviour into context involves placing observations, experiences and interpretations into a larger perspective.[8] An example of this may be of measuring satisfaction with a meal in a restaurant. A questionnaire can break down components of the experience in the restaurant and quantify the extent of satisfaction with these. But what effect did the 'atmosphere' have upon the experience? What role did the type of music, the colour and style of furniture, aromas coming from the kitchen, other people in the restaurant, the mood when entering the restaurant, feelings of relaxation or tension as the meal went on, contribute to the feeling of atmosphere? Building up an understanding of the interrelationship of the context of consumption allows the qualitative researcher to build up this holistic view. This can be done through qualitative observation and interviewing.

7 *Developing new theory.* This is perhaps the most contentious reason for conducting qualitative research. Chapter 11 details how causal research design through experiments helps to generate theory. Qualitative researchers may argue that there are severe limitations in conducting experiments upon consumers and that quantita-

tive approaches are limited to elaborating or extending existing theory. The development of 'new' theory through a qualitative approach is called 'grounded theory', which will be addressed later.

On the Companion Website see Professional Perspective 14 'Quality research' by Alicia Clegg. Alicia describes and illustrates applications of qualitative research and presents arguments that help in the understanding of why qualitative research may be of use or misleading to decision-makers.

Philosophy[9] and qualitative research

Positivist perspectives

Theory
A conceptual scheme based on foundational statements, or axioms, that are assumed to be true.

Operationalised
The derivation of measurable characteristics to encapsulate marketing phenomena, e.g. the concept of 'customer loyalty' can be operationalised through measurements such as the frequency of repeat purchases or the number of years that a business relationship has existed.

Empiricism
A theory of knowledge. A broad category of the philosophy of science that locates the source of all knowledge in experience.

Positivism
A philosophy of language and logic cosistent with an empiricist philosophy of science.

Paradigm
A set of assumptions consisting of agreed-upon knowledge, criteria of judgement, problem fields, and ways to consider them.

In Chapter 2 we discussed the vital role that **theory** plays in marketing research. Researchers rely on theory to determine which variables should be investigated, how variables should be **operationalised** and measured, and how the research design and sample should be selected. Theory also serves as a foundation on which the researcher can organise and interpret findings. Good marketing research is founded upon theory and contributes to the development of theory to improve the powers of explanation, prediction and understanding in marketing decision-makers.[10]

The dominant perspective of developing new theory in marketing research has been one of **empiricism** and more specifically **positivism**. The central belief of a positivist position is a view that the study of consumers and marketing phenomena should be 'scientific' in the manner of the natural sciences. Marketing researchers of this persuasion adopt a framework for investigation akin to the natural scientist. For many, this is considered to be both desirable and possible. A fundamental belief shared by positivists is the view that the social and natural worlds 'conform to certain fixed and unalterable laws in an endless chain of causation'.[11] The main purpose of a scientific approach to marketing research is to establish causal laws that enable the prediction and explanation of marketing phenomena. To establish these laws, a scientific approach must have, as a minimum, reliable information or 'facts'. The emphasis on facts leads to a focus upon objectivity, rigour and measurement.

As an overall research approach (using the description of a **paradigm** or research approach as developed in Chapter 2) qualitative research does not rely upon measurement or the establishment of 'facts' and so does not fit with a positivist perspective. However, if qualitative research is just seen as a series of techniques, they can be used to develop an understanding of the nature of a research problem, and to develop and pilot questionnaires. In other words, the positivist perspective of qualitative research is to see it as a set of techniques, applied as preliminary stages to more rigorous techniques that measure, i.e. questionnaires. This use of qualitative techniques is fine but may be limiting. To conduct in-depth interviews, focus groups or projective techniques, to understand the language and logic of target questionnaire respondents makes good sense. However, using qualitative techniques just to develop quantitative techniques can affect how those techniques are used. As an illustration, we will examine how focus groups may be conducted.

The term 'focus group discussion' is commonly used across all continents, yet it subsumes different ways of applying the technique. There are two main schools of thought, which may be termed 'cognitive' and 'conative'.

1 *Cognitive.* American researchers generally follow this tradition, which largely follows a format and interviewing style as used in quantitative studies. 'American-style groups' is shorthand in Europe for large groups (10 respondents on average), a structured procedure and a strong element of external validation. Within the cogni-

tive approach, the analysis or articulation has been worked on before, and so the interviews are largely meant to confirm or expand on known issues.

2 *Conative.* European researchers generally follow this tradition. This style assumes a different starting point, one that emphasises exploration, with analysis taking place during and after the group. There is less structure to the questions, with group members being encouraged to take their own paths of discussion, make their own connections and let the whole process evolve.

Table 6.1 summarises the differences between the US (cognitive) and European (conative) approaches to conducting focus groups. Note the longer duration of the European approach to allow the exploration to develop. To maintain the interest and motivation of participants for this time period, the interview experience must be stimulating and enjoyable.

Table 6.1 The two schools of thought about 'focus group discussions'[12]

Characteristics	Cognitive	Conative
Purpose	Demonstration	Exploration
Sample size	10–12	6–8
Duration	1.5 hours	1.5 to 6 hours
Interviewing	Logical sequence	Opportunistic
Questions	Closed	Open
Techniques	Straight question, questionnaires, hand shows, counting	Probing, facilitation, projectives, describing
Response required	Give answers	Debate issues
Interviewer	Moderator	Researcher
Observer's role	To get proof	To understand
Transcripts	Rarely necessary	Usually full
Analysis	On the spot	Time-consuming
Focus of time	Pre-planning	Post-fieldwork
Accusations against other style	'Formless'	'Over-controlling'
Suited for	Testing or proving ideas	Meaning or understanding
Output	To be confirmed in quantitative studies	Can be used in its own right to support decision-makers

International marketers have always been aware that qualitative research as it developed in the US and Europe involves quite different practices, stemming from different premises and yielding different results. American-style qualitative research started from the same evaluative premise as quantitative research but on a smaller scale. This made it cheaper, quicker and useful for checking out the less critical decisions. European-style qualitative research started from the opposite premise to quantitative research: it was developmental, exploratory and creative rather than evaluative. It was used as a tool of understanding, to get underneath consumer motivation.[13]

The American style uses a detailed discussion guide which follows a logical sequence and is usually strictly adhered to. The interviewing technique involves closed questions and straight answers. This type of research is used primarily to inform about behaviour and to confirm hypotheses already derived from other sources. For this reason, clients who have attended groups often feel they do not need any further analysis; the group interaction supplies the answers. Transcripts are rarely necessary and reports are often summarised or even done away with altogether.

The European style is used primarily to gain new insight; it also works from a discussion guide, but in a less structured way. The interviewing technique is opportunistic and probing. Projective techniques are introduced to help researchers understand underlying motivations and attitudes. Because the purpose is understanding, which requires a creative synthesis of (sometimes unconscious) consumer needs and brand benefits, analysis is time-consuming and usually involves full transcripts.

In the above descriptions of American and European traditions of applying qualitative techniques, it is clear to see the American perspective being positivist, i.e. aiming to deliver a 'factual' impression of consumers. The facts may be established, but they may not be enough – they may not provide the richness or depth of understanding that certain marketing decision-makers demand. So, although a positivist perspective has a role to play in developing explanations, predictions and understanding of consumers and marketing phenomena, it has its limitations and critics. The following quote from the eminent qualitative practitioner Peter Cooper cautions us of what we really mean by the term 'qualitative':

> 'There is much qualitative research that still hangs on the positivist model or is little more than investigative journalism. Competition also comes from the media with increasing phone-ins and debates described as "research". We need to be careful about the abuse of what goes under the title "qualitative".'[14]

The dominance of positivist philosophy in marketing research has been and is being challenged by other philosophical perspectives, taken and adapted from disciplines such as anthropology and sociology. These perspectives have helped marketing researchers to develop richer explanations and predictions and especially an understanding and a meaning as seen through the eyes of consumers.

Interpretivist perspectives

In general there are considered to be two main research paradigms that are used by marketing researchers.[15] These are the positivist paradigm and the interpretivist paradigm (though these are by no means the only research paradigms that may be adopted by marketing researchers).[16] Table 6.2 presents alternative names that may be used to describe these paradigms.

Table 6.2 Alternative paradigm names[17]

Positivist	Interpretivist
Quantitative	Qualitative
Objectivist	Subjectivist
Scientific	Humanistic
Experimentalist	Phenomenological
Traditionalist	Revolutionist

Whilst it may be easier to think of these as quite clear, distinct and mutually exclusive perspectives of developing valid and useful marketing knowledge, the reality is somewhat different. There is a huge array of versions of these paradigms, presented by philosophers, researchers and users of research findings. These versions change depending upon the assumptions of researchers and the context and subjects of their study, i.e. the ultimate nature of the research problem. It has been argued[18] that both positivist and interpretivist paradigms are valid in conducting marketing research and help to shape the nature of techniques that researchers apply.

In order to develop an understanding of what an interpretivist paradigm means, Table 6.3 presents characteristic features of the two paradigms.

Table 6.3 Paradigm features[19]

Issue	Positivist	Interpretivist
Reality	Objective and singular	Subjective and multiple
Researcher-respondent	Independent of each other	Interacting with each other
Values	Value-free = unbiased	Value-laden = biased
Researcher language	Formal and impersonal	Informal and personal
Theory and research design	Simple determinist	Freedom of will
	Cause and effect	Multiple influences
	Static research design	Evolving design
	Context-free	Context-bound
	Laboratory	Field/ethnography
	Prediction and control	Understanding and insight
	Reliability and validity	Perceptive decision-making
	Representative surveys	Theoretical sampling
	Experimental design	Case studies
	Deductive	Inductive

Comparison of positivist and interpretivist perspectives

The paradigms can be compared through a series of issues. The descriptions of these issues do not imply that any particular paradigm is stronger than the other. In each issue there are relative advantages and disadvantages specific to any research question under investigation. The issues are dealt with in the following paragraphs.

Reality. The positivist supposes that reality is 'out there' to be captured. It thus becomes a matter of finding the most effective and objective means possible to draw together information about this reality. The interpretivist stresses the dynamic, respondent-constructed and evolving nature of reality, recognising that there may be a wide array of interpretations of realities or social acts.

Researcher–respondent. The positivist sees the respondent as an 'object' to be measured in a consistent manner. The interpretivist may see respondents as 'peers' or even 'companions', seeking the right context and means of observing and questioning to suit individual respondents. Such a view of respondents requires the development of rapport, an amount of interaction and evolution of method as the researcher learns of the best means to elicit information.

Values. The positivist seeks to set aside their own personal values. Their measurements of respondents are being guided by established theoretical propositions. The task for the positivist is to remove any potential bias. The interpretivist recognises that their own values affect how they question, probe and interpret. The task for the interpretivist is to realise the nature of their values and how these affect how they question and interpret.

Researcher language. In seeking a consistent and unbiased means to measure, the positivist uses a language in questioning that is uniformly recognised. This uniformity may emerge from existing theory (to allow comparability of findings) or from their vision of what may be relevant to their target group of respondents. Ultimately, the positivist imposes a language and logic upon target respondents in a consistent manner. The interpretivist seeks to draw out the language and logic of target respondents. The language they use may differ between respondents and develop in different ways as they learn more about a topic and the nature of respondents.

Causality
Causality applies when the occurrence of X increases the probability of the occurrence of Y.

Extraneous variables
Variables other than dependent and independent variables which may influence the results of an experiment.

Determinism
A doctrine espousing that everything that happens is determined by a necessary chain of causation.

Target population
The collection of elements or objects that possess the information sought by the researcher and about which inferences are made.

Reliability
The extent to which a measurement reproduces consistent results if the process of measurement were to be repeated.

Validity
The extent to which a measurement represents characteristics that exist in the phenomenon under investigation.

Case study
A detailed study based upon the observation of the intrinsic details of individuals, groups of individuals and organisations.

Evolving research design
A research design where particular research techniques are chosen as the researcher develops an understanding of the issues and respondents.

Theoretical sampling
Data gathering driven by concepts derived from evolving theory and based on the concept of 'making comparisons'.

Theory and research design. In the development of theory, the positivist seeks to establish **causality** (discussed in detail in Chapter 11) through experimental methods. Seeking causality helps the positivist to explain phenomena and hopefully predict the recurrence of what has been observed in other contexts. There are many **extraneous variables** that may confound the outcome of experiments, hence the positivist will seek to control these variables and the environment in which an experiment takes place. The ultimate control in an experiment takes place in a laboratory situation. In establishing causality through experiments, questions of causality usually go hand in hand with questions of **determinism**, i.e. if everything that happens has a cause, then we live in a determinist universe.

The positivist will go to great pains to diagnose the nature of a research problem and establish an explicit and set research design to investigate the problem. A fundamental element of the positivist's research design is the desire to generalise findings to a **target population**. Most targeted populations are so large that measurements of them can only be managed through representative sample surveys. The positivist uses theory to develop the consistent and unbiased measurements they seek. They have established rules and tests of the **reliability** and **validity** of their measurements and continually seek to develop more reliable and valid measurements.

In the development of theory, the interpretivist seeks to understand the nature of multiple influences of marketing phenomena through **case studies**. The search for multiple influences means focusing upon the intrinsic details of individual cases and the differences between different classes of case. This helps the interpretivist to describe phenomena and hopefully gain new and creative insights to ultimately understand the nature of consumer behaviour in its fullest sense. The consumers that interpretivists focus upon live, consume and relate to products and services in a huge array of contexts, hence the interpretivist will seek to understand the nature and effect of these contexts on their chosen cases. The contexts in which consumers live and consume constitute the field in which the interpretivist immerses themselves to conduct their investigations. In understanding the nature and effect of context upon consumers, the interpretivist does not consider that *everything* that happens has a cause and that we live in a determinist universe. There is a recognition and respect for the notion of free will.

The interpretivist will go to great pains to learn from each step of the research process and adapt their research design as their learning develops. The interpretivist seeks to diagnose the nature of a research problem but recognises that a set research design may be restrictive and so usually adopts an **evolving research design**. A fundamental element of the interpretivist's research design is the desire to generalise findings to different contexts, such as other types of consumer. However, rather than seeking to study large samples to generalise to target populations, the interpretivist uses **theoretical sampling**. This means that the data gathering process for interpretivists is driven by concepts derived from evolving theory, based on the notion of seeking out different situations and learning from the comparisons that can be made. The purpose is to go to places, people or events that will maximise opportunities to discover variations among concepts. The interpretivist uses theory initially to help guide which cases they should focus upon, the issues they should observe and the context of their investigation. As their research design evolves they seek to develop new theory and do not wish to be 'blinkered' or too focused on existing ideas. The interpretivist seeks multiple explanations of the phenomena they observe and creates what they see as the most valid relationship of concepts and, ultimately, theory. Interpretivists seek to evaluate the strength of the theory they develop. The strongest means of evaluating the strength of interpretivist theory lies in the results of decision-making that is based on the theory. Interpretivists continually seek to evaluate the

worth of the theories they develop. A principal output of research generated by an interpretivist perspective should therefore be findings that are accessible and intended for use. If they are found meaningful by decision-makers and employed successfully by them, this may constitute further evidence of the theory's validity. If employed and found lacking, questions will have to be asked of the theory, about its comprehensibility and comprehensiveness and about its interpretation. If it is not used, the theory may be loaded with validity but have little value.

Summarising the broad perspectives of positivism and interpretivism

The positivist seeks to establish the legitimacy of their approach through **deduction**. In a deductive approach, the following process unfolds:

- An area of enquiry is identified, set in the context of well-developed theory, which is seen as vital to guide the researcher, ensuring that they are not naive in their approach and do not 'reinvent the wheel'.
- The issues to focus an enquiry upon emerge from the established theoretical framework.
- Specific variables are identified that the researcher deems should be measured, i.e. hypotheses are set.
- An 'instrument' to measure specific variables is developed.
- Respondents give answers to set and specific questions with a consistent language and logic.
- The responses to the set questions are analysed in terms of a prior established theoretical framework.
- The researcher tests theory according to whether their hypotheses are accepted or rejected. From testing theory in a new context, they seek to incrementally develop existing theory.

Such a process means that positivists reach conclusions based upon agreed and measurable 'facts'. The building and establishment of 'facts' forms the premises of deductive arguments. Deductive reasoning starts from general principles from which the deduction is to be made, and proceeds to a conclusion by way of some statement linking the particular case in question.

A deductive approach has a well-established role for existing theory; it informs the development of hypotheses, the choice of variables and the resultant measures.[20] Whereas the deductive approach starts with theory expressed in the form of hypotheses, which are then tested, an inductive approach avoids this, arguing that it may prematurely close off possible areas of enquiry.[21]

The interpretivist seeks to establish the legitimacy of their approach through **induction**. In an inductive approach, the following process unfolds:

- An area of enquiry is identified, but with little or no theoretical framework. Theoretical frameworks are seen as restrictive, narrowing the researcher's perspective, and an inhibitor to creativity.
- The issues to focus an enquiry upon are either observed or elicited from respondents in particular contexts.
- Respondents are aided to explain the nature of issues in a particular context.
- Broad themes are identified for discussion, with observation, probing and in-depth questioning to elaborate the nature of these themes.
- The researcher develops their theory by searching for the occurrence and interconnection of phenomena. They seek to develop a model based upon their observed combination of events.

Deduction
A form of reasoning in which a conclusion is validly inferred from some premises, and must be true if those premises are true.

Induction
A form of reasoning that usually involves the inference that an instance or repeated combination of events may be universally generalised.

Such a process means that interpretivists reach conclusions without 'complete evidence'. With the intense scrutiny of individuals in specific contexts that typify an interpretivist approach, tackling large 'representative' samples is generally impossible. Thus, the validity of the interpretivist approach is based upon 'fair samples'. The interpretivist should not seek only to reinforce their own prejudice or bias, seizing upon issues that are agreeable to them and ignoring those that are inconvenient. If they are to argue reasonably they should counteract this tendency by searching for conflicting evidence. [22] Their resultant theory should be subject to constant review and revision.

Ethnographic research

It is clear that an interpretive approach does not set out to test hypotheses but to explore the nature and interrelationships of marketing phenomena. The focus of investigation is a detailed examination of a small number of cases rather than a large sample. The data collected are analysed through an explicit interpretation of the meanings and functions of consumer actions. The product of these analyses takes the form of verbal descriptions and explanations, with quantification and statistical analysis playing a subordinate role. These characteristics are the hallmark of a research approach that has developed and been applied to marketing problems over many years in European marketing research. This research approach is one of ethnographic research.

Ethnography
A research approach based upon the observation of the customs, habits and differences between people in everyday situations.

Ethnography as a general term includes observation and interviewing and is sometimes referred to as participant observation. It is, however, used in the more specific case of a method which requires a researcher to spend a large amount of time observing a particular group of people, by sharing their way of life.[23] The origins of ethnography are in the work of nineteenth-century anthropologists who travelled to observe different pre-industrial cultures. An example in a more contemporary context could be the study of death rituals in Borneo, conducted over two years by the anthropologist Peter Metcalf.[24] Today, 'ethnography' encompasses a much broader range of work, from studies of groups in one's own culture, to experimental writing, to political interventions. Moreover, ethnographers today do not always 'observe', at least not directly. They may work with cultural artefacts such as written texts, or study recordings of interactions they did not observe at first hand.[25] In management research, issues related to corporate culture and the management of change are increasingly studied using ethnographic research through observing, interviewing and gathering written material within the organisations of their study. Before we develop an understanding of the approach in marketing research, it is worth summarising the aims of ethnographic research.

- *Seeing through the eyes of.* Viewing events, actions, norms and values from the perspective of the people being studied.
- *Description.* Attending to mundane detail to help understand what is going on in a particular context and to provide clues and pointers to other layers of reality.
- *Contextualism.* The basic message that ethnographers convey is that whatever the sphere in which the data are being collected, we can understand events only when they are situated in the wider social and historical context.
- *Process.* Viewing social life as involving interlocking series of events.
- *Flexible research designs.* Ethnographers' adherence to viewing social phenomena through the eyes of their subjects has led to a wariness regarding the imposition of prior and possibly inappropriate frames of reference on the people they study. This leads to a preference for an open and unstructured research design which increases the possibility of coming across unexpected issues.

■ *Avoiding early use of theories and concepts.* Rejecting premature attempts to impose theories and concepts which may exhibit a poor fit with participants' perspectives.[26] This will be developed further in this chapter when we examine grounded theory.

The use of ethnographic approaches is rapidly developing in marketing research. Decision-makers are finding great support from the process and findings, as the following example illustrates.

Mind games[27]

Since the 1950s, different research approaches have tended to reflect the different beliefs of marketers. Procter & Gamble would have one set of beliefs and Unilever another, and their research approaches would mirror their respective beliefs. So market research tended to be driven by how clients saw the world in general. Because problems have always been identified in clients' terms, research has always been undertaken on that basis, using only clients' vocabulary. The result has been that companies all ask the same questions, get the same answers and end up manufacturing the same products. Now they are asking why this is. The result has been a move away from agenda-driven research and a desire to find out what is actually happening in the world.

One means to discover 'what is happening in the world' is through the use of ethnography. Ethnographic research investigates the way in which people behave in their own environment, and how they interact with the world around them. So it could involve a researcher observing a family at home or watching how women choose clothes in a retail outlet.

Ethnographic research is emphatically not about pursuing a client-driven agenda. Today's market researchers need these anthropological skills to enable them to understand and translate the hidden and not-so-hidden messages, which their subjects reveal. Fewer questions, fewer clipboards, and more waiting and watching. The new breed of market researchers and their informants will be setting their own agendas in future. ■

Ethnography cannot reasonably be classified as just another single method or technique. In essence, it is a research discipline based upon culture as an organising concept and a mix of both observational and interviewing tactics to record behavioural dynamics. Above all, ethnography relies upon entering respondents' natural life worlds – at home, while shopping, at leisure and in the workplace. The researcher essentially becomes a naive visitor in that world by engaging respondents during realistic product usage situations in the course of daily life.

Whether called on-site, observational, naturalistic or contextual research, ethnographic methods allow marketers to delve into actual situations in which products are used, services are received and benefits are conferred. Ethnography takes place not in laboratories but in the real world. Consequently, clients and practitioners benefit from a more holistic and better nuanced view of consumer satisfactions, frustrations and limitations than in any other research method.[28]

A growing trend is for marketers to apply ethnographic methods in natural retail or other commercial environments. There are several objectives that lie behind these studies, one of which is orientated towards a detailed ecological analysis of sales behaviour. In other words, all of the elements that comprise retail store environments – lighting, smells, signage, display of goods, the location, size and orientation of shelving – have an impact upon the consumer experience and their ultimate buying behaviour. The ethnographer's role is to decode the meaning and impact of these ecological elements. Often, these studies utilise time-lapse photography as a tool for behavioural observation and data collection over extensive periods of time and avoid actual interaction with consumers, as illustrated in the following example.[29]

Top of the Pops[30]

The point of purchase (POP) is a manufacturer's last opportunity to have an effect on their customers' decisions. Awareness of the crucial role of in-store influences is growing, and several POP companies have started offering detailed research on how customers react at the point of sale. To achieve this awareness, Electronic Surveillance of Behaviour (ESOB) gives a detailed understanding of how consumers behave in a shop. Kevin Price, Managing Director of *Coutts Design*, has formed a partnership with *The In-Store Audit* to utilise ESOB.

Says Kevin, 'with ESOB, shoppers are tracked remotely on video around the store and their movements and actions are followed. Because this technique is fairly unobtrusive, we are able to capture natural shopper behaviour as people are not being followed around by a researcher.' He goes on to explain the complexity of the computer software. 'The cameras are specially modified and they record a large sample size. They can measure consumer behaviour from entry to exit, following customers around the store and noting the items they touch and the visual cues that they give and get. The cameras may operate for between 10 and 14 days. The information is then analysed and the key clips from the video are used to reinforce the key points that have emerged from the analysis.' ■

One of the key elements of the above example is the context in which the consumer is behaving. The researcher observes shoppers, taking in and reacting to their retail experience, behaving naturally in the set context. The context of shoppers does not just mean the retail outlet they visit. The processes of choosing and buying products, of using products or giving them as gifts, of reflecting upon and planning subsequent purchases are all affected by contextual factors. Context operates on several levels, including the immediate physical and situational surroundings of consumers, as well as language, character, culture and history. Each of these levels can provide a basis for the meaning and significance attached to the roles and behaviour of consumption.

'Can we divorce the ways we buy, use and talk about products from the cultural and linguistic context within which economic transitions occur? The answer is an emphatic *no*'.[31]

The ethnographer may observe the consumer acting and reacting in the context of consumption. They may see a shopper spending time reading the labels on cat food, showing different brands to their partner, engaged in deep conversation, pondering, getting frustrated and putting tins back on the shelf. They may see the same shopper more purposefully putting an expensive bottle of cognac into their shopping trolley without any discussion and seemingly with no emotional attachment to the product. The ethnographer may want to know what is going on. How may the consumer explain their attitudes and motivations behind this behaviour? This is where the interplay of observation and interviewing helps to build such a rich picture of consumers. In questioning the shopper in the above example, responses of 'we think that Remy Martin is the best' or 'we always argue about which are the prettiest cat food labels' would not be enough. The stories and contexts of how these assertions came to be would be explored. The ethnographer does not tend to take simple explanations for activities that in many circumstances may be habitual to consumers. Ethnographic practice takes a highly critical attitude towards expressed language. It challenges our accepted words and utterances at face value, searching instead for the meanings and values that lie beneath the surface. In interviewing situations, typically this involves looking for gaps between expressed and non-verbal communication elements. For example, if actual practices and facial and physical gestures are inconsistent with a subject's expressed attitudes towards the expensive cognac, we are challenged to discover both the reality behind the given answer and the reasons for the 'deception'.

Ethnographic research is also effective as a tool for learning situationally and culturally grounded language, the appropriate words for everyday things as spoken by

various age or ethnic groups. Copywriters and strategic thinkers are always pressed to talk about products and brands in evocative and original ways. Ethnography helps act as both a discovery and an evaluation tool.[32]

The ethnographer is expected to critically analyse the situations they observe. The critique or analysis can be guided by theory but in essence the researcher develops a curiosity, thinks in an abstract manner and at times steps back to reflect and see how emerging ideas connect. By reacting to the events and respondents as they face them, to draw out what they see as important, the ethnographer has the ability to create new explanations and understandings of consumers. This ability to develop a new vision, to a large extent unrestricted by existing theory, is the essence of a grounded theory approach, which is explained and illustrated in the next section.

Grounded theory

Grounded theory
Theory derived from data, systematically gathered and analysed.

The tradition of **grounded theory** was developed by Glaser and Strauss in the late 1950s and published in their seminal work in 1967.[33] At that time, qualitative research was viewed more as impressionistic or anecdotal, little more than 'soft science' or journalism.[34] It was generally believed that the objective of sociology should be to produce scientific theory, and to test this meant using quantitative methods.[35] Qualitative research was seen to have a place, but only to the extent to which it developed questions which could then be verified using quantitative techniques. Glaser and Strauss accepted that the study of people should be scientific, in the way understood by quantitative researchers. This meant that it should seek to produce theoretical propositions that were testable and verifiable, produced by a clear set of replicable procedures. Glaser and Strauss defined theory as:

... theory in sociology is a strategy for handling data in research, providing modes of conceptualisation for describing and explaining. The theory should provide clear enough categories and hypotheses so that crucial ones can be verified in present and future research; they must be clear enough to be readily operationalised in quantitative studies when these are appropriate.[36]

The focus upon developing theory was made explicit in response to criticisms of ethnographic studies that present lengthy extracts from interviews or field observations. Strauss sought to reinforce his view of the importance of theory, illustrated by the following quote:

… much that passes for analysis is relatively low level description. Many quite esteemed and excellent monographs use a great deal of data, quotes or field note selections. The procedure is very useful when the behaviour being studied is relatively foreign to the experiences of most readers or when the factual assertions being made would be under considerable contest by sceptical and otherwise relatively well-informed readers. Most of these monographs are descriptively dense, but alas theoretically thin. If you look at their indexes, there are almost no new concepts listed, ones that have emerged in the course of research.[37]

In contrast to the perhaps casual manner in which some ethnographers may be criticised for attempts at developing theory, the grounded theorist follows a set of systematic procedures for collecting and analysing data. A distinctive feature of the approach is that the collection of data and the analysis that takes place occur simultaneously, with the aim of developing general concepts, to organise data and to integrate these into a more general, formal set of categories. The research process involves progressing through the following stages.[38]

Microscopic examination of data. This is usually completed in a group situation where different individuals focus on sections of interviews and are asked to interpret what they are seeing. These are written up on a flipchart to examine the range of perspectives but also connections among the issues, problems or themes. At this stage these issues are implicit and not systematically worked out.

Coding. This procedure aims to organise data into a set of themes or codes (qualitative coding will be covered in more detail in Chapter 9). This is done firstly by 'open coding'. Researchers are then encouraged to think about different dimensions of the open-coded categories, termed 'dimensionalising', and to find links between categories by 'axial coding'. Different events and situations are observed, to build up a complete picture of the variations within any theoretical category through 'theoretical sampling', i.e. more data would be gathered in a direction driven by the concepts derived from evolving theory and based on the concept of 'making comparisons' – looking for further instances of the derived theory that present a contradictory view, to a point where there are no contradictory views. Eventually, categories are refined so that a theoretical framework emerges, termed 'selective coding'. Coding procedures help in the following manner to:

- build rather than test theory;
- provide researchers with analytic tools for handling masses of raw data;
- help analysts to consider alternative meanings of phenomena;
- be systematic and creative simultaneously;
- identify, develop, and relate the concepts that are the building blocks of theory.[39]

Using qualitative data analysis software. The process of coding described above would typically be conducted using proprietary software. The software essentially helps the researcher to code data in a consistent manner, to search for themes and codes and examine the context in which they emerge in a transcript of an interview or number of interviews. The connection of codes and themes helps to establish the nature of grounded theory. The software allows for the manipulation of codes and themes so that they may be viewed from many different perspectives. From this, the researcher develops further issues and individuals to investigate, and the interpretation and theory emerge. The software helps to manage and manipulate the data, and the researcher uses this as support to develop the theory. (Qualitative data analysis packages are covered in more detail in Chapter 9.)

Presenting diagrams. Grounded researchers use diagrams to represent the relationship of theoretical categories. This helps enormously to focus the individual categories and their interrelationships. Many software packages help to build such diagrams, helping the researcher to visualise their data as they are building up their theory. The use of diagrams is therefore vital in developing theory as well as in the presentation of theory when analysis is complete.

Attempting to gain an objective viewpoint

For the grounded theorist, data collection and analysis occur in alternating sequences. Analysis begins with the first interview and observation, which leads to the next interview or observation, followed by more analysis, more interviews or fieldwork, and so on. It is the analysis that drives the data collection. Therefore there is a constant interplay between the researcher and the research act. Because this interplay requires immersion in the data, by the end of the enquiry the researcher is shaped by the data, just as the data are shaped by the researcher. The problem that arises during this

mutual shaping process is how one can become immersed in the data and still maintain a balance between objectivity and sensitivity. Objectivity is necessary to arrive at an impartial and accurate interpretation of events. Sensitivity is required to perceive the subtle nuances and meanings of data and to recognise the connections between concepts. Both objectivity and sensitivity are necessary for making discoveries. Objectivity enables the researcher to have confidence that his or her findings are a reasonable, impartial representation of a problem under investigation, whereas sensitivity enables creativity and the discovery of new theory from data.[40]

During the analytic process, grounded researchers attempt to set aside their knowledge and experience to form new interpretations about phenomena. Yet, in their everyday lives, they rely on knowledge and experience to provide the means for helping them to understand the world in which they live and to find solutions to problems encountered. Most researchers have learned that a state of complete objectivity is impossible and that in every piece of research, quantitative or qualitative, there is an element of subjectivity. What is important is to recognise that subjectivity is an issue and that researchers should take appropriate measures to minimise its intrusion into their investigations and analyses.

In qualitative research, objectivity does not mean controlling the variables. Rather it means an openness, a willingness to listen and to 'give voice' to respondents, be they individuals or organisations. It means hearing what others have to say, seeing what others do and representing these as accurately as possible. It means developing an understanding of those they are researching, whilst recognising that researchers' understandings are often based on the values, culture, training and experiences that they bring from all aspects of their life; these can be quite different from those of their respondents.[41] As well as being open to respondents, the qualitative researcher reflects upon what makes them, as observers, 'see' and 'listen' in particular ways. This usually means that, while working on a particular project, the researcher keeps a diary or journal. This diary is used to make notes about the conditions of interviews and observations, of what worked well and what did not, of what questions they would have liked to ask but did not think of at the time. As the researcher reads through their diary in the analysis process, the entries become part of the narrative they explore, they reveal to themselves and to others the way they have developed their 'seeing' and 'listening'. Research diaries will be covered in more detail in examining qualitative data analysis in Chapter 9.

Developing a sensitivity to the meanings in data

Having sensitivity means having insight into, and being able to give meaning to, the events and happenings in data. It means being able to see beneath the obvious to discover the new. This quality of the researcher occurs as he or she works with data, making comparisons, asking questions, and going out and collecting more data. Through these alternating processes of data collection and analysis, meanings that are often elusive at first later become clearer. Immersion in the data leads to those sudden insights.[42] Insights do not just occur haphazardly; rather, they happen to prepared minds during interplay with the data. Whether we want to admit it or not, we cannot completely divorce ourselves from who we are and what we know. The theories that we carry around in our heads inform our research in multiple ways, even if we use them quite un-self-consciously.[43]

Ultimately, a grounded theory approach is expected to generate findings that are meaningful to decision-makers, and appropriate to the tasks they face. As with other interpretivist forms of research, if it is found meaningful by decision-makers and employed successfully by them, there is further evidence of the theory's validity. Another qualitative approach that is absolutely meaningful to decision-makers in that its primary focus is to deliver actionable results is called action research.

Action research

Background

The social psychologist Kurt Lewin had a main interest in social change and specifically in questions of how to conceptualise and promote social change. Lewin is generally thought to be the person who coined the term **action research** and gave it meanings that are applicable today.[44] In action research, Lewin envisaged a process whereby one could construct a social experiment with the aim of achieving a certain goal.[45] For example, in the early days of the Second World War, Lewin conducted a study, commissioned by US authorities, on the use of tripe as part of the regular daily diet of American families. The research question was: 'To what extent could American housewives be encouraged to use tripe rather than beef for family dinners?' Beef was scarce and was destined primarily for the troops.

Lewin's approach to this research was to conduct a study in which he trained a limited number of housewives in the art of cooking tripe for dinner. He then surveyed how this training had an effect on their daily cooking habits in their own families. In this case, action research was synonymous with a 'natural experiment', meaning that the researchers in a real-life context invited participants into an experimental activity. This research approach was very much within the bounds of conventional applied social science with its patterns of authoritarian control, but it was aimed at producing a specific, desired social outcome.

The above example can be clearly seen from a marketing perspective. It is easy to see a sample survey measuring attitudes to beef, to tripe, to feeding the family and to feelings of patriotism. From a survey, one can imagine advertisements extolling the virtues of tripe, how tasty and versatile it is. But would the campaign work? Lewin's approach was not just to understand the housewives' attitudes but to engage them in the investigation and the solution – to *change* attitudes and behaviour.

Lewin is credited with coining a couple of important slogans within action research that hold resonance with the many action researchers that practise today. The first is 'nothing is as practical as a good theory' and the second is 'the best way to try to understand something is to change it'. In action research it is believed that the way to 'prove' a theory is to show how it provides in-depth and thorough understanding of social structures, understanding gained through planned attempts to invoke change in particular directions. The appropriate changes are in the proof.

Lewin's work was a fundamental building block to what today is called action research. He set the stage for knowledge production based on solving real-life problems. From the outset, he created a new role for researchers and redefined criteria for judging the quality of the enquiry process. Lewin shifted the researcher's role from being a distant observer to involvement in concrete problem-solving. The quality criteria he developed for judging a theory to be good, focused on its ability to support practical problem-solving in real-life situations.[46]

From Lewin's work has developed a rich and thriving group of researchers who have developed and applied his ideas throughout the world. In management research, the study of organisational change with the understanding and empowerment of different managers and workers has utilised action research to great effect. There has been little application of action research in marketing research, though that is changing. Marketing researchers and marketing decision-makers alike are learning of the nature of action research, the means of implementing it and the benefits it can bestow.

Action research
A team research process, facilitated by a professional researcher(s), linking with decision-makers and other stakeholders who together wish to improve particular situations.

Approach

Action research is a team research process, facilitated by one or more professional researchers, linking with decision-makers and other stakeholders who together wish to improve particular situations. Together, the researcher and decision-makers or stakeholders define the problems to be examined, generate relevant knowledge about the problems, learn and execute research techniques, take actions, and interpret the results of actions based on what they have learned.[47] There are many iterations of problem definition, generating knowledge, taking action and learning from those actions. The whole process of iteration evolves in a direction that is agreed by the team.

Action researchers accept no *a priori* limits on the kinds of research techniques they use. Surveys, statistical analyses, interviews, focus groups, ethnographies and life histories are all acceptable, if the reason for deploying them has been agreed by the action research collaborators and if they are used in a way that does not oppress the participants.

Action research is composed of a balance of three elements. If any one of the three is absent, then the process is not action research.

- *Research.* Research based on any quantitative or qualitative techniques, or combination of them, generates data and, in its analyses and interpretation, shared knowledge.
- *Participation.* Action research involves trained researchers who serve as facilitators and 'teachers' to team members. As these individuals set their action research agenda, they generate the knowledge necessary to transform the situation and put the results to work. Action research is a participatory process in which everyone involved takes some responsibility.
- *Action.* Action research aims to alter the initial situation of the organisation in the direction of a more self-managed and more rewarding state for all parties.

An example of an action research team in marketing terms could include:

- *Marketing researchers*: trained in a variety of qualitative and quantitative research techniques, and with experience of diagnosing marketing and research problems.
- *Strategic marketing managers*: decision-makers who work at a strategic level in the organisation and have worked with researchers, as well as those who have no experience of negotiating with researchers.
- *Operational marketing managers*: decision-makers who have to implement marketing activities. These may be the individuals who meet customers on a day-to-day basis and who really feel the impact and success of marketing ideas.
- *Advertising agency representatives*: agents who have worked with strategic decision-makers. They may have been involved in the development of communications campaigns to generate responses from target groups of consumers.
- *Customers*: existing customers who may be loyal and have had many years of experience of the company (initiating and funding the action research) and its products and perhaps even its personnel.
- *Target customers*: potential customers who may be brand switchers or even loyal customers to competitive companies.

Figure 6.2[48] illustrates how action research may be applied. This model of action research is taken from the subject area of the management of change, which is relevant to many of the problems faced by marketing decision-makers. The process aims to create a learning community in a team such as that above. The team develops an understanding of issues to the extent that they make sound judgements and take effective action to implement the changes they wish to make.
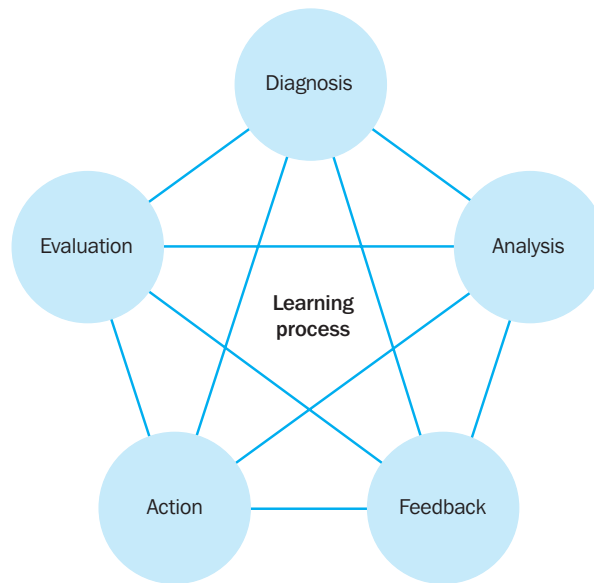
Figure 6.2
The action research approach

The process in Figure 6.2 can be described as follows.

■ *Diagnosis.* The present state of affairs would be set out, including the perceived barriers to change and an initial broad statement of desired direction for the organisation. Diagnosis would include documenting the change process and all data gathering activities such as secondary data gathering, surveys, interviews or observations.

■ *Analysis.* An initial interpretation of data gathered would be made. From this the issues to be tackled would be identified. Summary findings and the development of a framework, with set tasks for team members in subsequent data gathering, would be drawn up.

■ *Feedback.* Data analyses would be fed back for examination and discussion in the team. 'Ownership' of the diagnosis would be developed to formulate a commitment to action.

■ *Action.* Individual courses of action and the development of broader strategies would be formulated.

■ *Evaluation.* There would be an ongoing review of methods and outcomes. The effectiveness of any action would be evaluated against agreed criteria and critical success factors.

All of these stages are interrelated, so there is no definitive path that the team would take. In subsequent iterations of activities, the team could move around the stages in any order that suits their needs.

In marketing research, problems such as 'changing the organisation to have a greater focus on customer satisfaction' or 'examining the means to motivate and reward sales personnel' could be tackled using action research. To see an illustration of action research in marketing, go to the Companion Website and read Professional Perspective 15 'Researching marketing competencies with Action Research' by Stan Maklan and Georgia Field. Stan and Georgia present their own perspectives of why an action research approach was appropriate and beneficial to a novel Internet business.

Ethics in marketing research

The researcher and the client must respect respondents when conducting qualitative research. This should include protecting the anonymity of respondents, not misleading or deceiving them, conducting research in a way not to embarrass or harm the respondents, and using the research results in an ethical manner.[49]

The above advice from the Market Research Society contrasts sharply with a controversial approach to retail research that sometimes goes by the name of 'guerrilla ethnography' or 'street research'. This involves observing and talking with consumers in their natural habitats. The researcher in this case commonly does not identify their role as a researcher, nor do they formally state the objectives behind their interaction with consumers. Instead, through the normal course of chatting with fellow customers or sales personnel, an attempt is made to glean information about customer preferences, sales cues and customer language. The benefit here is that the social distance and formal barriers between researcher and subject are broken down and interaction is more 'natural' and less subject to contrivance. The main objection expressed by critics is the potential invasion of privacy and somewhat manipulative structure of interaction, and the confusion that may be caused by not being absolutely open to those being observed or questioned.[50]

In Professional Perspective 13 on the Companion Website, the Chairman of Heineken exhorts the virtues of observing consumers. He positively encourages his managers to visit bars to see consumers enjoying their products in their natural context. This he sees as part of a natural curiosity that decision-makers should have of their customers. Decision-makers naturally meet many of their customers on a day-to-day basis, especially sales personnel, and much is to be learned from these encounters. As long as these encounters respect the anonymity of respondents, do not mislead or deceive them, and in no way embarrass or harm them, these encounters are a healthy part of conducting business.

It is up to the researcher to be aware of the harm that respondents may suffer and the damage such harm will inflict on all parties involved in an investigation. A reminder of why researchers must be aware of the potential for harm is presented in the following example.

e x a m p l e

The need to pay greater attention to issues of the balance of power, to ensure public cooperation[51]

There are two primary reasons why researchers must be aware of the potential harm to respondents and the need to foster a continual willingness to cooperate in research:

1 Qualitative research is doubly dependent on public cooperation. Not only do we rely on respondents' cooperation in taking part, but the non-directive and open-ended nature of qualitative questioning techniques means that we also rely on their being positively engaged by the process, willing and eager to apply their minds, and happy to reveal their thoughts.
2 The nature of qualitative interviews and groups means that moderators will always try to encourage respondents to reveal more. This brings with it a responsibility to ensure that safeguards are in place, to deter coercion and avoid the potential danger of abuse of power and control over respondents in qualitative research. ■

Internet and computer applications

The Internet presents huge opportunities for the qualitative researcher. Computer-mediated communication (CMC) can be used to run online versions of semi-structured or in-depth interviews, 'observation' of virtual communities, the collection of personal documents, and focus groups. Applying these techniques using the Internet rather than face to face presents a whole array of advantages and challenges.[52] These may be summarised as follows.

Advantages

- *Extending access to participants.* Provided that potential participants have access to the technology, researchers can cross time and space barriers. They can reach a much wider geographical span of respondents, and also populations that are normally hard to reach, such as people with disabilities.
- *Sensitive subjects.* For some respondents the sensitivity of the subject being studied may mean that they would not discuss it in a face-to-face manner. The anonymity and distance can help to draw out views from such respondents.
- *Interest groups.* A variety of online formats, such as chat rooms, mailing lists and conferences, focus on specific topics, drawing together geographically dispersed participants who may share interests, experiences or expertise.
- *Cost and time savings.* Issues such as the time and travelling expenses of researchers, the hire of venues and the costs of producing transcripts can make face-to-face interviewing an expensive option for many researchers, especially if they are using qualitative approaches for the first time. The Internet dramatically reduces or eliminates many of these costs and thus makes qualitative approaches more accessible to a wider array of companies and decision-makers.
- *Handling transcripts.* As interviews or observations are built up through dialogue on the Internet, many of the potential biases or mistakes that occur through audio recordings can be eliminated.

Challenges

- *Computer literacy for the researcher.* Applying qualitative research on the Internet means that some degree of technical expertise is required of the researcher. The extent of expertise depends upon which techniques are being used. For example, moderators of focus groups online will have to learn about the capabilities of the chosen software for running a focus group. They will also have to learn about the specific skills in making the group experience work well online, perhaps by participating in online research conferences or gaining exposure to alternative discussion practices online.
- *Making contact and recruitment.* Establishing contact online requires a mutual exchange of email addresses. There is an array of techniques that can encourage potential respondents to reveal their address (which will be developed in Chapters 7 and 8) but in essence the researcher has to develop rapport and trust to draw the most out of respondents.
- *Interactive skills online.* Even if the researcher develops their skills online, it must be remembered that respondents use their computing equipment with varying degrees of expertise.
- *Losing access.* A key challenge for online studies is to sustain electronic connection with respondents for the whole period of the qualitative research process. This is a reminder that, unlike a survey which may be a short, one-off contact with a respondent, qualitative techniques may unfold and evolve over time and involve returning to respondents as issues develop and theory emerges.

The issues above lay out the broad advantages and challenges of the Internet for the qualitative researcher. As we develop more detailed descriptions and evaluations of qualitative techniques in Chapters 7 to 9, we will examine these points in more detail.

Summary

Qualitative and quantitative research should be viewed as complementary. Unfortunately, many researchers and decision-makers do not see this, taking dogmatic positions in favour of either qualitative or quantitative research. The defence of qualitative approaches for a particular marketing research problem, through the positive benefits it bestows and through explaining the negative alternatives of a quantitative approach, should be seen as healthy, as should the defence of quantitative approaches. Business and marketing decision-makers use both approaches and will continue to need both.

Qualitative and quantitative approaches to marketing research are underpinned by two broad philosophical schools, namely positivism and interpretivism. The central belief of a positivist position is a view that the study of consumers and marketing phenomena should be 'scientific' in the manner of the natural sciences. Marketing researchers of this persuasion adopt a framework for investigation akin to that of the natural scientist. The interpretivist researcher does not set out to test hypotheses but to explore the nature and interrelationships of marketing phenomena. The focus of investigation is a detailed examination of a small number of cases rather than a large sample. The data collected are analysed through an explicit interpretation of the meanings and functions of consumer actions. The product of these analyses takes the form of verbal descriptions and explanations, with quantification and statistical analysis playing a subordinate role.

In examining qualitative approaches, ethnography as a general term includes observation and interviewing and is sometimes referred to as participant observation. It is, however, used in the more specific case of a method which requires a researcher to spend a large amount of time observing a particular group of people, by sharing their way of life. The ethnographer is expected to critically analyse the situations they observe. The critique and the analysis can be guided by theory but in essence the researcher develops a curiosity, thinks in an abstract manner and at times steps back to reflect and see how emerging ideas connect. By reacting to the events and respondents as they face them, to draw out what they see as important, the ethnographer has the ability to create new explanations and understandings of consumers.

Some ethnographers may be criticised in their attempts at developing theory. In response, the grounded theorist follows a set of systematic procedures for collecting and analysing data. A distinctive feature of a grounded theory approach is that the collection of data and its analysis take place simultaneously, with the aim of developing general concepts, to organise data and integrate these into a more general, formal set of categories.

Ethnographic techniques and a grounded theory approach can be applied in an action research framework. Action research is a team research process, facilitated by one or more professional researchers, linking with decision-makers and other stakeholders who together wish to improve particular situations. Together, the researcher and decision-makers or stakeholders define the problems to be examined, generate relevant knowledge about the problems, learn and execute research techniques, take actions, and interpret the results of actions based on what they have learned. There are many iterations of problem definition, generating knowledge, taking action and learning from those actions. The whole process of iteration evolves in a direction that is agreed by the team.

The Internet presents huge opportunities for the qualitative researcher. Computer-mediated communication (CMC) can be used to run online versions of semi-structured or in-depth interviews, 'observation' of virtual communities, the collection of personal

documents, and focus groups. Being able to conduct such techniques online offers great opportunities to the qualitative marketing researcher and also many challenges.

When conducting qualitative research, the researcher and the client must respect respondents. This should include protecting the anonymity of respondents, honouring all statements and promises used to ensure participation, and conducting research in a way not to embarrass or harm the respondents.

Questions

1 What criticisms do qualitative marketing researchers make of the approaches adopted by quantitative marketing researchers, and vice-versa?

2 Why is it not always possible or desirable to use quantitative marketing research techniques?

3 Evaluate the differences between a European and an American approach to qualitative research.

4 Describe the characteristics of positivist and interpretivist marketing researchers.

5 In what ways may the positivist and interpretivist view potential research respondents?

6 What role does theory play in the approaches adopted by positivist and interpretivist marketing researchers?

7 What does ethnographic research aim to achieve in the study of consumers?

8 Why may marketing decision-makers wish to understand the context of consumption?

9 Describe and illustrate two research techniques that may be utilised in ethnographic research.

10 What stages are involved in the application of a grounded theory approach?

11 Is it possible for marketing researchers to be objective?

12 Why may the Kurt Lewin case of action research be deemed an application of marketing research?

13 Describe the key elements to be balanced in the application of action research.

14 Describe the five interrelated phases of an action research approach.

15 What do you see as the key advantages and challenges of conducting qualitative research using the Internet?

Notes

1 McElhatton, N., 'A research commitment more than skin deep', *Research* (May 1994), 10.

2 Clarke, A., 'Research takes an inventive approach', *Marketing* (13 September 2001), 25.

3 Strauss, A. and Corbin, J.M., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd edn (Sage, 1998), 27–34.

4 Cooper, P., 'Consumer understanding, change and qualitative research', *Journal of the Market Research Society* 41(1) (January 1999), 3.

5 Sykes, W., 'Validity and reliability in qualitative market research: a review of the literature', *Journal of the Market Research Society* 32(3), 289.

6 De Groot, G., 'Qualitative research: deeply dangerous or just plain dotty?' *European Research* 14(3) (1986), 136–41.

7 Flaster, C., 'First get the language right, then tell them a story', *ResearchPlus* (November 1997), 11.

8 Gilmore, A. and Carson, D., ' "Integrative" qualitative methods in a service context', *Marketing Intelligence and Planning* 14(6) (June 1996), 21.

9 It is recognised that this topic is treated in a superficial manner and is really a basic introduction to the key ideas. Students who are interested in developing these ideas further should see: Hunt, S.D., *Modern Marketing Theory* (Cincinnati, OH: South Western Publishing, 1991), Potter, G., *The Philosophy of Social Science: New Perspectives* (Harlow:

Prentice Hall, 2000) and Silverman, D., *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction* (London: Sage, 2001).

10 Hunt, S., *Marketing Theory: The Philosophy of Marketing Science* (Homewood, IL: Irwin, 1983), 2.

11 Hughes, J.A., *Sociological Analysis: Methods of Discovery* (London: Nelson, 1976), 19.

12 Goodyear, M., 'Divided by a common language: diversity and deception in the world of global marketing', *Journal of the Market Research Society* 38(2) (April 1996), 105.

13 Broadbent, K., 'When East meets West, quite what does "qualitative" mean?', *ResearchPlus* (March 1997), 14.

14 Cooper, P., 'Consumer understanding, change and qualitative research', *Journal of the Market Research Society* 41(1) (January 1999), 3.

15 Hussey, J. and Hussey, R., *Business Research* (Basingstoke: Macmillan Business, 1997).

16 See Healy, M. and Perry, C., 'Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm', *Qualitative Market Research: An International Journal* 3(3) (2000), 118–26.

17 Hussey, J. and Hussey, R., *Business Research* (Basingstoke: Macmillan Business, 1997).

18 Channon C., 'What do we know about how research works?', *Journal of the Market Research Society* 24(4) (1982), 305–15.

19 Creswell, J.W., *Research Design: Qualitative and Quantitative Approaches* (Thousand Oaks, CA: Sage, 1994).

20 Ali, H. and Birley, S., 'Integrating deductive and inductive approaches in a study of new ventures and customer perceived risk', *Qualitatitive Market Research: An International Journal* 2(2) (1999) 103–10.

21 Bryman, A., *Quantity and Quality in Social Research* (London: Unwin Hyman, 1988).

22 St Aubyn, G., *The Art of Argument* (London: Christophers, 1956), 61–75.

23 Travers, M., *Qualitative Research through Case Studies* (London: Sage, 2001), 4.

24 Metcalf, P., *A Borneo Journey into Death: Berawan Eschatology from its Rituals* (Kuala Lumpur: Abdul Majeed, 1991).

25 Silverman, D., *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*, 2nd edn (London: Sage, 2001), 45.

26 Silverman, D., *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*, 2nd edn (London: Sage, 2001), 46.

27 Reid, T., in Flack, J.-A., 'Mind games', *Marketing Week* (16 August 2001), 41–3.

28 Mariampolski, H., 'The power of ethnography', *Journal of the Market Research Society* 41(1) (January 1999), 79.

29 Mariampolski, H., 'The power of ethnography', *Journal of the Market Research Society* 41(1) (January 1999), 84.

30 Witthaus, M., 'Top of the Pops', *Marketing Week* (21 September 2001), 75–8.

31 Mariampolski, H., 'The power of ethnography', *Journal of the Market Research Society* 41(1) (January 1999), 82.

32 Mariampolski, H., 'The power of ethnography', *Journal of the Market Research Society* 41(1) (January 1999), 81.

33 Glaser, B. and Strauss, A., *The Discovery of Grounded Theory* (Chicago, IL: Aldine, 1967).

34 Blumer, H., 'Sociological analysis and the variable', in Blumer, H., *Symbolic Interactionism: Perspective and Method* (Berkeley, CA: University of California Press, 1969), 127–39.

35 Travers, M., *Qualitative Research through Case Studies* (London: Sage, 2001).

36 Glaser, B. and Strauss, A., *The Discovery of Grounded Theory* (Chicago, IL: Aldine, 1967), 3.

37 Strauss, A., Fagerhaugh, S., Suczek, B. and Wiener, C., *The Social Organisation of Medical Work* (Chicago, IL: University of Chicago Press, 1985).

38 This outline summarises a quite detailed procedure. For a definitive practical guideline to performing grounded theory see: Strauss, A. and Corbin, J.M., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd edn (London: Sage, 1998).

39 Strauss, A. and Corbin, J.M., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd edn (London: Sage, 1998), 7.

40 Strauss, A. and Corbin, J.M., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd edn (London: Sage, 1998), 53.

41 Bresler, L., 'Ethical issues in qualitative research methodology', *Bulletin of the Council for Research in Music Education* 126 (1995), 29–41; Cheek, J., 'Taking a view: qualitative research as representation', *Qualitative Health Research* 6 (1996), 492–505.

42 Strauss, A. and Corbin, J.M., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd edn (London: Sage, 1998), 46–7.

43 Sandelowski, M., 'Theory unmasked: the uses and guises of theory in qualitative research', *Research in Nursing and Health* 16 (1993), 213–18.

44 Greenwood, D.J. and Levin, M., *Introduction to Action Research: Social Research for Social Change* (Thousand Oaks, CA: Sage, 1998), 17.

45 Lewin, K., 'Forces behind food habits and methods of change', *Bulletin of the National Research Council* 108 (1943), 35–65.

46 Greenwood, D.J. and Levin, M., *Introduction to Action Research: Social Research for Social Change* (Thousand Oaks, CA: Sage, 1998), 19.

47 Greenwood, D.J. and Levin, M., *Introduction to Action Research: Social Research for Social Change* (Thousand Oaks, CA: Sage, 1998), 4.

48 Bate, S. P., 'Changing the culture of a hospital: from hierarchy to network', *Public Administration* 78(3) (2000), 487.

49 *MRS Code of Conduct* (The Market Research Society).

50 Mariampolski, H., 'The power of ethnography', *Journal of the Market Research Society* 41(1) (January 1999), 84.

51 Murphy, D., 'Fishing by the Net', *Marketing* (22 August 1996), 25.

52 Mann, C. and Stewart, F., *Internet Communication and Qualitative Research: A Handbook for Researching Online* (London: Sage, 2000), 17–38.

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Qualitative research: focus group discussions

Objectives

After reading this chapter, you should be able to:

1 understand why the focus group is defined as a direct qualitative research technique;

2 describe focus groups in detail, with an emphasis on planning and conducting focus groups;

3 evaluate the advantages, disadvantages and applications of focus groups;

4 describe alternative ways of conducting qualitative research in groups;

5 discuss the considerations involved in conducting qualitative research in an international setting, extending the contrast between European and US traditions of running focus groups;

6 understand the ethical issues involved in conducting focus groups;

7 describe the difference between real-time and non-real-time online focus groups.

The best moderators of focus groups are those that create a spirit of spontaneity and a passion for the issues under discussion.

Overview

In this chapter, we start by presenting a means of classifying qualitative research techniques and we examine the implications of such classification. The characteristics of the focus group are presented along with their advantages and disadvantages. The manner in which focus groups should be planned and conducted is then presented. Running successful focus groups depends upon the skills of a moderator, i.e. the person who manages the group, and the ensuing discussion. We present the qualities needed in moderators to get the most out of focus group discussions. There are variations on the main theme of running a focus group; these are described as well as other qualitative group activities. Running focus groups using the Internet is a rapidly developing technique; we describe how focus groups can be run in 'real-time' and 'non-real-time'. In Chapter 6 we contrasted the purpose and different ways of running focus groups in the US and in Europe; this contrast is developed and illustrated further in examining international marketing research issues. Several ethical issues that arise in running focus groups are identified.

The following example illustrates how using focus groups helps researchers and decision-makers to understand the problems faced by consumers. The problems are expressed in consumers' own words and in ways that words cannot convey. The example also illustrates that researchers and decision-makers may think that they know of all the issues they should be questioning, but once the exploration starts, respondents can reveal new issues that they perceive to be of more importance.

example

There's trouble in classes for kids who wear glasses[1]

Dollond & Aitchison, the optician, which prides itself on its professionalism, decided to look at a small but significant sector of its market – children. Avoiding the tendency of many clients to talk only to parents, the company was determined that children should have their say too.

They commissioned Clarke Research, which does much work with children and families, to conduct qualitative research with children, aged from 5 to 12 years, and their parents. Research was carried out in 12 focus groups, each children's group having four friendship

Finding your glasses easy to lose and uncomfortable is not just a problem for kids.

pairs, and including boys and girls spread by age and geographically. Each children's group was followed by a group of their parents (the children were taken to a local McDonald's while this went on).

Using projective techniques, Clarke Research asked children to describe the experience of wearing specs. This included asking them to draw pictures, to use words and cartoons to describe feelings and attitudes, and to make up mood boards. Some of what was learnt had been anticipated. Children described the experience of losing their specs, forgetting them, finding them uncomfortable, and getting them dirty. What had not been expected was that all the children spoken to had received some degree of harassment, teasing or outright bullying from non-specs wearing children. ■

Classifying qualitative research techniques

A classification of qualitative research techniques is presented in Figure 7.1.

These techniques are classified as either direct or indirect, based on whether the true purpose of the project is known to the respondents. A **direct approach** is not disguised. The purpose of the project is disclosed to the respondents or is otherwise obvious to them from the questions asked. Focus groups and in-depth interviews are the major direct techniques. Even though the purpose of the project is disclosed, the extent to which the purpose of the research is revealed at the start of a focus group or in-depth group may vary. Suppose that the researcher wanted to understand how respondents felt about the Benetton brand, what their views were of Benetton advertising campaigns, the style and quality of Benetton clothes, how 'cool' the brand was, the importance of its being an Italian company – to name but a few issues that could be tackled. Rather than stating these objectives or even that the study was for Benetton right at the start, the researcher may initially hide these issues. If revealed at the start, respondents may focus straight on to these issues and not the surrounding contextual issues that may reveal the 'relative' impact of the Benetton brand. Thus the researcher may initially reveal that the discussion is going to be about 'what clothes mean to you'. The researcher may explore what respondents feel to be good and poor examples of clothing advertisements and why. What types of clothing and accessories do respondents see as stylish, how important is it to wear stylish clothes, how important is it to wear 'cool' clothes? – drawing out examples of brands to illustrate these views. Italy as a country could be explored in terms of characteristics of Italians or Italian design and style. If respondents bring up Benetton in the discussion, the researcher can then focus upon specific questions about the brand, contrast it with other brands and clearly see which subjects generated positive or negative views of Benetton. Respondents may deduce that the study is being conducted for

Direct approach
A type of qualitative research in which the purposes of the project are disclosed to the respondent or are obvious given the nature of the interview.
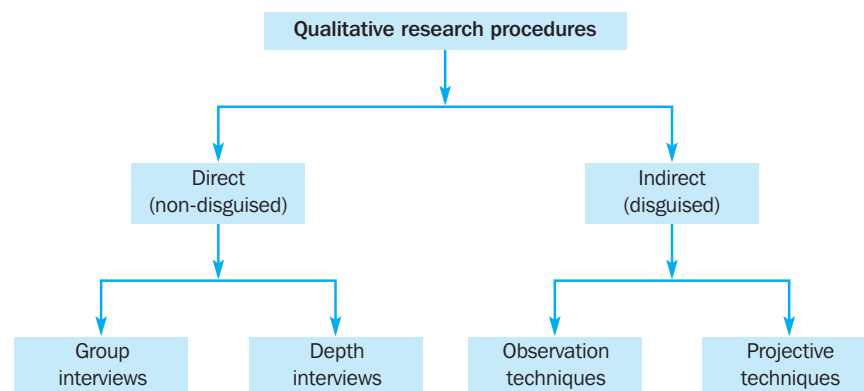
Figure 7.1
A classification of qualitative research techniques

Benetton as the discussion proceeds, this may be apparent by the end of the discussion, or the researcher may clarify this point and explain why it was not revealed at the beginning.

In using focus groups or in-depth interviews, the researcher employs a direct approach but has control over how much 'directness' they reveal at the start of the discussion. The researcher must consider what 'frame of mind' they want respondents to be in at the start of the discussion, as a too narrow or set focus at the start can impede the thought processes and creativity of the respondents and the success of the discussion.

Indirect approach
A type of qualitative research in which the purposes of the project are disguised from the respondents.

In contrast, research that takes an **indirect approach** totally disguises the purpose of the project. In an indirect approach, the researcher wants respondents to behave as naturally as possible without any impediment of research purposes. In observation or ethnographic techniques, consumers may be seen shopping, choosing products, using products, interacting with other people and objects, hopefully in a natural environment and a natural manner. The 'respondent' may not know that they are being observed, or if they do and have agreed to be observed, may not really know why. The purpose of using projective techniques (presented in Chapter 8) is to discover underlying motivations, beliefs, attitudes or feelings regarding consumer behaviour. The techniques allow indirect questioning to allow respondents to discover novel ways to think about and express their feelings, where direct questioning would fail.

Figure 7.1 presents a useful way to remember which qualitative techniques tend towards directness and indirectness. Another way of thinking about these issues would be to visualise a continuum with 'totally direct' at one extreme and 'totally indirect' at the other. Qualitative techniques may then be positioned on this continuum and the implications of that position addressed. The implications for the researcher are as follows.

■ *Ethical.* What are the ethical issues concerning revealing what a study is about? Would a respondent get involved in the study if they knew what it was really about? Will a respondent feel cheated or abused by not being told the purpose or finding it out as they go along?

■ *Data richness.* If the respondent knows what a study is about, to what extent does this 'close' their mind or destroy their creativity? Qualitative techniques aim to draw out deeply held views, issues that may be difficult to conceive or express. Researchers need to be able to get respondents in the right frame of mind to be able to elicit this rich data. To what extent does revealing the purpose of a study impede this process?

The researcher cannot resolve this issue by stating, for example, that 'they will only use direct techniques' to resolve the ethical issues. Successful focus groups and in-depth interviews can utilise certain observation techniques. As an example, consider recording a simple answer of 'no' to a question. This 'no' may be interpreted in different ways, depending upon facial expressions ('was the respondent smiling?'), the tone of their voice ('was it sharp and direct?'), their posture and body language ('were they hunched and hiding their face?'), or their positioning and reactions to others around them ('were they seeking support from others of the same view, by gestures directed towards those respondents?'). The researcher can use and manipulate scenarios to observe respondents as a means of interpreting the answers they give to questions. The same can be said of projective techniques, all of which can be used to great effect in focus groups and in-depth interviews.

The researcher ultimately has to work out the extent of directness or indirectness of their chosen qualitative techniques and address the ethical and data richness issues before setting out the detail of how they will administer their qualitative techniques. These issues may be unique in each investigation, depending upon the nature of

respondents being studied and the questions they face. In practice, the qualitative researcher may resolve the best means to administer a technique by experimenting and adapting; these issues are tackled later in this chapter.

Focus group discussions

Focus group
A discussion conducted by a trained moderator among a small group of respondents in an unstructured and natural manner.

Moderator
An individual who conducts a focus group interview, by setting the purpose of the interview, questioning, probing and handling the process of discussion.

A **focus group** is a discussion conducted by a trained moderator in a non-structured and natural manner with a small group of respondents. A **moderator** leads and develops the discussion. The main purpose of focus groups is to gain insights by creating a forum where respondents feel sufficiently relaxed to reflect and to portray their feelings and behaviour, at their pace and using their language and logic. The value of the technique lies in discovering unexpected findings, often obtained from a free-flowing group discussion. Focus groups are the most important qualitative marketing research procedure, being used extensively in new product development, advertising development and image studies. They are so popular that many marketing research practitioners consider this technique synonymous with qualitative research.[2] Given their importance and popularity, we describe the salient characteristics of focus groups in detail.[3]

Characteristics

The major characteristics of a focus group are summarised in Table 7.1.

Table 7.1 Characteristics of focus groups

Key benefit	Group members 'feed' off each other and creatively reveal ideas that the researcher may not have thought of or dared to tackle
Key drawback	Group members may feel intimidated or shy and may not reveal anything
Group size	6 to 12
Group composition	Homogeneous, respondents pre-screened
Physical setting	Relaxed, informal atmosphere
Time duration	1.5 to 6 hours
Recording	Use of audiocassettes, videotapes and notes from observations
Moderator	Observational, interpersonal and communication skills

One of the main characteristics and key benefits lies in the amount of creative discussion and other activities that may be generated. Group members have the time to reflect upon the discussion and range of stimuli that may be presented to them. The stimuli may come from other group members or from the moderator. Using their intuition and imagination, group members can explain how they feel or behave, in words they are comfortable with and using logic that is meaningful to them. The key drawback lies in how intimidating the group scenario may be to certain individuals. Many individuals may be self-conscious in expressing their ideas, feeling they may be ridiculed by others, or they may be shy and unable to freely express themselves in a group. A focus group is generally made up of 6–12 members. Groups of fewer than six are unlikely to generate the momentum and group dynamics necessary for a successful session. Likewise, groups of more than 12 may be too crowded and may not be conducive to a cohesive and natural discussion.[4] Large groups have a tendency to splinter into sub-groups as group members compete to get their views across.

A focus group generally should be homogeneous in terms of demographic and socio-economic characteristics. Commonality among group members avoids inter-

actions and conflicts among group members on side issues.[5] An amount of conflict may draw out issues or get respondents to rationalise and defend their views in a number of ways; it can also mean that the discussion does not get stale with everybody agreeing with each other and setting a scenario where genuine disagreement gets stifled. However, major conflicts should and can be avoided by the careful selection of respondents. Thus, for many topics, a women's group should not combine married homemakers with small children, young unmarried working women and elderly divorced or widowed women, because their lifestyles are substantially different. Moreover, the respondents should be carefully screened to meet certain criteria. These criteria are set by the researcher to ensure that respondents have had adequate experience with the object or issue being discussed. Other criteria for selection can include characteristics of respondents, such as whether they own their own homes, are of a particular age or even whether they smoke more than 40 cigarettes a day. Respondents who have already participated in numerous focus groups should not be included. These so-called professional respondents are atypical, and their participation leads to serious validity problems.[6]

The physical setting for the focus group is also important. A relaxed, informal atmosphere helps group members to forget they are being questioned and observed. What is meant by a relaxed, informal atmosphere may change depending upon the type of respondent and the subject being tackled. Examples of what 'relaxed and informal' means can include the home of a friend within a particular community, a works canteen, a village hall, a room in a leisure centre, a meeting room in a hotel or a purpose-built discussion group room. The poor acoustics and hard seats of a works canteen may not seem relaxed and informal. To group respondents, however, it may be the place where they are happy to talk and willing to open up to a moderator. The practitioner Wendy Gordon (see Professional Perspective 1 on the Companion Website) contends that the issue of 'real' context as opposed to simulation has always been a thorny issue. She has seen an increase in the tendency to conduct groups wherever the product is seen, bought or used rather than in recruiter living rooms or purpose-built viewing facilities. An example of this could be using part of a furniture store to discuss issues around house decoration, furnishings and cleaning or maintaining the home. Such a setting may set a very strong frame of reference to start the focus group and provide lots of stimuli. This example does not mean that all research needs to be conducted *in situ*, but that the technique can be designed to allow the findings from the real and the research environments to inform the overall recommendations.[7] Light refreshments should be served before the session and made available throughout; these become part of the context of relaxation. The nature of these refreshments largely depends upon how long the discussion lasts, the nature of tasks faced by the respondents and the ethical viewpoint of the researcher.

Although a focus group may last from one to six hours, a duration of one and a half to two hours is typical. When a focus group lasts up to six hours, respondents may be performing a series of projective techniques such as building 'mood boards' or 'role playing'. Lasting this length of time, a break for a meal may be planned. Otherwise, a flow of drinks and snacks may be made available. The opportunities and problems that occur by serving alcoholic drinks in focus groups will be discussed in the 'ethics in marketing research' section later in this chapter.

The lengthy period of discussion in a focus group is needed to establish rapport with the respondents, to get them to relax and be in the right frame of mind, and to explore in depth their beliefs, feelings, ideas, attitudes and insights regarding the topics of concern. Focus group discussions are invariably recorded, mostly using audiotape but often on videotape, for subsequent replay, transcription and analysis. Videotaping has the advantage of recording facial expressions and body movements,

but it can increase the costs significantly. Frequently, where focus groups are conducted in purpose-built studios, decision-makers as 'clients' observe the session from an adjacent room using a two-way mirror or through video transmission. Video transmission technology also enables the clients to observe focus group sessions live from a remote location. Care must be taken with both audio and video recording in terms of how comfortable respondents are with being recorded and what effects recorders have on how much they relax and honestly portray how they feel, especially when projective techniques are used, which some respondents may find embarrassing. Many moderators can give rich examples of how the most interesting points to emerge from a focus group occur when recorders are switched off at the end of the discussion. This happens when a group of respondents have really become involved in the subject of discussion and have enjoyed talking to their fellow respondents. Even when the moderator has finished the discussion, some respondents carry on discussing the issues between themselves as they put their coats on, leave the room and even perhaps as they walk to their cars. Moderators can hear issues discussed in this informal manner that they wish had been tackled with the full group.

The moderator plays a vital role in the success of a focus group. The moderator must establish rapport with the respondents and keep the discussion flowing, including the **probing** of respondents to elicit insights. In addition, the moderator may have a central role in the analysis and interpretation of the data. Therefore, the moderator should possess skill, experience, knowledge of the discussion topic, and an understanding of the nature of group dynamics.

Probing
A motivational technique used when asking questions to induce the respondents to enlarge on, clarify or explain their answers.

Advantages and disadvantages of focus groups

Focus groups offer several advantages over other data collection techniques. These may be summarised by the 10 *S*s:[8]

1 *Synergy*. Putting a group of people together will produce a wider range of information, insight and ideas than will individual responses secured privately.

2 *Snowballing*. A bandwagon effect often operates in a group discussion in that one person's comment triggers a chain reaction from the other respondents. This process facilitates a very creative process where new ideas can be developed, justified and critically examined.

3 *Stimulation*. Usually after a brief introductory period, the respondents want to express their ideas and expose their feelings as the general level of excitement over the topic increases in the group.

4 *Security*. Because the respondents' feelings may be similar to those of other group members, they feel comfortable and are therefore willing to 'open up' and reveal thoughts where they may have been reluctant if they were on their own.

5 *Spontaneity*. Because respondents are not required to answer specific questions, their responses can be spontaneous and unconventional and should therefore provide an accurate idea of their views.

6 *Serendipity*. Ideas are more likely to arise unexpectedly in a group than in an individual interview. There may be issues that the moderator had not thought of. The dynamics of the group can allow these issues to develop and be discussed. Group members, to great effect, may clearly and forcibly ask questions that the moderator may be reluctant to ask.

7 *Specialisation*. Because a number of respondents are involved simultaneously, the use of a highly trained, but expensive, interviewer is justified.

8 *Scientific scrutiny*. The group discussion allows close scrutiny of the data collection process in that observers can witness the session and it can be recorded for later analysis. Many individuals can be involved in the validation and interpretation of the collected data.

9 *Structure.* The group discussion allows for flexibility in the topics covered and the depth with which they are treated. The structure can match the logical structure of issues from the respondents' perspective as well as the language and expressions they are comfortable with.

10 *Speed.* Because a number of individuals are being interviewed at the same time, data collection and analysis proceed relatively quickly.

The disadvantages of focus groups may be summarised by the five *M*s:

1 *Misjudgement.* Focus group results can be more easily misjudged than the results of other data collection techniques. As discussed in Chapter 6, as a qualitative technique, focus groups can evolve through a line of questioning and probing. The specific direction of questioning and the ultimate interpretation of findings can be susceptible to bias.

2 *Moderation.* As well as being great fun to moderate, focus groups can be difficult to moderate. Much depends upon the 'chemistry' of the group in terms of how group members get on with each other and draw ideas and explanations from each other. Even moderators with many years of experience can get into difficulty with particular group members who disrupt the discussion. The quality of the results depends upon how well the discussion is managed and ultimately on the skills of the moderator.

3 *Messiness.* The unstructured nature of the responses makes coding, analysis and interpretation difficult in comparison with the structure of quantitative techniques. Focus group data tend to be messy and need either strong theoretical support or the discipline of a grounded theory approach to ensure that decision-makers can rely upon the analyses and interpretations.

4 *Misrepresentation.* Focus group results concentrate on distinct target groups, describing them and contrasting them to other groups or types of respondent. Trying to generalise to much wider groups, in the same manner as with a quantitative survey based on a representative sample, can be very misleading.

5 *Meeting.* There are many problems in getting potential respondents to agree to take part in a focus group discussion. Even when they have agreed to participate, there are problems in getting focus group respondents together at the same time. Running focus groups on the Internet has helped to resolve these problems to some extent, but for some target groups even this does not offer a solution. An example is in conducting business research with managers as respondents. Given the amount of travel and tight schedules that many managers have, getting them together at the same time is very difficult. With many managers reluctant to reveal their company's behaviour and plans in front of other managers, one can see that the focus group may be very difficult to administer in getting managers to meet up and discuss issues.

Planning and conducting focus groups

The procedure for planning and conducting focus groups is described in Figure 7.2.

Planning begins with an examination of the marketing research problem(s) and objectives. In most instances, the problem has been defined by this stage, but it is vital to ensure that the whole process is founded upon a clear awareness of the gaps in the knowledge of marketing decision-makers. Given the problem definition, the objectives of using focus groups should be clarified. There should be a clear understanding of what information can be elicited and what the limitations of the technique are.

```
┌─────────────────────────────────────────────────────────┐
│        Clarify marketing research problem(s) and objectives        │
└─────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────┐
│      Clarify the role of focus groups in fulfilling those objectives      │
└─────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────┐
│         Specify the issues to be developed in the focus groups         │
└─────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────┐
│       Specify the types of target respondents to make up groups       │
└─────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────┐
│       Specify the location(s) in which to conduct the focus groups       │
└─────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────┐
│                     Recruit group members                     │
└─────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────┐
│                   Run an experimental group                   │
└─────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────┐
│                   Conduct the focus groups                   │
└─────────────────────────────────────────────────────────┘
                            ↓
┌─────────────────────────────────────────────────────────┐
│               Analyse data and present findings               │
└─────────────────────────────────────────────────────────┘
```

Figure 7.2
Procedure for planning and conducting focus groups

Topic guide
A list of topics, questions and probes that are used by a moderator to help them manage a focus group discussion.

The next step is to develop a list of issues, or **topic guide**, that are to be tackled in the focus groups. This list may be a series of specific questions but is more likely to be a set of broad issues that can be developed into questions or probes as the focus group actually takes place. Specific questions may be of help to the moderator who feels that a consistent set of points needs to be presented to different groups in order to allow clear comparisons to be made. Specific questions also act as a 'prop' when the discussion is failing; indeed some group respondents may initially feel that their role is to react to specific questions. However, treating the whole discussion as a means to present set questions may stifle the creativity and spontaneity that are the hallmarks of successful focus groups. The moderator should open the discussion with a general question to make respondents comfortable with the subject and the purpose of the research, and then present specific questions, issues and probes that can develop as the moderator tunes into the dynamics of the group. There may be additional, new issues that develop and, indeed, issues that group members do not see as being appropriate, and these can be discussed. The emphasis should be upon an evolution and learning process rather than administering a consistent set of questions.

The types of group members to take part in the discussions are then specified. From this specification, a questionnaire to screen potential respondents is prepared. Typical information obtained from the questionnaire includes product familiarity and knowledge, usage behaviour, attitudes towards and participation in focus groups, and standard demographic characteristics. With the types of respondents specified, consideration must be taken of what would make them relaxed and comfortable.

Having decided on the location of the focus groups, the actual recruitment of group members progresses. This is one of the most difficult tasks, as potential group members may be sceptical of what may happen at the group, sometimes fearing that they are exposing themselves to a hard-sell campaign of time-share holidays or home improvements! If individuals have attended a focus group beforehand, the process of recruitment is easier, but getting the right type of respondent together at the right

place and time can prove difficult. With the screening questionnaire, recruitment may take place on a face-to-face basis through street interviews or through database details by telephone. One traditional approach is to give the specification of group members to an individual in whose home the discussions are to take place. That individual then recruits respondents who fit that specification from their local community. The advantage of this approach is their ability to persuade respondents that the process is a bona fide research process and is going to be rewarding in many ways; ultimately they make sure that potential respondents actually attend. The big disadvantage is ensuring that those recruited match the screening questionnaire requirements. Whichever method of recruiting respondents is used, even when individuals have said that they will participate in the group, a telephone follow-up is necessary to remind and motivate group members.

Group respondents have to be rewarded for their attendance. Usually they enjoy the experience immensely once they are there, but that does not ensure that they attend in the first place. Attendance can be rewarded with cash, a donation to charity or a gift. The following example illustrates the difficulties involved in recruitment and a marketing researcher's creative solution to the problem.

<table>
<tr><td>example</td><td></td></tr>
</table>

So how do you upstage a Ferrari owner?[9]

Researching an upmarket, socially active audience is difficult at the best of times. The target is opinionated, demanding, often resistant to research and almost impossible to reach. So, when we got the brief to conduct focus groups among Ferrari, Porsche, top Mercedes and other exotic sports car owners, we were tempted to panic. We knew we could find them, but how could we persuade them to participate?

We realised the one thing that would link our target, who were also defined as keen drivers, not just poseurs, was their love of cars and desire to know the latest news about new models (and to try them out if possible).

That's why we decided to offer the carrot of a drive around a race and proving track and the opportunity to meet the design and management team at our famous sports car maker. If anything might motivate people, who clearly already had sufficient money to indulge a very expensive taste, it should be this package. It worked like a dream, and we had great success getting the right people to come and, more importantly, to participate. ■

Experimental group
An initial focus group run to test the setting of the interview, the opening question, the topic guide and the mix of respondents that make up a group.

The first focus group to be run should be seen as an **experimental group**. All aspects of running the group should be evaluated. Were the group members relaxed and comfortable in the chosen location, i.e. did the context work as intended? How did they react to the tape recorder, video or two-way mirror, i.e. at what point did they seem to relax and forget that they were being recorded? What types of member interacted well or not, and what issues helped or hindered interaction? How did the initial question work in opening up and developing the discussion? How did the topic guide work, were there issues missing or issues that individuals would not tackle? How comfortable was the moderator handling the topics, did they have to interject to liven the discussion? How much did the moderator have to know about the subject to have credibility with the respondents? With a reflection of these issues, any necessary alterations can be made to the way that the remaining focus groups are administered. There may be very useful information that emerges from the experimental group that can be included in the main analysis. However, if the group does not work well, the information gleaned may be of little use but the lessons learnt are invaluable in running the remaining groups.

Finally, the focus groups can be actually run. The question arises here of how many groups should be run. Beyond the first experimental group, the number of groups needed can vary. The extent to which comparisons are sought in analyses can determine how many groups are needed. Seeking comparisons means recruiting respondents with different backgrounds or experiences. If there are a great variety of

types of individual that make up a target market, then many homogeneous groups may be needed to reflect the variety of types, e.g. a group of 18–25-year-old single male car owners compared with groups of women or groups of older males, married men or non-car owners. The definition of these distinct target groups to question is entirely bound in the nature of the research problem.

A confounding factor in the definition of these groups is the extent to which different types of respondents will mix together in a single group. In an experimental focus group that explored attitudes and behaviour related to sports activities in the City of Bath, distinct target groups of respondents did not work well together. Older respondents who participated in more 'gentle' sports activities did not particularly appreciate, listen to or respect the views of younger respondents. In addition, there was a gender split in that male respondents showed very little respect for the views of female respondents. The older male respondents were particularly patronising to younger females. As a result it was decided to run groups that separated younger and older respondents and males and females. This meant an increase in the total number of focus groups conducted but also meant that the remaining focus groups worked well and that the views of distinct target groups were clearly presented.

If target respondents are geographically widespread, then many groups may be needed to represent this diversity. The following example of Disneyland Paris illustrates the comparisons made in focus group analyses. The important markets they concentrated upon were France, Germany and Britain and so focus groups would be needed to represent each country. Further analyses of two distinct target groups were also needed. Thus, a minimum of two groups per country, i.e. six groups plus an experimental group, would be needed, which would be doubled if it were felt to be important to run exclusively male and female groups.

example

An MR boost helps Disneyland Paris take off into profit[10]

Back in the summer of 1994 it's fair to say the team at Disneyland Paris were having a tough time of it: they were off the launch pad, but not yet up to full speed. Two years after opening, attendance figures, although impressive, had not lived up to expectations and there was even press speculation about the future of the park.

It was at this time that the Disney team asked The Added Value Company's Paris office to carry out research to understand the potential role of a new attraction. Even by Disney standards, Space Mountain looked as though it would be an amazing experience. It started with a catapult launch in a rocket ship up the outside of a mountain, with the daring space travellers experiencing zero gravity at the top. You were then plunged inside the darkness of the mountain, to travel through space at breathtaking speed, through meteorite storms and past flying asteroids.

Qualitative research groups were used, based on the need to get in-depth and rich consumer reactions to the new concept and to highlight which features were the most motivating. Groups were split between people who had already been to the theme park and those who had not visited, with two targets by lifestyle: parents of children aged 5–15 and young adults aged 25–35. Research was carried out in France, Germany and Britain, with an internal 'data merger' meeting used to share findings from the three countries and develop pan-European findings and conclusions. ■

Another factor to be considered in determining the number of focus groups to conduct relates to whether the researcher is using focus groups as part of a grounded theory approach. With such an approach (as described in Chapter 6) theoretical sampling is adopted whereby further examples or instances are sought that may contradict the nature of the emerging grounded theory. The complexity of the developing grounded theory and the extent to which the qualitative researcher makes sense of the issues they are exploring will ultimately determine the number of discussions needed.

This can be contrasted to focus groups conducted from a positivist perspective, focusing upon generating an understanding of issues to be confirmed in a subsequent survey. In the latter scenario, the researcher may not need to continue to search for more contradictory perspectives. Whichever paradigm underpins the application of focus groups, resources permitting one should conduct additional discussion groups until the moderator can anticipate what will be said. The last sentence is a reminder of the final factor determining the number of discussions – the time and money afforded by the client.

In summary, the number of focus groups that should be conducted on a single subject depends on the following factors.

■ The extent to which comparisons are sought
■ The different types of respondent to be targeted and how well they mix together
■ The geographic spread of respondents
■ The paradigm that underpins the focus group
■ The time and budget available.

Another dimension of running the groups, beyond the actual number of groups, is the nature of stimuli that the moderator chooses to input. At certain times, examples of particular products or brands may be introduced for respondents to examine, or even taste or sample if the nature of the product permits. Advertising material such as brochures, posters or even video recordings of television or cinema adverts can be shown and a response generated. One of the most frequently used forms of stimuli is the **mood board**.

Mood board
A collage created in a focus group setting. Focus group respondents are asked to snip words and pictures from magazines that they see as representing the values a particular brand is perceived to have. In some circumstances, collages can also be made up from audio and video tapes.

The 'mood board' is really the creation of a collage. Focus group respondents are given a pile of magazines and asked to snip words and pictures from them. The direction they are given is to select images and words that they think represent characteristics of a brand, a consumer type or lifestyle or whatever issue the researcher wishes to be illustrated in this manner. The resultant collage can then be used to stimulate discussion and to help draw together ideas and connect them. Creative marketing decision-makers in copywriting or advertisement development may develop many ideas directly from the collages or mood boards. The mood board has two main functions:

■ *Reference point.* The moderator can use the mood board to reflect upon the discussion, in which case issues can emerge that were not so apparent in the heat of a discussion.
■ *Enabling device.* The mood board gets respondents to loosen up and talk more freely. The focus group is not to get respondents to talk rationally but to display what 'feels right' to them. The collage can help to express feelings they may not be able to put into words, or enable those words to have more clarity. This can happen by questioning what is included in a mood board as well as what is omitted.[11]

The following example illustrates that the mood board can develop beyond two-dimensional images.

Art of the matter[12]

Paul Walton, Chairman of new product development consultancy The Value Engineers, explains: 'In the early stages of a new product or brand reassessment project, words might be appropriate. Beyond the words come picture collages and as you learn more and you start to give the brand a clearer identity, you introduce mock-ups of packaging and other props to 'three-dimensionalise' the world.'

Alex Authers, Research Director at branding consultancy New Solutions, argues the case for stimuli beyond the mood board: 'We've moved on from static visuals, to using videos. It's often useful to show a series of fast-edited clips set to some sort of soundtrack. People are

now much more video literate. So, instead of having a mood board, we have a mood video. Video collages are particularly good for exploring the emotional resonances of brands.'

Other props used in qualitative research include swatches of material and even fragrances. Authers says smells and colours can 'help to create a mood and evoke a positioning'. ■

The final stage in planning and conducting focus groups involves the analysis of data. Chapter 9 discusses qualitative data analysis in more detail. However, at this point there are two essential points to note:

1 *Evolving analysis.* Focus groups can change and develop in terms of the issues discussed and the stimuli used to draw out views from respondents. The changes are made as the moderator generates and develops new ideas as each focus group progresses. The moderator makes observations and notes to help them as the discussion progresses and also for when they are over. These observations and notes are part of the total analysis in that they decide which issues to probe, which issues to drop and the form of summarising issues that may be presented to groups at certain stages of the discussion.

2 *Not just the narrative.* If the discussion is recorded then transcripts can be produced which can be analysed with proprietary software. These transcripts form a major part of the analysis procedure but the accumulation and reflection upon observations and notes forms a key part of the analysis.

The moderator

Throughout this chapter we have referred to the moderator as an individual who conducts a focus group discussion, by setting the purpose of the discussion, questioning, probing and handling the process of discussion. This individual may be the researcher handling the project. More likely it will be someone who specialises in the technique, or given the number of groups to run and the time allowed to complete them, a number of specialist moderators will be employed. Whoever is to undertake the task of 'moderating' will require the following qualities: [13]

1 *Kindness with firmness.* The moderator must quickly develop an empathy with group members. From this the moderator should show kindness to make respondents feel welcome, combined with a firmness to stop particular individuals taking over the discussion.

2 *Permissiveness.* The moderator must be permissive, allowing the flow of discussion to develop as the group sees fit. However, he or she must be alert to signs that the group's cordiality or purpose is disintegrating.

3 *Involvement.* The moderator must encourage and stimulate intense personal involvement. In certain circumstances, this may mean becoming involved in the actual discussion itself. This can happen if a tendency for 'group speak' emerges. 'Group speak' happens when little debate or creativity in ideas develops, as particular individuals may not wish to be seen as going against a perceived group norm.

4 *Incomplete understanding.* The moderator must encourage respondents to be more specific about generalised comments by exhibiting a feigned naivety or incomplete understanding.

5 *Encouragement.* The moderator must encourage unresponsive members to participate.

6 *Flexibility.* The moderator must be able to improvise and alter the planned outline amid the distractions of the group process.

7 *Sensitivity.* The moderator must be sensitive enough to guide the group discussion at an intellectual as well as emotional level. He or she must also be attuned to mood changes and issues that fire up enthusiastic responses or conversely cause the discussion to dry up.

8 *Observation.* As the group progresses, notes must be made of ideas or questions to come back to, interpretations of particular silences or bouts of laughter, and how group members are interacting with each other. These observations help the group discussion to progress well and the interpretation of the discussion to have greater meaning.

Other variations of focus groups

Focus groups can use several variations of the standard procedure. These include:

- *Two-way focus group.* This allows one target group to listen to and learn from a related group. In one application, physicians viewed a focus group of arthritis patients discussing the treatment they desired. A focus group of these physicians was then held to determine their reactions.[14]
- *Dual-moderator group.* This is a focus group discussion conducted by two moderators. One moderator is responsible for the smooth flow of the session, and the other ensures that specific issues are discussed.
- *Duelling-moderator group.* Here also there are two moderators, but they deliberately take opposite positions on the issues to be discussed. This allows the researcher to explore both sides of controversial issues. It also encourages respondents who may support a particular perspective to express their views without the fear that they will be 'attacked' by the rest of the group.
- *Respondent-moderator group.* In this type of focus group, the moderator asks selected respondents to play the role of moderator temporarily to improve group dynamics.
- *Client-respondent group.* Client personnel are identified and made part of the discussion group. Their primary role is to offer clarifications that will make the group process more effective.
- *Mini group.* These groups consist of a moderator and only four or five respondents. They are used when the issues of interest require more extensive probing than is possible in the standard group of 6 to 12.

Other types of qualitative group discussions

Brainstorming

Traditional brainstorming[15] has been used for several decades, especially in the context of management or marketing issues. Whether formal or informal, the process is the same: think of as many ideas as you can and say them out loud; leave the evaluation until later; build on and combine others' ideas; be as imaginative as possible, the wilder the ideas the better. The group moderator seeks to nurture an atmosphere of creativity, tapping into the intuition of respondents, generating novel ideas and connections between ideas.

When it works well, ideas flow freely from an interplay that may never have occurred if the group had not brainstormed together.

Two problems plague traditional brainstorming: production blocking and evaluation apprehension.

- *Production blocking* occurs when a group member has an idea, but someone else is talking. When it's their turn, they have forgotten the idea, or think their idea is redundant or not that good. If the group is large or dominated by talkative people, they lose interest and do not say what they think.

■ *Evaluation apprehension* occurs when respondents become anxious about what others think of their thoughts. Ideas may be censored, as there is a fear of being labelled as odd. When respondents feel this apprehension, they do not produce as many new and potentially useful ideas but keep them to themselves and therefore defeat the purpose of brainstorming.

Industrial group discussions

As noted earlier, the focus group has limited use in industrial or business research. Getting managers together at the same time and place is a big operational problem. Getting them to be open about their companies in a group scenario is also very difficult to achieve. In the GlobalCash Project, group discussions were seen as important in order to explore and develop issues that could not be measured in a questionnaire. What was measured in the questionnaire produced many statistics whose analyses pointed to further elaboration. Developing an elaboration of the statistical findings could be tackled in a group discussion. To overcome the above issues of getting managers to sit down together, part of the incentive to complete the questionnaire was an invitation to take part in a 'closed forum for questionnaire respondents'. This meant that a date was set to present findings from the survey exclusively to questionnaire respondents. The date and meeting place of The Management Centre in Brussels were established well in advance to allow managers the chance to put the date in their diaries. The day started with an initial session of presenting statistical findings to the whole group of managers who had responded to the questionnaire. As questions were fielded in this forum, particular topics of interest were identified and focus groups named as 'workshops' were built around these topics. By mid-morning, groups of around 10–12 managers were together tackling a theme of importance to them. They were moderated by an academic researcher from one of the 20 business schools taking part in the GlobalCash Project. With a loose set of topics to develop into questions and probes, the format for a focus group was achieved. In the afternoon, the same format continued with a presentation of questionnaire findings and further group discussions. By splitting the day in this manner, managers could attend and contribute to two subject areas of interest to them. Finally, individual managers could be identified who could be interviewed in depth on their own at a later date. The one-to-one in-depth interviews in the GlobalCash Project will be illustrated in the next chapter.

International marketing research

The term 'focus group' is commonly used across all continents, yet it subsumes approaches that are different. For many European marketing researchers, one of the biggest headaches associated with focus groups is its name. The technique has been disparaged by a popular media impression that implies that focus groups are a means to get answers to set questions. Many see this as the US approach to conducting focus groups and thus see the name as an unwanted Americanism that has displaced what they see as the favoured term, 'group discussion'.

> *The best people at moderating focus groups are ones who can create that spirit of spontaneity. You can't do it with a crowded agenda. You can't do it if you're focused, which is why focus is wrong.* [16]

The above quotation implies a criticism of a US approach that can be highly structured rather than an approach that is truly spontaneous and exploratory in its fullest sense. It is not advocated at this stage that the word 'focus' be removed, but it is a reminder that there are different research approaches that can affect how the

technique is administered. US and European approaches were presented in Chapter 6 to illustrate how different philosophies that underpin research techniques can affect how the technique is applied.[17] In examining the use of focus groups, two main schools of thought were presented: the *cognitive* approach, which largely follows a format and interviewing style as used in quantitative studies and is generally used by American researchers; and the *conative* approach, which has less structure to the questions, with group members being encouraged to take their own paths of discussion, make their own connections and let the whole process evolve, and is generally used by European researchers. It is not advocated that one particular approach is better than the other. Each approach has its own distinctive strengths and weaknesses depending upon why the technique is being used, the nature of the respondents and the nature of researchers and decision-makers who are to use the data.

Viewing laboratory
A room where a focus group may be conducted and simultaneously observed, usually by using a two-way mirror.

For example, in the USA, observing focus groups through the use of purpose-built focus group rooms or a **viewing laboratory** with a two-way mirror is a *policy*. Typically there will be five to eight observers, sometimes more, with agency and marketing company matching level for level to maintain a balance of power. American researchers and clients defend the value of observers. People with different perspectives can listen in a way that a moderator cannot, since the moderator is often 'dipping in for the moment' on one specific issue while the clients 'have the brand in their bones'. Probably the most deeply felt reason for being there is just being there; there is no substitute for the touchy-feely benefit of experiencing the consumer first-hand. More than ever, marketing people are isolated in their small worlds (probably in atypical New York or Chicago), making assumptions about their users. Seeing them is a reality check. Hearing tonality, watching body language, observing the consumer interact with the product, enriches the learning process. Sometimes brand people will get rejuvenated and creatives will get inspired. [18]

e x a m p l e

A client's view of observing the focus group through the two-way mirror[19]

'It's a strange experience, sitting on the other side of a two-way mirror watching your product being analysed in a focus group. It's like being locked out of your house and peering through the windows as burglars rifle through your drawers and take the piss out of your record collection.'

'You want to bang on the glass and plead with them to stop, but then they say something nice about your product and you want to grab the person next to you and shout "I thought of that!".'

'You tell yourself not to take the criticism too personally, to be professional, but it's hard when you've had sleepless nights agonising over minute details only to see them ignored or ripped apart in the space of half an hour.'

'It's great to get raw feedback like this, but remember just how raw it is. One group can say completely different things to another and some may just like the sport of making the suits behind the mirror suffer.' ■

At face value, the above description of the use of viewing labs seems to show a feature of US practice that has many benefits, but consider the following problems:

■ *Incomplete data.* Clients may not have seen all the groups, leaving gaps in the total picture. Observers admit that they focus on the most lively and self-serving points; that they identify positively and negatively with certain responses and filter accordingly; that they are attentive to ideas that affirm their positions and dismissive of contrary viewpoints.

■ *Instant analysis.* Rather than waiting for a moderator's report or even the conclusions of the group, observers often jump to conclusions on little evidence or, worse still, stop listening once they have an 'impression' of the results.

▶

- *Moderator reflection*. Often the moderator gets trapped into debriefing immediately after the group and taking positions that are hard to retreat from but which might be very different after a thoughtful review.
- *Effect on respondents*. There is little research on the effect of the two-way mirrors on respondents. The setting is not exactly conducive to natural expression and respondents may alter their responses for effect, stay silent or be self-conscious.

Supporters of the European approach to focus groups may console themselves that they can overcome these drawbacks. They may argue that the context in which they run groups is more conducive to relaxed respondents, that moderators are not under such pressure to produce 'instant analysis' and that they can take more time to reflect and interpret the data generated. But in Europe, and especially in Britain (which has had the biggest tradition of running focus groups in people's homes), marketing researchers are embracing the use of the viewing lab. The following example illustrates how the viewing lab is increasingly being used in Britain.

example

Through the looking glass[20]

Which is the better marketing research environment: a member of the public's front room or a viewing lab? Traditionally, in Britain, the former was the unchallenged leader when it came to new product development and advertising campaign research. However, during the 1990s, Britain has taken a leap into viewing-lab territory.

Lucy Bannister, a director of qualitative research firm Davies Riley Smith Maclay, sees Britain catching up with the rest of the world.

> 'Clients in Britain have always thought of in-home research first because that is the way things developed historically. But I cannot think of any other country where this is the case, especially in Europe or in the US. Viewing labs have always been the norm outside Britain and marketers abroad think it is bizarre that we have this strange bias.' ■

Some focus group participants quickly learn to behave quite naturally in front of two-way mirrors.

The conclusion from this example is that marketing researchers should not take a dogmatic position, believing that their approach is the most 'correct'. Focus groups vary enormously in how they are planned, administered and analysed, from the very structured interview in a studio bristling with technology, to a passionate dialogue conducted globally over the Internet, to a riotous exchange on a tropical beach. The marketing researcher should be open-minded to learn from focus group practices in different countries and be able to critically evaluate why they have been administered in that way.

Ethics in marketing research

A number of ethical issues related to running focus groups have emerged in this chapter. The focus group can be a direct qualitative research technique where the purpose of the discussion is made clear to respondents before it starts. However, the focus group can incorporate observational and projective techniques that introduce elements of indirectness, i.e. elements or all of the purpose of the discussion being hidden from respondents. This is where the marketing researcher faces a real ethical dilemma. If they fully reveal the purpose of their study, would this put off potential respondents whose views are important to the success of the study? Even if they do manage to recruit potential respondents, the researcher has to consider how they may feel when the real purpose of the study becomes apparent through the nature of the discussion or by the moderator revealing all at the end. The nature of the dilemma faced by marketing researchers is that by revealing too much at the start of the study they may compromise the quality of their discussion. A full revelation may not be conducive to respondents reflecting, making new connections and expressing themselves about issues that may be deeply held, difficult to conceive and express.

Researchers should take all reasonable precautions to ensure that respondents are in no way adversely affected or embarrassed as a result of the focus group. What is meant by an adverse effect and embarrassment will largely depend upon the issues being explored and how they are perceived by target respondents. Some respondents may find personal hygiene issues very embarrassing to talk about in a group scenario, but not financial issues. Other individuals may be very frank about their sexual behaviour, while others would be shocked at the notion of talking about sex with a group of strangers. The researcher must get to know how the issues they wish to explore are perceived by their target respondents by examining literature, secondary data and the use of experimental focus groups.

Another major ethical problem that is research issue and respondent specific is the use of alcoholic drinks during focus groups. For certain groups of respondents, relaxing and socialising in a comfortable context, drinking wine or beer is very natural. Researchers with experience of running many focus groups would argue that serving alcoholic drinks can help to reduce tension in certain respondents and give them the confidence to express their particular viewpoint. Other researchers would argue that this is unethical practice, that in effect the researcher is 'drugging' the respondents. Whatever the researcher decides is right for the type of respondents, the issues they are questioning them about and the context in which the discussion takes place, there are practical problems involved with serving alcohol. Controlling the flow of alcohol and how much is given to certain respondents may take attention away from the discussion. If control is not exerted, particular respondents may get out of hand and disrupt or even destroy the discussion. Researchers could be

▶

accused of not taking reasonable precautions to ensure that respondents are in no way adversely affected or embarrassed as a result of the focus group, should the use of alcohol be abused by certain respondents.

Finally, in no circumstances should researchers use audio or video recording or two-way mirrors in focus groups without gaining the consent of respondents. It must be made clear to respondents that recording or observation equipment is to be used and why it needs to be used. Respondents must then be free to decline any offer to take part in a discussion.

Internet and computer applications

The development of the Internet, the increasing numbers of individuals with access and who are comfortable using the Internet, has presented many opportunities to run focus groups online. For online focus groups, virtual facilities can be used, providing the same facilities as real-life facilities including 'rooms' such as a reception room, discussion room and client backroom. Online focus groups can be conducted in two ways: 'real-time' or 'non-real-time'.[21]

■ *Real-time.* In this method, all respondents are online at the same time. The transmission of messages is immediate or as close to immediate as can be. As one respondent types in their message, it is transmitted to the group as a whole. Other respondents can read the message and can reply as soon as they receive it. Real-time focus groups can be highly interactive with a fast 'passionate' exchange of views. This is the big advantage of real-time groups, that the passion for a subject and the nature of rapid exchange can develop a very creative atmosphere. Respondents do not have to wait for others to comment in order to send further messages, so as they have been stimulated to think in particular ways, their thoughts and ideas can flow straight out on to the screen. The drawback is that the respondent with the most proficiency at typing has the power to dominate the discussion. Compared with face-to-face focus groups where the moderator and audio and video recordings can track respondent statements and the responses stimulated, this cannot be done online in real time. The distinction of replying and sending becomes a blur as respondents do not take their turn as they would face to face. This means that much of the structure of conversation is lost.

■ *Non-real-time.* In this method there is no requirement for respondents to be online at the same time. This method is analogous to the use of email. Unlike email, though, this form of focus group is conducted using a 'conference site' as opposed to individual email addresses. Respondents can send their views about a particular issue to a conference folder. Messages can be read in real time but this does not usually happen. The reality is that all responses are archived in the folder and can be opened and responded to by other respondents. This method of running focus groups can overcome difficulties in running focus groups in different time zones and therefore is an excellent tool in conducting international research. Respondents with weak typing skills do not lose their voice or feel intimidated. The major benefit of this approach is that akin to face-to-face focus groups, respondents can 'sit back' and reflect upon how they really feel about an issue. Even better than in face-to-face scenarios, they can take time to express how they really feel about an issue.

Online focus groups have a number of logistical requirements. Depending upon how the online focus group is to be conducted, respondents may have to gain access to proprietary conferencing software. Real-time and non-real-time focus groups can operate from particular Websites, requiring potential respondents to have a Web browser and to know the Web address from which the group will be run.

In summary, the key benefits of running focus groups online include the following:[22]

- More potential respondents can be recruited through the growing use of the Internet, and the growing ease of conducting discussions online.
- Respondents can be made to feel that they have the ability to contribute; their confidence can be quickly built up.
- Conflicts in face-to-face focus groups that may stem from respondents taking a dislike to other respondents from their physical appearance can be avoided.
- A great breadth of information may be collected, through the types of respondent that can be recruited and the geographic spread of respondents.
- The practical difficulties of getting individuals together at the same time in the same location can be overcome.
- The nature of a discussion location that is 'comfortable' to the respondent is largely overcome by each respondent setting the conditions that they feel comfortable in.

Summary

In direct qualitative methods, respondents are able to discern the true purpose of the research, whereas indirect methods disguise the purpose of the research to some extent. Focus groups can be a completely direct qualitative method, though by the use of observational and projective techniques elements of indirectness can be introduced. In deciding how 'direct' a focus group should be, researchers have to face ethical issues and questions related to the richness of data that they can draw from respondents. Focus group discussions are the most widely used qualitative research technique.

Focus groups are conducted by a moderator in a relaxed and informal manner with a small group of respondents. The moderator leads and develops the discussion. In a focus group, respondents can portray their feelings and behaviour, using their own language and logic. The value of the technique lies in the unexpected findings that emerge when respondents are allowed to say what they really feel. An experimental focus group should always be run at the start of any series of discussions. The researcher needs to understand how comfortable target respondents feel in the chosen location for discussion, how the opening question works, the topic guide, the probes, the mix of respondents and the refreshments, including any alcoholic drinks. In a nutshell, one could argue that one should continue running focus groups until nothing new is learned from target respondents. This perspective oversimplifies how diverse and spread out different respondents may be and how well they mix together in a group situation. The time and budget available will also mean that a cut-off point has to be drawn and the strongest analysis and interpretation has to be made at that point. Development in the use of the Internet means that group discussions can take place away from the traditional 'viewing' or meeting room. The maxim of respondents being relaxed and comfortable with the context and means of communicating still applies.

There are a variety of different styles of running focus groups. The variations and adaptations are used to draw the best out of particular types of respondent, tackling particular types of issue. The most significant development in the variation of technique has been afforded by developments in Internet technology and take-up. Respondents can be targeted from all over the world, discussing issues online in 'real-time' or 'non-real-time'. Another factor that affects the style of running focus groups lies in the underlying philosophy for conducting research. US and European styles can be broadly encapsulated as being respectively 'highly structured' or 'highly spontaneous'. There is no one absolute correct method to administer a focus group; the researcher should understand the factors that will make the technique work for their particular research problem and the type of respondents they have to work with. A key element of ensuring that the focus group works well lies in the ethical issues of how much is revealed to respondents before they get involved and during the discussion. Getting respondents to relax and be open may involve the use of alcoholic drinks that for many researchers creates no problems but for some creates ethical and practical problems.

Questions

1 Why may marketing researchers not wish to fully reveal the purpose of a focus group discussion with respondents before it starts?

2 What are the key benefits and drawbacks of conducting focus group discussions?

3 What are the difficulties in conducting focus groups with managers or professionals?

4 What determines the questions, issues and probes used in a focus group?

5 Evaluate the purpose of running an experimental focus group discussion.

6 What does a 'comfortable setting' mean in the context of running a focus group?

7 To what extent can a moderator achieve an 'objective detachment' from a focus group discussion?

8 Why is the focus group moderator so important to the success of a focus group discussion?

9 What are the relative advantages and disadvantages of being able to covertly observe a focus group discussion?

10 What can the researcher do to make potential respondents want to take part in a focus group?

11 What determines the number of focus groups that should be undertaken in any research project?

12 Describe the purpose and benefits of using stimulus material in a focus group.

13 What is the difference between a dual moderator and a duelling moderator group?

14 Describe the opportunities and difficulties that may occur if alcoholic drinks are served during focus group discussions.

15 Evaluate the benefits and limitations of conducting focus group discussions on the Internet.

Notes

1 Clarke, B. and Saunders, M., 'There's trouble in classes for kids who wear glasses', *ResearchPlus* (November 1997), 7.

2 Garee, M.L. and Schori, T.R., 'Focus groups illuminate quantitative research', *Marketing News* 31(12) (9 June 1997), H25.

3 Drayton, J. and Tynan, C., 'Conducting focus groups – a guide for first-time users', *Marketing Intelligence and Planning* 6(1) (1988), 5–9.

4 The group size of 6 to 12 is based on rules of thumb. For more discussion, see Dachler, H.P., 'Qualitative methods in organization research', *Organizational Studies* 18(4) (1997), 709–24.

5 Mazella, G.F., 'Show-and-tell focus groups reveal core bloomer values', *Marketing News* 31(12) (9 June 1997), H8; 'Focus groups: consumers speak out', *Frozen Food Age* 43(11) (June 1995) 709–24; Nelson, J.E. and Frontczak, N., 'How acquaintanceship and analyst can influence focus group results', *Journal of Advertising* 17 (1988), 41–8.

6 Kahn, H., 'A professional opinion', *American Demographics* (Tools Supplement) (October 1996), 14–19.

7 Gordon, W., 'New life for group discussions', *ResearchPlus* (July 1993), 1.

8 Greenbaum, T.L., *The Handbook for Focus Group Research* (Newbury Park, CA: Sage, 1997).

9 Ellis, R., 'So how do you upstage a Ferrari owner?', *ResearchPlus* (November 1994), 10.

10 Tayler, D., 'An MR boost helps Disneyland Paris take off into profit', *ResearchPlus* (June 1997), 6.

11 Croft, M., 'Art of the matter', *Marketing Week* (9 October 1997), 71.

12 *Ibid.*

13 Katcher, B.L., 'Getting answers from the focus group', *Folio: The Magazine for Magazine Management* (Special Sourcebook Issue for 1997 Supplement) 25(18) (1997), 222; and an adaptation from Chase, D.A., 'The intensive group interviewing in marketing', *MRA Viewpoints* (1973).

14 Silverstein, M., 'Two-way focus groups can provide startling information', *Marketing News* (4 January 1988), 31.

15 Gallupe, R.B. and Cooper, W.H., 'Brainstorming electronically', *Sloan Management Review* 35(1) (Fall 1993), 27.

16 Savage, M., 'Soft focus', *Research* (September 1999), 32–3.

17 Note that the expression 'US' or 'European focus group' does not mean that such methods are exclusively used in these continents. The term is used to show that there is a greater propensity for a particular approach to running a focus group in a particular manner in each of these continents.

18 Sonet, T., 'See the USA, through the looking glass', *ResearchPlus* (June 1994), 6.

19 Dwek, R., 'Through the looking glass', *Marketing* (11 September 1997), 37.

20 *Ibid.*

21 Mann, C. and Stewart, F., *Internet Communication and Qualitative Research: A Handbook for Researching Online* (London: Sage, 2000), 101–2.

22 Sweeney, J.C., Soutar, G.N., Hausknecht, D.R., Dallin, R.F. and Johnson, L.W., 'Collection of information from groups: a comparison of two methods', *Journal of the Market Research Society* 39(2) (April 1997), 397.

Qualitative research: depth interviewing and projective techniques

Objectives

After reading this chapter, you should be able to:

1 understand why the depth interview is defined as a direct qualitative research technique and observation and projective techniques are defined as indirect techniques;

2 describe depth interview techniques in detail, citing their advantages, disadvantages and applications;

3 explain how theory may be used to create structure to questioning and analysis in depth interviewing by reference to the laddering technique and the repertory grid technique;

4 describe projective techniques in detail and compare association, completion, construction and expressive techniques;

5 understand the language problems that should be considered by international qualitative researchers;

6 understand the ethical dilemmas faced by qualitative research practitioners.

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

With no social pressure to conform to group responses, respondents can be questioned in-depth in a context that allows them to really express how they feel.

Overview

Having discussed qualitative research in terms of its nature and approach, and evaluated focus groups as the main marketing research technique, we now move on to describe and evaluate other qualitative techniques.

We start by describing and evaluating the procedure of depth interviewing. The role of the depth interviewer is described and the advantages and disadvantages of the technique are summarised. In the process of conducting depth interviews, the techniques of 'laddering' and 'repertory grid' can be applied to help to structure the elicitation and analysis process. Laddering and repertory grid techniques will be described and illustrated. The indirect qualitative association, completion, construction and expressive projective techniques are described and illustrated. The applications of these techniques are detailed, followed by an evaluation of their advantages and disadvantages. This is developed into an overall summary of the relative strengths and weaknesses of qualitative techniques under the headings of 'focus groups', 'depth interviews', 'projective techniques' and 'ethnographic techniques'. The GlobalCash Project is used to illustrate why depth interviewing may be used in international marketing research, demonstrating some of the cultural issues of applying the technique. We summarise the relative gains that can be administered on a one-to-one basis, with an emphasis on techniques.

The following example illustrates a subject that for some individuals is difficult to discuss in a group: personal hygiene issues and even their fantasies when bathing. The subject requires a qualitative technique to generate data of a kind that will support creative advertising decisions. The subject matter, however, is ideal for the application of depth interviews and projective techniques where time can be spent to build trust and rapport between interviewer and respondent. Depth interviews allow respondents talking about bathing the time and space to express ideas about events that normally they just get on with and do not have to justify to anyone!

example

Soaps look for a fresh way to work consumers into a lather[1]

In studies of bath soaps, respondents invariably say that a good soap makes them feel 'clean and fresh'. However, they often have difficulty explaining what 'clean and fresh' means. Copywriters trying to find a new way to talk about freshness in their advertising do not find such results supportive or relevant to the decisions they have to take. Hence, respondents have been probed on a one-to-one basis through depth interviews, about all the things 'clean and fresh' meant to them: the times they felt this way, their mental pictures, the moods and feelings connected with it, what music and colours come to mind, and even what fantasies it evoked.

Escape from ordinary life was one of the main themes that emerged from the depth interviews – getting away from the cramped rushed city, to the free, relaxed, unhindered life, surrounded by nature, in the country. The words and images sparked by this theme offered new ideas for creative advertising ideas. ■

Depth interviews

Procedure

Depth interview
An unstructured, direct, personal interview in which a single respondent is probed by an experienced interviewer to uncover underlying motivations, beliefs, attitudes and feelings on a topic.

Like focus groups, depth interviews are an unstructured and direct way of obtaining information but, unlike focus groups, depth interviews are conducted on a one-on-one basis. A **depth interview** is an unstructured, direct, personal interview in which a single respondent is probed by an experienced interviewer to uncover underlying motivations, beliefs, attitudes and feelings on a topic.[2]

179

A depth interview may take from 30 minutes to over an hour. It may occur on a one-off basis or it may unfold over a number of meetings between an interviewer and a respondent. Once the interviewer has gained access to a potential respondent, the interviewer should begin by explaining the purpose of the interview, showing what the respondent will get out of taking part in the interview and explaining what the process will be like. To illustrate the technique in the context of the GlobalCash Project, beyond the introduction the interviewer may ask a Finance Director a general introductory question such as 'Has your main bank helped you to prepare for the introduction of the euro?' The interviewer then encourages the subject to talk freely about his or her attitudes towards different banks' offerings. After asking the initial question, the interviewer uses an unstructured format, guided by a topic guide to remind them of important subject areas to cover. The subsequent direction of the interview is determined by the respondent's initial reply, the interviewer's probes for elaboration, and the respondent's answers.

As with the focus group topic guide and the moderator managing that guide, spontaneity ensures that the process is creative and meaningful to the respondent. Suppose that the respondent replies to the initial question by saying 'Bank X, our bank, has not been really helpful in developing our plans.' The interviewer might then pose a question such as 'Why do you think they have not been helpful?' If the answer is not very revealing, e.g. 'They are not customer orientated, never have been', the interviewer may ask a probing question, such as 'What would you see as being a customer orientated bank?' This question could open up a whole series of issues such as 'trust', 'relationship development' or 'technical support', to name a few. Such an exchange of questions and answers could emerge from a heading of *Bank awareness of their customers' needs* on their topic guide. The interviewer will keep an eye on their topic guide to ensure they tackle all the issues they feel to be important, but the specific wording of the questions and the order in which they are asked is influenced by the respondent's replies.

Probing is of critical importance in obtaining meaningful responses and uncovering hidden issues. Probing can be done by asking general questions such as 'Why do you say that?', 'That's interesting, can you tell me more?' or 'Would you like to add anything else?'[3] Probing can also be more specific, an example in the above scenario being 'At what point do you feel your relationship with Bank X started to deteriorate?' One of the main success factors of specific probing is that the researcher understands something about the nature of the subject they are researching. This means they appreciate the significance of particular revelations, understand the language (even technical language and jargon in certain areas, like banking) and have a credibility with the respondent that encourages them to open up with the interviewer.

The interviewer must be alert to the issues that they wish to go through but also the issues that the respondent is willing to talk about, and must listen carefully to and observe which issues fire enthusiasm in the respondent. The questions and probes they put to respondents should follow the interest and logic of the respondent, making them feel motivated to respond in a manner that suits them. As with a focus group discussion, the respondent should feel comfortable and relaxed, which could mean holding the interview in their office, their home, a bar, a sports club – any context in which the respondent will feel comfortable and more willing to be reflective, honest and open. Answering in a manner that suits the respondent helps to make the interview more comfortable and relaxed. For a great amount of business research, the depth interview is the best way to gain access and to talk to managers. Much of the interviewing takes place in their office at a time that is convenient to them. Researchers can also observe characteristics of the manager in his or her office environment that can be of help in their analyses. Examples of this could be levels of

formality in the workplace, reports and books that the manager has around them for reference, the manager's use of IT equipment, or the tidiness of the workplace. These observations would be entirely based upon the purpose of the study, but the context of the office can be of help to the manager and the researcher. In order to make the above process work, the interviewer should:

1 Do their utmost to develop an empathy with the respondent.
2 Make sure the respondent is relaxed and comfortable.
3 Be personable to encourage and motivate respondents.
4 Note issues that interest the respondent and develop questions around these issues.
5 Not be happy to accept brief 'yes' or 'no' answers.
6 Note where respondents have not explained clearly enough issues that need probing.

In marketing research that focuses upon managers or professionals as illustrated above, the context of the depth interview helps to set the frame of mind of the respondent. The context should also help the respondent and interviewer to relax, engendering an atmosphere to explore and develop issues that they feel to be relevant. The following arguments help to focus on the issues faced by the interviewer coping with managers and professionals, trying to find the right context to allow these respondents to express how they really feel.[4] The depth interview helps to overcome:

1 *Hectic schedules.* The best respondents tend also to be the busiest and most successful people. They can make time for an interview, but are rarely able to spare the much greater time needed for them to come to a group discussion at some location away from their office. So groups exclude the best respondents.
2 *Heterogeneity.* Whereas mothers evaluating nappy ads, or beer drinkers the latest lager, have only their personal preferences to consider, it is very different for an executive evaluating copiers, airline advertisements or computer software. This is because their reactions are complicated by the type of job they do and who they work for. The group discussion is dependent on the group's composition being fairly homogeneous; the job backgrounds of business people make them too varied to be entirely comfortable in a group.
3 *Live context.* A lot of information comes from seeing the respondent at his or her desk, which is missed in a group discussion. Work schedules pinned to the wall, the working atmosphere, the freebies from suppliers on the desk, the way coffee is served, help fill out the picture.
4 *Interviewer reflection.* Groups do not allow the researcher enough thinking time. Two groups, each taking an hour and a half over successive evenings, do not even begin to compare with two or three full days of non-stop interviewing. Individual interviews give much more scope for experimentation. If one way does not work, it's only one respondent, not a whole group, that is affected.

Another major application of depth interviews where the context of the interview plays a major part is in the interviewing of children. Researchers into children and teenagers spend considerable time working out the best research approach. Debates proliferate on the most appropriate research designs and interviewing techniques: depth interviews versus group, mini groups versus standard groups, friendship pairs versus stranger groups, association projective techniques versus expressive and so on. These debates make researchers focus on the implications of how research is conducted and which technique provides the best results. One vital element is often overlooked, and that is the issue of context: in other words, the need to ensure that the situation in which children are interviewed is relevant to the research needs.

Like adults, children have multifaceted personalities. The same teenager can be sullen at home, obstructive in the classroom, but the life and soul of the peer group.

The essential difference between children and adults, however, is the extent to which different aspects of the persona can be accessed on one occasion and in one situation, the research setting. Adults have insight into the different roles and behaviour, which they adopt in different contexts, and can project other aspects of themselves which they bring into the research situation. Children and young teenagers, on the other hand, react to the moment and thus project only one aspect of themselves. They lack the maturity, experience and self-knowledge to draw on other parts of themselves and therefore find it almost impossible to know, let alone admit or explore, how they might behave in another circumstance. [5]

The above evaluation of the importance of context also shows that question formulation, the very heart of marketing research, can be more important in interviewing children than when dealing with adults. A straightforward 'What do you think about X?' may throw children into confusion. A formulation such as 'If you were to call your friend about this, what would you tell them?' would be more likely to produce illuminating responses.

Friendship pair
A technique used to interview children as two friends or classmates together.

➤➤➤
See Professional Perspective 16.

An effective technique when working with children is the use of **friendship pairs** – interviewing two friends or classmates together. This helps to cut out lying because children are not alone with strangers and because if one tells a lie, the other tells on them. The ingrained honesty of children arguably makes them easier to research than adults, who of course are far more accomplished exponents of deception. Go to the Companion Website and read Professional Perspective 16 'Out of the mouths of babes' by Marsha Hemingway. Marsha describes many of the difficulties of researching children and many creative solutions to make interviewing fun and informative.

Advantages and challenges of depth interviews

Depth interviews have the following advantages. They can:

1 Uncover greater depth of insights than focus groups. This can happen through concentrating and developing an issue with the individual. In the group scenario, interesting and knowledgeable individuals cannot be solely concentrated upon.

Friendship pairs: helps to cut out lying, because if one tells a lie, the other tells on them.

2 Attribute the responses directly to the respondent, unlike focus groups where it is often difficult to determine which respondent made a particular response.

3 Result in a free exchange of information that may not be possible in focus groups because there is no social pressure to conform to group response. This makes them ideally suited to sensitive issues, especially commercially sensitive issues.

4 Be easier to arrange than the focus group as there are not so many individuals to coordinate and the interviewer can travel to the respondent.

The following are not necessarily disadvantages, but really the challenges that researchers face when using this very valuable technique:

1 The lack of structure makes the results susceptible to the interviewer's influence, and the quality and completeness of the results depend heavily on the interviewer's skills. As with all qualitative techniques, the interviewer needs to develop an awareness of the factors that make them 'see' in a particular way.

2 The length of the interview, combined with high costs, means that the number of depth interviews in a project tends to be few. If few depth interviews can be managed, the researcher should focus upon the quality of the whole research experience. 'Quality' in this context means the qualities that the respondent possesses in terms of richness of experience and how relevant their experiences are to the study; the quality of drawing out and getting respondents to express themselves clearly and honestly; and the quality of analysis in terms of interpretation of individual respondents and individual issues evaluated across all the interviews conducted.

3 The data obtained can be difficult to analyse and interpret. Many responses may not be taken at face value; there can be many hidden messages and interpretations in how respondents express themselves. The researcher needs a strong theoretical awareness to make sense of the data or the technical means to develop theory if using a grounded theory approach. As well as the transcripts of the interview, additional observations add to the richness and multifaceted analyses and potential interpretations.

The following technique illustrates how a theoretical awareness can help to develop structure, elicit or draw more out of respondents and help to make sense of the data they generate. It is an illustration of how the three challenges outlined above can be overcome.

The laddering technique

The depth interview can be driven by topic guide, made up of a just a few topics covering a very broad range of issues. From these few topics, the nature of questions, the order of questions and the nature of probes can be driven by the interviewer's perception of what will draw the best out of respondents. Alternatively, a depth interview can be semi-structured where parts of the interview use consistent and highly structured questions, with set response categories, interspersed with open-ended questions involving probes that again suit the nature of the respondent. There are other variations on the technique that can help the interviewer and respondent to apply useful structure to issues that can be very 'messy' and unstructured. One of the most popular techniques to apply structure is called laddering.

Laddering
A technique for conducting depth interviews in which a line of questioning proceeds from product characteristics to user characteristics.

The laddering technique is made up of a linking or ladder of elements that represent the link between products and the consumer's perception process. It enables an understanding of how consumers translate product attributes, through personal meanings associated with them.[6] Theories of consumer behaviour act as a foundation to this approach, based on the contention that consumption acts produce consequences for the consumer. Consumers learn to associate these consequences to

Laddering techniques should release respondents from fixed attitudes, perceptions and values.
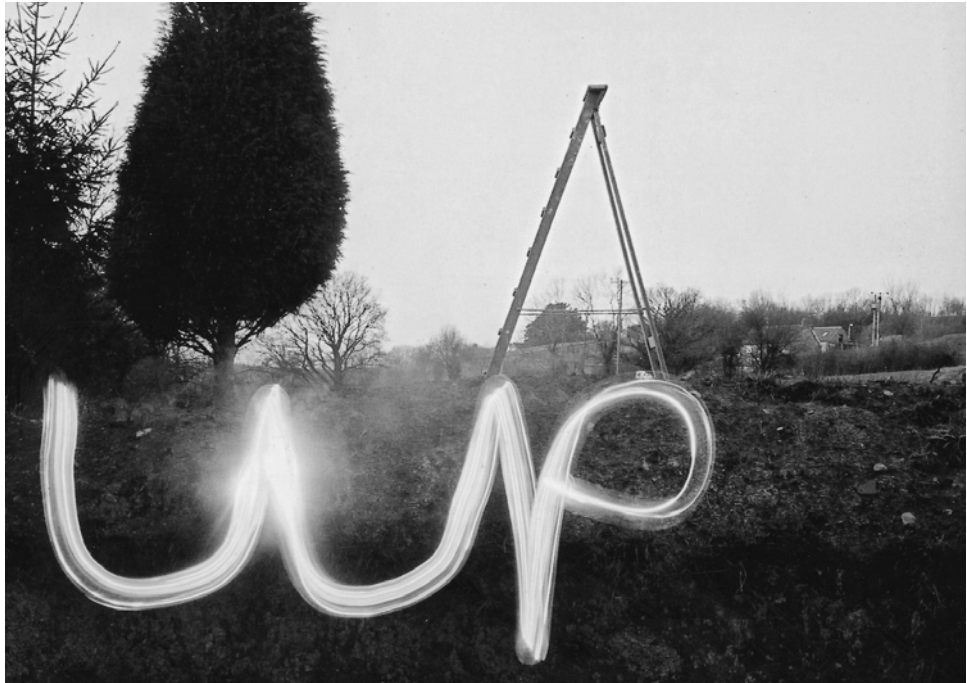
specific product attributes. These associations are reinforced through consumers' buying behaviour and, as a result, they learn to choose products that have certain attributes in order to obtain the consequences they desire. There are two features of this theory that the researcher using laddering technique will focus upon:

1 *Motivation.* The laddering technique focuses upon generating insight to the motives behind consumption of certain products. It represents a more contemporary approach to classical motivation research.[7] The laddering technique aims to stimulate respondents to reflect upon their buying behaviour in a way unconnected from their usual behaviour.

2 *Cognitive structure.* This feature is also a development of theory, in this case Means-End-Chain (MEC), developed by Guttman.[8] MEC starts from a point of consumer motivation. It contends that motivation leading towards a form of behaviour is derived from the connection between tangible product attributes and more abstract cognitive structures that involve the physical and emotional consequences derived from these attributes. At a more abstract level, the consequences lead to values held by the consumer.

The laddering technique is therefore designed to identify and follow the chain of:

$$\text{Attributes} > \text{Consequences} > \text{Values} \ (A - C - V)$$

The depth interview using the laddering technique is based on comparisons of the consumer's choice alternatives. These can include, for example, different products used for the same purpose, such as an electric toothbrush and a conventional toothbrush, and/or varieties in a product line such as full-fat and skimmed milk, and/or product brands such as Heineken and Amstel beer, and/or kinds of packaging such as wine in bottles and in cartons. Other elements that can affect consumer choices can be added to the above list. In making the comparisons, the interviewer poses a question that hopes to encourage the respondent to put aside their established rationale and reflect upon their consumption behaviour, in ways that they would not normally do. In other words, questions are posed that make respondents think about consump-

tion from other points of view to try to release respondents from fixed attitudes, perceptions and values. The interviewer's role is to build up a rapport with respondents and get them to relax and feel comfortable to respond with whatever comes to mind. The interview revolves around three basic questions based on the A – C – V chain. The questions posed would be:

1 *Values*: How important is this for you? (e.g. health)
2 *Consequences*: What does this difference mean? (e.g. not fattening)
3 *Attributes*: What is different about these alternatives? (e.g. low calories)

From the attribute of 'low calories' in comparing different product varieties in a product line such as full-fat and skimmed milk, a further and negative consequence of 'watery taste' could be elicited, followed by the values of 'unnatural'. The interviewer aims to build a great repertoire of chains through these levels spontaneously, with the interviewer choosing the right comparisons to draw out the next link in the chain. The resulting responses are analysed to establish possible categories of attributes, consequences and values. The categories are then organised according to the qualitative comments made between the relationship of an attribute, its consequence and value. The result is an association matrix that graphically displays the categories and the connections that have emerged.

Laddering requires an interviewer with experience of depth interviewing, with a realisation of what will relax a respondent and get them in a frame of mind to 'play' with the associations sought. The interviewer needs to appreciate the theoretical foundations to the technique,[9] not only to help them determine what they should aim to elicit from respondents but also to help generate sound and meaningful analyses.

The repertory grid technique

Another widely used technique that applies structure to qualitative depth interviewing is the repertory grid technique (RGT). This technique was originally developed by George Kelly in 1955[10] to explore the meanings that people attach to the world around them, that they find particularly hard to articulate. As with the laddering technique, there is a theoretical underpinning, personal construct psychology.[11] Given the debate around this theory, there are a number of variations of the technique and arguments about whether it should be implemented and analysed in a positivist or interpretivist manner,[12] but in essence the stages involved in the repertory grid technique are:

1 Element selection
2 Construct elicitation
3 Element comparisons
4 Data analysis

Element selection. The elements selected will depend upon the nature of consumer behaviour that the interviewer wishes to examine. In a study that wished to understand '*broad-based patterns of consumer behaviour*'[13] the elements chosen included 30 generic products and services such as newspapers, holidays, chocolate bars, eggs, alcoholic drinks, shoes, toothpaste, savings accounts, restaurant meals and shampoo. In another study that wished to understand '*the process of effective new product development*'[14] the elements chosen included 30 successful new products and services such as Slim Fast, Pull Ups, Loed Tea, Ultraglide, Baby Jogger, Gourmet Coffee, Zantac, Paragliders, MTV and Carmen Soft. These elements should be chosen by the respondents, not just chosen and presented by the interviewer. There should be some homogeneity in the elements in the sense that the respondent sees the elements relating to and representative of the behaviour being studied.

Construct elicitation. Having selected elements that the respondent believes to encapsulate the behaviour being studied, the interviewer now seeks to understand what connects them together. The first stage of this involves the interviewer selecting three of the chosen elements at random and then presenting to the respondent small cards with a summary of these elements. The respondent is then asked to describe how they see two of the three to be alike and how the third may be different. The researcher selects different 'triads' to the point where all the elements have been contrasted or the respondent cannot describe further 'similarities' or 'differences'. Construct elicitation therefore draws out the respondents' perspective of the important features that encapsulate a particular form of consumer behaviour.

Element comparisons. The constructs elicited from respondents are now turned into bipolar descriptions in a manner similar to the semantic differential scale described and illustrated in Chapter 12, page 305. For example, suppose that in a study of '*the process of effective new product development*', three elements were compared: MTV, Gourmet Coffee and Paragliders. The respondent may say that MTV and Gourmet Coffee are alike in that there are many similar competitive products that they can be compared with, whereas the Paraglider is different in that there are few competitive 'personal' means to fly long distances. Whether this is factually correct or whether the interviewer agrees with this is immaterial; it is how the respondent sees it that is important. Such a belief could be turned into a bipolar scale as illustrated below:

No other products like this in the market	— (1)	— (2)	— (3)	— (4)	— (5)	— (6)	— (7)	Lots of competitive products

Other scales would be constructed to cover the ways that respondents compared the chosen elements, i.e. the constructs elicited would be turned into a series of bipolar scales. Respondents would then be expected to evaluate all the original elements using these scales. If they felt that Gourmet Coffee has many competitors they could tick a line at the (6) or (7) point. So, if there were 30 elements that encapsulate the behaviour under study and 20 constructs elicited to evaluate these elements, a total of 30×20 or 600 ratings would need to be performed.

Data analysis. A grid can be assembled for each respondent to represent in the above example, the 30×20 ratings. Again, an illustration of this comes from a small extract of the '*the process of effective new product development*' study. The following represents an extract of the total response from one respondent.

Construct	*MTV*	*Gourmet Coffee*	*Paraglider*
Market newness	6	6	2
Company newness	3	7	1
Technology newness	2	4	6

With a number of these completed grids, factor analysis (Chapter 22) can be performed to discover the important underlying factors or dimensions that encapsulate a particular form of behaviour.[15] The analysis can continue by performing cluster analysis (Chapter 23) to explore patterns of similarity or dissimilarity in different types of respondent, i.e. to discover and describe groups of respondents who may view particular forms of behaviour in similar ways.

The above process can take a great deal of time to select the elements, to elicit the means to compare the elements and to evaluate them. It requires an amount of rapport

between the interviewer and the respondent and patience in working through the stages. All of the stages may not be completed in one interview. It may require the interviewer to return to the respondent to develop the next stage, especially in transferring the elicited elements to bipolar descriptions to allow all the elements to be evaluated.

Applications of depth interviews

Applying depth interviews, with high or low levels of structure, with a strong theoretical foundation such as the laddering or repertory grid technique, or with the desire to generate grounded theory, presents challenges but also many rewards. There are many marketing decisions that can be made with support from marketing researchers using the broad array of techniques under the heading of 'depth interviews'. The following summarises the applications.[16]

1 Interviews with professional people (e.g. finance directors using banking services).
2 Interviews with children (e.g. attitudes towards a theme park).
3 Detailed probing of the respondent (e.g. new product development for cars).
4 Discussion of confidential, sensitive or embarrassing topics (e.g. personal hygiene issues).
5 Situations where strong social norms exist and where the respondent may be easily swayed by group response (e.g. attitudes of university students towards sports).
6 Detailed understanding of complicated behaviour (e.g. the purchase of fashion or 'high-status' goods).
7 Interviews with competitors, who are unlikely to reveal the information in a group setting (e.g. travel agents' perceptions of airline travel packages).
8 Situations where the product consumption experience is sensory in nature, affecting mood states and emotions (e.g. perfumes, bath soap).

In the application of depth interviews, the researcher can use other techniques to help maintain the interest of respondents, to make the experience more enjoyable for the respondent and themselves alike, and ultimately to draw out the true feelings of respondents. A set of techniques that helps to achieve all this, that have been applied with great success over many years, are the body of indirect techniques called 'projective techniques'.

Projective techniques

Projective technique
An unstructured and indirect form of questioning that encourages respondents to project their underlying motivations, beliefs, attitudes or feelings regarding the issues of concern.

Both focus groups and depth interviews are direct approaches in which the true purpose of the research can be disclosed to the respondents or is otherwise obvious to them. **Projective techniques** are different from these techniques in that they attempt to totally disguise the purpose of the research. A projective technique is an unstructured, indirect form of questioning that encourages respondents to project their underlying motivations, beliefs, attitudes or feelings regarding the issues of concern.[17] In projective techniques, respondents are asked to interpret the behaviour of others rather than to describe their own behaviour. In interpreting the behaviour of others, it is contended that respondents indirectly project their own motivations, beliefs, attitudes or feelings into the situation. Thus, the respondent's attitudes are uncovered by analysing their responses to scenarios that are deliberately unstructured, vague and ambiguous. The more ambiguous the situation, the more respondents project their emotions, needs, motives, attitudes and values, as demonstrated by work in clinical psychology on which projective techniques are based.[18] As in psychology, these techniques are classified as association, completion, construction and expressive. Each of these classifications is discussed below.[19]

Association techniques

In **association techniques**, an individual is presented with a stimulus and asked to respond with the first thing that comes to mind. Word association is the best known of these techniques. In **word association**, respondents are presented with a list of words, one at a time, and are asked to respond to each with the first word that comes to mind. The words of interest, called test words, are interspersed throughout the list, which also contains some neutral, or filler, words to disguise the purpose of the study. For example, in the GlobalCash study, individual banks may be examined with test words such as location, speed, error, quality and price. The subject's response to each word is recorded verbatim and responses are timed so that respondents who hesitate or reason out (defined as taking longer than three seconds to reply) can be identified. The interviewer, not the respondent, records the responses.

The underlying assumption of this technique is that association allows respondents to reveal their inner feelings about the topic of interest. Responses are analysed by calculating:

1 The frequency with which any word is given as a response.
2 The amount of time that elapses before a response is given.
3 The number of respondents who do not respond at all to a test word within a reasonable period.

Those who do not respond at all are judged to have an emotional involvement so high that it blocks a response. It is often possible to classify the associations as favourable, unfavourable or neutral. An individual's pattern of responses and the details of the response are used to determine the person's underlying attitudes or feelings on the topic of interest, as shown in the following example.

e x a m p l e

Dealing with dirt

Word association was used to study women's attitudes towards detergents. Below is a list of stimulus words used and the responses of two women of similar age and household status. The sets of responses are quite different, suggesting that the women differ in personality and in their attitudes towards housekeeping. Ms M's associations suggest that she is resigned to dirt. She sees dirt as inevitable and does not do much about it. She does not do hard cleaning, nor does she get much pleasure from her family. Ms C sees dirt too, but is energetic, factual-minded and less emotional. She is actively ready to combat dirt, and she uses soap and water as her weapons.[20]

Stimulus	Ms M	Ms C
Washday	Everyday	Ironing
Fresh	And sweet	Clean
Pure	Air	Soiled
Scrub	Does not; husband does	Clean
Filth	This neighbourhood	Dirt
Bubbles	Bath	Soap and water
Family	Squabbles	Children
Towels	Dirty	Wash

These findings suggest that the market for detergents could be segmented based on attitudes. Firms (such as Procter & Gamble) that market several different brands of washing powders and detergents could benefit from positioning different brands for different attitudinal segments. ■

Mrs C sees dirt, perhaps too much. She is energetic, factual-minded and less emotional. She is actively ready to combat dirt.

There are several variations to the standard word association procedure illustrated here. Respondents may be asked to give the first two, three or four words that come to mind rather than only the first word. This technique can also be used in controlled tests, as contrasted with free association. In controlled tests, respondents might be asked 'What banks come to mind first when I mention 'hi-tech"?' More detailed information can be obtained from completion techniques, which are a natural extension of association techniques.

Completion techniques

Completion technique
A projective technique that requires respondents to complete an incomplete stimulus situation.

Sentence completion
A projective technique in which respondents are presented with a number of incomplete sentences and are asked to complete them.

In **completion techniques**, respondents are asked to complete an incomplete stimulus situation. Common completion techniques in marketing research are sentence completion and story completion.

Sentence completion is similar to word association. Respondents are given incomplete sentences and are asked to complete them. Generally, they are asked to use the first word or phrase that comes to mind, as illustrated in the GlobalCash Project.

e x a m p l e
GlobalCash Project

Sentence completion

In the context of the launch of the euro, the following incomplete sentences may be used.

A manager who does not plan to track their company's transactions using the Internet is

A manager who selects a new bank based on the lowest price is

ABN AMRO is most preferred by

When I think of dividing business between banks, I

This example illustrates one advantage of sentence completion over word association: respondents can be provided with a more directed stimulus. Sentence completion may provide more information about the subjects' feelings than word association. Sentence completion is not as disguised as word association, however, and many respondents may be able to guess the purpose of the study. A variation of sentence completion is paragraph completion, in which the respondent completes a paragraph beginning with the stimulus phrase. A further expanded version of sentence completion and paragraph completion is story completion.

In **story completion**, respondents are given part of a story, enough to direct attention to a particular topic but not to hint at the ending. They are required to give the conclusion in their own words, as in the following example.

Story completion
A projective technique in which respondents are provided with part of a story and are required to give the conclusion in their own words.

example
GlobalCash Project

Story completion

A finance director had been conducting business with a leading Norwegian bank for over 10 years. Her company was planning to make more use of the Internet to manage transactions. The manager that was her main contact at the bank knew little of the applications and benefits of the Internet for her business. So, she spent three months in talks with a variety of software suppliers, primarily trying to get over the problem of integrating new software with the systems she already had. After going through a major selection process and being at the point where she was about to commit her company, another department of the bank contacts her with an Internet solution that solves her integration problems at a much cheaper price compared with her selection process.

What is the finance director's response? . . . .

Why? . . . .

The respondent's completion of this story will reveal characteristics of the relationship she 'enjoys' with the bank. Why, after such a lengthy relationship, may the bank take the business for granted? What role is expected of the bank manager? What must happen in a relationship to get it to the point when switching to another bank is inevitable, whatever the costs? ■

Construction techniques

Construction techniques are closely related to completion techniques. Construction techniques require the respondents to construct a response in the form of a story, dialogue or description. In a construction technique, the researcher provides less initial structure to the respondents than in a completion technique. The two main construction techniques are picture response techniques and cartoon tests.

Construction technique
A projective technique in which respondents are required to construct a response in the form of a story, dialogue or description.

The roots of **picture response techniques** can be traced to the thematic apperception test (TAT), which consists of a series of pictures of ordinary as well as unusual events. In some of these pictures, the persons or objects are clearly depicted, while in others they are relatively vague. The respondent is asked to tell stories about these pictures. The respondent's interpretation of the pictures gives indications of that individual's personality. For example, an individual may be characterised as impulsive, creative, unimaginative, and so on. The term thematic apperception test is used because themes are elicited based on the subject's perceptual interpretation (apperception) of pictures.

Picture response technique
A projective technique in which respondents are shown a picture and are asked to tell a story describing it.

In **cartoon tests**, cartoon characters are shown in a specific situation related to the problem. Respondents are asked to indicate what one cartoon character might say in response to the comments of another character. The responses indicate the respondents' feelings, beliefs and attitudes towards the situation. Cartoon tests are simpler to administer and analyse than picture response techniques.

Cartoon tests
Cartoon characters are shown in a specific situation related to the problem. Respondents are asked to indicate the dialogue that one cartoon character might make in response to the comment(s) of another character.

Expressive techniques

Expressive techniques
Projective techniques in which respondents are presented with a verbal or visual situation and are asked to relate the feelings and attitudes of other people to the situation.

In **expressive techniques**, respondents are presented with a verbal or visual situation and asked to relate the feelings and attitudes of other people to the situation. The respondents express not their own feelings or attitudes, but those of others. The two main expressive techniques are role playing and third-person technique.

Role playing
Respondents are asked to assume the behaviour of someone else or a specific object.

In **role playing**, respondents are asked to play the role or to assume the behaviour of someone else. The researcher assumes that the respondents will project their own feelings into the role.[21] A major use of role playing is in uncovering the nature of a brand personality. The following example shows why the brand personality is of importance to the marketer and how an understanding of personality may be used in advertising.

example

It's a bit of an animal[22]

Brand personality
A brand has a man-made personality and it can survive only if those responsible for it think long term, safeguard its consistency, and ensure its adherence to and compatibility with the needs and attitudes of those for whom it caters. The brand's personality is sacrosanct. Its owners need to guard against squatters, parallel traders, plagiarists and the threat of the brand becoming generic. Above all, they must guard against inconsistency.[23]

The oddball snack product Peperami, strongly flavoured and meat based, in a world where sweet confectionery snacks are the norm, has been an enormous success. Advertising for the product could not be developed without first obtaining real insights into how consumers related to Peperami. Research revealed that there was a widespread view of the brand's personality. Children and adults alike referred to it as bizarre, mischievous, anarchic, impulsive, rebellious and manic. More than 50 rough advertising concepts were then tested to establish how the brand personality could be expressed, from mere eccentricity to naked aggression. What emerged was a swaggering character with a mischievous desire to shock. All the ingredients were there for the advertising solution – the manic, animated Peperami, and the pay-off line, 'It's a bit of an animal'. ■

Brand personality can be uncovered using role playing. In a group discussion scenario, participants may be asked to play out the personality of a brand. In the Peperami example, the setting could be a cocktail bar after work on a Friday evening, with individuals acting out a Peperami brand, Pringles, a Snickers Bar, and a Lidl (store) branded packet of salted peanuts. What the individuals as a brand do, what they say, how they interact with each other in the cocktail scenario, allow an expression of personality that straight questioning may not reveal. Video recording of the event, played back to the group, acts as a means to discuss and elicit further meaning of a brand's personality, highlighting any positive and negative associations of the brand.

Third-person technique
A projective technique in which respondents are presented with a verbal or visual situation and are asked to relate the beliefs and attitudes of a third person in that situation.

In the **third-person technique**, respondents are presented with a verbal or visual situation and are asked to relate the beliefs and attitudes of a third person rather than directly expressing personal beliefs and attitudes. This third person may be a friend, a neighbour, a colleague, or any person that the researcher chooses. Again, the researcher assumes that the respondents will reveal personal beliefs and attitudes while describing the reactions of a third party. Asking an individual to respond in the third person reduces the social pressure to give an acceptable answer.

Advantages and disadvantages of projective techniques

Projective techniques have a major advantage over the unstructured direct techniques (focus groups and depth interviews): they may elicit responses that subjects would be unwilling or unable to give if they knew the purpose of the study. At times, in direct questioning, the respondent may intentionally or unintentionally misunderstand, misinterpret or mislead the researcher. In these cases, projective techniques can

increase the validity of responses by disguising the purpose. This is particularly true when the issues to be addressed are personal, sensitive or subject to strong social norms. Projective techniques are also helpful when underlying motivations, beliefs and attitudes are operating at a subconscious level.[24]

Projective techniques suffer from many of the disadvantages of unstructured direct techniques, but to a greater extent. These techniques generally require personal interviews with individuals who are experienced interviewers and interpreters, hence they tend to be expensive. Furthermore, as in all qualitative techniques, there is the risk of interpretation bias. With the exception of word association, all are open-ended techniques, making the analysis and interpretation more problematic.

Some projective techniques such as role playing require respondents to engage in what may seem to be unusual behaviour. Certain respondents may not have the self-confidence or the ability to fully express themselves with these techniques. In role playing, for example, the skills of acting may make one respondent more articulate at expressing their feelings compared with others. The same may be said of techniques where pictures and cartoons are put together and interpreted, in that distinctive skills may make certain respondents more adept and comfortable in expressing themselves. To counter this, one could argue that there is a great amount of skill required in expressing oneself in an open-ended depth interview. One could point to fiction writers or poets who are able to encapsulate particular feelings most clearly and succinctly, which again is enormously skilful.

With such skill requirements, the disadvantage of the technique lies in the nature of the respondents who agree to participate, and how characteristic they are of distinct target markets.

Applications of projective techniques

Projective techniques are used less frequently than unstructured direct methods (focus groups and depth interviews). A possible exception may be word association, which is commonly used to test brand names and occasionally to measure attitudes about particular products, brands, packages or advertisements. As the examples have shown, projective techniques can be used in a variety of situations.[25] The usefulness of these techniques is enhanced when the following guidelines are observed. Projective techniques should be used:

1 Because the required information cannot be accurately obtained by direct questioning.
2 In an exploratory manner to elicit issues that respondents find difficult to conceive and express.
3 To engage respondents in the subject, by having fun in expressing themselves in interesting and novel ways.

Comparison between qualitative techniques

To summarise comparisons between qualitative techniques, Table 8.1 gives a relative comparison of focus groups, depth interviews, projective techniques and ethnographic approaches (qualitative observation).

The nature of qualitative research is such that, within the broad categories above, there are numerous variations of the techniques with distinct strengths and weaknesses in eliciting and representing consumer feelings. Really, it is not possible to say that one technique is better or worse than the other. Faced with a given problem, it would seem to be the case of deciding which technique is the most appropriate to represent consumers.[26] What may affect this choice is the confidence that marketing

Table 8.1 A comparison of focus groups, depth interviews, projective techniques and ethnographic techniques

Criteria	Focus groups	Depth interviews	Projective techniques	Ethnographic techniques
Degree of structure	Can vary from highly to loosely structured	Can vary from highly to loosely structured	Tends to be loosely structured	Loosely structured, though can have a framework to guide observation
Probing of individual respondents	Low	High	Medium	None when used in isolation and in a covert manner
Moderator bias	Medium	Relatively high	Low to high	None when used in isolation and in a covert manner
Uncovering subconscious information	Low	Medium to high	High	High
Discovering innovative information	High	Medium	Low	Medium
Obtaining sensitive information	Low	Medium	High	High
Involving unusual behaviour or questioning	No	To a limited extent	Yes	Perhaps on the part of the observer

decision-makers may have in particular techniques. Thus, for example, any number of arguments may be made for the use of a projective technique as being the best way to indirectly tackle a sensitive issue. If the marketer who has to use the research findings does not believe it to be a trustworthy technique, then other, perhaps less appropriate, techniques may have to be used.

International marketing research

One of the major contributors to the success of depth interviews and projective techniques is getting the context of questioning right. The context of questioning can have two key components.

The first component is the actual location, such as an office, hotel room or even a bar. Given the issues that are to be tackled and the characteristics of the target respondents, the location plays a significant role in helping the respondent to relax and feel comfortable about responding in a manner that the interviewer is looking for. For some respondents, e.g. a young graduate manager, being questioned about issues relating to 'night life and entertainment', talking in a noisy, smoky bar may be far more natural than talking in a hotel room. The same respondents being questioned about cleaning their homes may be far more comfortable in a quiet 'home' environment, though perhaps not their own home. The interviewer has to work out what location will work best for their target respondents and the issues that will be discussed. They have to appreciate that this location may change between different respondents and may also change as they discover what effect location has on getting the most out of respondents.

The second component of the context of questioning is the protocol of conducting the interview. The protocol can include the clothes the interviewer wears, the manner in which they greet respondents, introduce the interview, conduct the interview and terminate it. This means that, even if an interviewer is technically adept at questioning, they may still not get the most out of respondents. For example, an interviewer who dresses formally when the respondent sees themselves 'off duty' may result in an interview that is tense and lacking in spontaneity.

There are no firm rules about the balance of respondent characteristics, related to issues to be questioned and the context of questioning. The interviewer has to be aware of the balance, and make adjustments as they learn of what works well in drawing out a quality response. Much of what works in a particular context is culturally bound. There can be striking differences between countries in how comfortable respondents may feel talking about issues in their home, for example. There can be striking differences in the protocol of clothing, greeting people, questioning in a direct manner and giving 'gifts' as a reward for taking part in an interview. The problem of 'context' was paramount in the following example. This example illustrates a case in which depth interviews helped to probe in detail the views of professionals, which given the detail of banking relationships could be deemed as highly complex and sensitive.

e x a m p l e

GlobalCash Project

Why are you going to make more use of electronic banking?

The issues tackled in the GlobalCash Project were guided by major pan-European banks from Britain, Finland, France, Germany, Norway, Spain and the USA. These banks formed a steering committee for the project that helped shape the research design. Each bank had different information requirements to support their marketing strategies, which meant that there was a big demand for a wide array of questions to be asked. This was achieved to a great extent but it did not allow great depth to be achieved in many question areas. For example, the questionnaire asked companies what plans they were making to change in particular activities. An example of a response to this question was 'to make greater use of electronic banking'. Statistical analyses could reveal the proportion who were to make greater use of electronic banking, the types of industry they operated in, the types of existing system they had and the banks they worked with. However, the questionnaire could not question *why* they were going to make more use of electronic banking.

Depth interviews were of great importance in probing the reasons behind this planned behaviour. A small sample of the questions and probes included the following questions: Was the greater use driven by new services offered by different banks? What role did the euro play in this decision? Did they foresee any integration between their old systems and new systems? Would additional hardware or software be purchased?

The whole GlobalCash questionnaire had many superficial questions that subsequently could be probed in depth. When around 50% of the responses to the postal survey were received, an interim report was presented to the steering committee. The issues that they deemed the most important to probe formed a topic guide for depth interviews.

With the completed questionnaires, profiles of the respondents could be analysed and those with behaviour relevant to the issues that needed to be explored in depth could be contacted. For example, 'sophisticated' users of cash management services could be profiled. These 'sophisticated' users were then contacted and invited to take part in a depth interview. Around 60 depth interviews were conducted, with a minimum of two interviews in any one country. ■

In the above example, each interview had the same interviewer who ensured a degree of consistency in the approach. This required the interviewer to visit 20 countries in order to conduct the 60 interviews. This interviewer had to have a

strong awareness of the technical issues of cash management banking in order to be seen as credible by the respondents, to be able to question and probe, and to appreciate the relevance of the responses elicited. He had to be aware of the context of interviewing in each of the target countries in order to get the most out of respondents. The first component of context, the actual location for the interviews, was consistent throughout each country. The interviews were held in the offices of the Finance Directors of Europe's largest companies, who made up his respondents. This was an environment where the respondents are naturally used to thinking and talking about the issues under question. It was also an environment where they could talk uninterrupted; the respondent was in control of the interview space and would relax. The second component of context, i.e. the protocol of the interview, was where the interviewer needed local support. Most of the technical language and terms used in cash management banking are 'American'. Many Finance Directors speak English, but some did not, and even if they did, at times they needed to express a view in their native tongue. To understand the protocol and language issues, the interviewer was supported by another interviewer in each country, drawn from local business schools. This interviewer helped to explain the protocol of individual countries, the peculiarities of greeting respondents, posing questions in particular ways and closing an interview. They could help with the translation of questions and responses and ultimately in interpreting the findings. Using one main interviewer allowed a full understanding of the interplay between the respondent, the issues and the context of interview. Using a fellow interviewer in individual countries allowed for the subtle characteristics of protocol and language to be incorporated into the whole process.

Understanding the interplay of respondent, issue and context is vital to the success of depth interviewing and projective techniques. Working in alien cultures in international markets makes this understanding more difficult. Investing in the time and means to develop this understanding ensures that the interviewer generates quality information, getting to really reflect what international respondents feel.

Ethics in marketing research

The essence of qualitative research is that consumers are examined in great depth. They may be questioned and probed in depth about subjects they hardly reflect upon on a day-to-day basis, never mind talk to strangers about. Great care must be taken not to upset or disturb respondents through such intensive questioning. In survey work, reassuring respondents that their responses will be confidential can work if there is a demonstrable link between their responses to questions and a means of aggregating all of the findings – making the individual response 'hidden'. In qualitative research, the process is far more intimate and the link between the response and the respondent is far more difficult to break. This is illustrated in the following example.

example   **Just how 'anonymous' is a quali respondent?**[27]

From both a practical and a methodological perspective, 'confidentiality' in qualitative research is a different concept from confidentiality as applied to survey work. In survey work the emphasis is upon anonymity, i.e. the identity of individual respondents should not be revealed. This creates two problems for qualitative research.

1 Anonymity cannot be promised in qualitative research, especially in the light of current practices where it is increasingly common for clients and others to come to groups, or hear audiotapes, see videotapes and other primary data. This issue further demands consideration of the question: where does the identity reside? In a name, a face, the voice, a turn of phrase, maybe? It is of course also pertinent to ask: do respondents actually care if their identity is revealed?

2 In quantitative research, respondent identity is generally unimportant. The very essence of sampling theory is that a sufficiently large and randomly chosen sample will represent the views, behaviour or attitudes of any known population as a whole. As such, the identity of any one individual is irrelevant to quantitative findings. In sharp contrast, in qualitative research the relationship between the specific individuals and their views is at the heart of analysis and interpretation. You cannot reach qualitative findings without having 'revealed' the individual as part of the research process. Therefore, confidentiality through anonymity is a methodologically untenable concept wherever anyone other than the interviewer is privy to any part of the research process. Yet at the same time, more clients attend groups and more groups take place in viewing facilities. ■

At the end of the second point above comes the reason why the marketing research industry is so concerned about how qualitative research respondents are handled. More consumers are being questioned in both domestic and business scenarios. If they are to 'open up' and reveal deeply held feelings, perhaps in front of a group of strangers, or if they are to take part in projective techniques that they may see as being unorthodox, they have to be reassured of how the data captured will be used. As well as the ethical questions of potentially damaging respondents come the problems of respondents either not willing to take part or, if they do, being very guarded with their responses.

Ethical questions also arise when videotaping sessions with the respondents. Some of the pertinent questions are how much to tell respondents and when the clients should be allowed access. When videotaping respondents, regardless of whether or not they were aware of the camera during the meeting, at the end of the meeting they should be asked to sign a written declaration conveying their permission to use the recording. This declaration should disclose the full purpose of the video, including who will be able to view it. If any respondent refuses, the tape should be either destroyed or edited to omit that respondent's identity and comments completely. The researcher should be sensitive to the comfort level of the respondents, and respect for the respondents should warrant restraint. When a respondent feels uncomfortable and does not wish to go on, the researcher should not aggressively probe or confront any further. It has also been suggested that respondents should be allowed to reflect on all they have said at the end of the interview and should be given the opportunity to ask questions. This may help return the respondents to their pre-interview emotional state.

Internet and computer applications

The Internet has opened up many possibilities to marketing researchers who wish to use depth interviews and projective techniques. Through the use of email, interviewers can reach and question respondents from all over the world. To be able to track down and talk to individuals with the desired qualities for a particular study, without the time and cost implications of travelling, presents a significant benefit of the Internet.

Being able to track down and talk to respondents, i.e. gaining access to qualitative respondents, is vital, but the quality of the discussion with these respondents should be considered.[28] In the case of depth interviews, a full dialogue can develop between an interviewer and respondent, either in real-time or in non-real-time, i.e. a series of emails over a period of time. The discussion, questions and probes can be tailored to specific respondents, allowing them time to reflect upon the issues and express their views in their own manner. The interviewer can present stimuli in terms of images or audio recordings that may help to elicit more from the respondent. It is possible to use Webcams to be able to observe respondents and for respondents to view interviewers, provided of course that both parties have this technology.

However, even with the use of Webcams, much of the non-verbal communication that makes depth interviews work is lost. Subtle changes in facial expression and body language may be missed. These non-verbal forms of communication are important in developing a rapport between interviewer and respondent and are vital in the development of dialogue and the analysis of data. The limitation of visual interaction can make projective techniques difficult to implement. For certain projective techniques such as the array of Completion and Construction techniques, the anonymity afforded by the distance between interviewer and respondent can be a positive feature. Some respondents may find the presence of an interviewer inhibiting when they are trying to think of and present a story completion. Where respondents feel inhibited, working through a response in their own space and time may be the ideal context in which to tackle particular issues.

In evaluating the worth of the Internet in depth interviewing and projective techniques we can return to the GlobalCash example in the International Marketing Research section. Meeting respondents face to face allowed a great richness of dialogue and understanding to be built up. There was no question about how successful the interviews were in understanding *why* respondents behaved in particular ways. However, consider the travel and other costs involved. Consider the time involved in conducting the interviews and just typing up the transcripts of the interviews. Compare the cost and timing requirements of meeting face to face in an international setting with a dialogue by email. The process could have been conducted by email much more quickly and cheaply. The dilemma faced is whether the respondents would have allowed such a dialogue in the first place and how much they would open up and develop a dialogue.

Go to the Companion Website and read Professional Perspective 20 'Live talk on the Web' by Tim Macer. He reviews a software package, Survey Guardian. This package allows for a real-time, 'chat style' on-screen interview.

➤➤➤
See Professional Perspective 20.

Summary

The direct qualitative technique of depth interviewing allows researchers to focus upon individuals with the qualities they deem to be important to their research objectives. With a 'quality individual' the researcher can question and probe to great depth and elicit a quality understanding of that individual's behaviour or feelings. The technique is well suited to tackling commercially and personally sensitive issues. It is also well suited to interviewing children. There are many types of interview that can be applied under the term 'depth interview'. They can range from the very open and unstructured to semi-structured exchanges. The application of structure in a depth interview can be founded on a theoretical underpinning of how individuals should be questioned and probed. Two relevant and widely used examples of depth interviews using a theoretical underpinning are laddering and repertory grid techniques. Laddering seeks to reveal chains of attributes, consequences and values that respondents associate with products. The repertory grid seeks to elicit the underlying elements and the connection of those elements, related to a particular form of consumer behaviour.

Indirect projective techniques aim to project the respondent's motivations, beliefs, attitudes and feelings on to ambiguous situations. Projective techniques may be classified as association (word association), completion (sentence completion, story completion), construction (picture response, cartoon tests) and expressive (role playing, third-person) techniques. Projective techniques are particularly useful when respondents are unwilling or unable to provide the required information by direct methods.

The qualitative researcher needs to develop an understanding of the interplay between characteristics of their target respondents, the issues they will be questioned about and the context in which they will be questioned. The context can be broken down into the physical location of the interview and the protocol of starting, running and terminating the interview. Building up this understanding is vital to the success of conducting depth interviews and projective techniques in international markets.

The marketing research industry is concerned about how qualitative research respondents are handled in interviews and observations. More consumers are being questioned in both domestic and business scenarios. If they are to 'open up' and reveal deeply held feelings, perhaps in front of a group of strangers, or if they are to take part in projective techniques that they may see as being unorthodox, they have to be reassured of how the data captured will be used.

The Internet has opened up many possibilities to conduct depth interviews and use projective techniques on a global basis. A rich dialogue between an interviewer and a respondent can be developed and recorded with much lower cost and time demands when compared with meeting face to face. A loss of the subtle eye contact and body language can be a price to pay for the savings afforded by using the Internet.

Questions

1. What is a depth interview? Summarise the process of administering a depth interview.

2. What are the major advantages of depth interviews?

3. What are the requirements of the researcher undertaking depth interviews? Why are these requirements particularly important when conducting interviews with managers?

4. Why may a structure be applied to the depth interview in the form of a laddering or repertory grid technique?

5 Describe the process of administering the repertory grid technique.

6 Evaluate the context and timing requirements that you think would be needed to make the repertory grid technique work.

7 Choose any particular application of a depth interview and present a case for why you think the technique may work much better than a focus group.

8 What are projective techniques? Under what circumstances should projective techniques be used?

9 Describe the word association technique. Give an example of a situation in which this technique is especially useful.

10 Describe the story completion technique. Give an example of the type of respondent and the context in which such a technique would work.

11 Describe the criteria by which marketing researchers may evaluate the relative worth of qualitative techniques.

12 Why is the context of questioning particularly important when conducting depth interviews in international marketing research?

13 Why may depth interviews or projective techniques upset or disturb respondents?

14 Describe a projective technique that you feel would work particularly well by email – without the use of Webcams.

15 What limitations are there to conducting depth interviews by email, compared with meeting respondents face to face?

Notes

1 De Chernatony, L. and Dall Olmo Riley, F., 'Brand consultants' perspectives on the concept of the brand', *Marketing and Research Today* 25(1) (February 1997), 45–52.

2 Harris, L.M., 'Expanding horizons', *Marketing Research: A Magazine of Management and Application* 8(2) (Summer 1996), 12.

3 'Looking for a deeper meaning', *Marketing* (Market Research Top 75 Supplement) (17 July 1997), 16–17.

4 Bloom, N., 'In-depth research? A bit controversial!' *ResearchPlus* (June 1993), 12.

5 Parke, A., 'When not to rely on the young imagination', *ResearchPlus* (March 1997), 14.

6 De Andrade Marseilles Reis, A.H.M., 'Laddering: an efficient tool for the formation of product positioning strategies and ways of communication', *The Dynamics of Change in Latin America*, ESOMAR Conference in Rio de Janeiro (1997).

7 Dichter, E., *The Strategy of Desire* (New York: Doubleday, 1960).

8 Guttman, J., 'Laddering: extending the repertory grid methodology to construct attribute-consequence-value hierarchies', in D.C. Heath (ed.), *Personal Values and Consumer Psychology* (Lexington, 1984).

9 E.g. Grunert, G.K. and Grunert, S.C., 'Measuring subjective meaning structures by the laddering method: theoretical considerations and methodological problems', *International Journal of Research in Marketing* 12 (1995).

10 Kelly, G.A., *The Psychology of Personal Constructs* (New York: Norton, 1955).

11 Kelly, G.A., *A Theory of Personality: The Psychology of Personality* (New York: Norton, 1969).

12 Marsden, D. and Littler, D., 'Exploring consumer product construct systems with the repertory grid technique', *Qualitative Market Research: An International Journal* 3(3) (2000), 127–44.

13 Van Raaij, W.F. and Verhallen, T.M.M., 'Domain specific market segmentation', *European Journal of Marketing* 28(10) (1994) 49–66.

14 Jin, Z.Q., Birks, D.F. and Targett, D., 'The context and process of effective NPD: a typology', *International Journal of Innovation Management* 1(3) (September 1997), 275–98.

15 Fransella, F. and Bannister, D., *Manual for Repertory Grid Technique* (London: Academic Press, 1977).

16 Sokolow, H., 'In-depth interviews increasing in importance', *Marketing News* (13 September 1985), 26.

17 Best, K., 'Something old is something new in qualitative research', *Marketing News* 29(18) (28 August 1995), 14; Kassarjian, H.H., 'Projective methods', in R. Ferber (ed.), *Handbook of Marketing Research* (New York: McGraw-Hill, 1974), 3.85–3.100.

18 Levy, S.J., 'Interpreting consumer mythology: structural approach to consumer behaviour focuses on story telling', *Marketing Management* 2(4) (1994), 4–9; Hollander, S.L., 'Projective techniques uncover real consumer attitudes', *Marketing News* (4 January 1988), 34.

19 Kennedy, M.M., 'So how'm I doing?', *Across the Board* 34(6) (June 1997), 53–4; Lindzey, G., 'On the classification of projective techniques', *Psychological Bulletin* (1959), 158–68.

20 Krivyakina, M., 'P&G develops new Tide product for high-efficiency machines', *Chemical Market Reporter* 251(13) (31 March 1997), 9; 'Interpretation is the essence of projective research techniques', *Marketing News* (28 September 1984), 20.

21 For issues involved in role playing, see Suprenant, C., Churchill, G.A. and Kinnear, T.C. (eds), 'Can role playing be substituted for actual consumption?', *Advances in Consumer Research* (Provo, UT: Association for Consumer Research, 1984), 122–6.

22 *Campaign* (28 October 1994), S13.

23 Van Mesdag, M., 'Brand strategy needs turning back to front', *Marketing Intelligence and Planning* 15(2–3) (Feb–March 1997), 157.

24 McGrath, M.A., 'Gender differences in gift exchange: new directions from projections', *Psychology & Marketing* 12(5) (August 1995), 371–93; Schnee, R.K., 'Quality research: going beyond the obvious', *Journal of Advertising Research* 28 (February–March 1988), RC-9–RC-12; Kerlinger, F.N., *Foundations of Behavioural Research* 3 (New York: Holt, Rinehart & Winston, 1986), 471.

25 For more on projective techniques, see Valentine, V. and Evans, M., 'The dark side of the onion: rethinking the meaning of 'rational' and 'emotional' responses', *Journal of the Market Research Society* 35 (April 1993), 125–44.

26 Colwell, J., 'Qualitative market research: a conceptual analysis and review of practitioner criteria', *Journal of the Market Research Society* 32(1) (1990), 33.

27 Imms, M., 'Just how "anonymous" is a quali respondent?', *Research* (April 1997), 20.

28 Pincott, G. and Branthwaite, A., 'Nothing new under the sun?', *International Journal of Market Research* 42(2) (2000), 137–55.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Qualitative research:
data analysis

Objectives

After reading this chapter, you should be able to:

1 understand the importance of qualitative researchers being able to reflect upon and understand the social and cultural values that shape the way they gather and interpret qualitative data;

2 describe the stages involved in analysing qualitative data;

3 describe the array of data types that qualify as qualitative data;

4 explain the nature and role of coding in the stage of reducing qualitative data;

5 appreciate the benefits of being able to display the meaning and structure that a qualitative researcher sees in their data;

6 understand why qualitative data analysis pervades the whole process of data gathering and why the stages of analysis are iterative;

7 appreciate the strengths and weaknesses of analysing data using qualitative analysis software and have the means to experiment further with a number of online demonstrations;

8 understand the ethical implications of the ways that qualitative researchers interpret data.

Qualitative analysis involves the process of making sense of data that are not expressed in numbers.

Overview

The application of qualitative techniques can see researchers changing direction as they learn what they should focus their attention on. The techniques, the nature of respondents and the issues explored can change and evolve as a project develops. This chapter starts by examining how the researcher reflects upon what happens to the way they perceive and observe as these changes occur. It discusses how these reflections form a key source of qualitative data to complement the narrative generated from interviews and observations.

The stages involved in the process of analysing qualitative data are outlined and described. The first stage of the process involves assembling qualitative data in its rich and varying formats. The second stage progresses on to reducing the data, i.e. selecting, classifying and connecting data that researchers believe to be of the greatest significance. A key element of this stage is the concept of coding. Coding qualitative data is described and illustrated in some detail, including coding for the concept of grounded theory as introduced in Chapter 6. The third stage involves display data, i.e. using graphical means to display the meaning and structure that a researcher sees in the data they have collected. Manual and electronic means of displaying data are discussed. The final stage involves verifying the data. The marketing researcher aims to generate the most valid interpretation of the data they collect, which may be supported by existing theories or through the concept of theoretical sampling. Though these stages seem quite distinct, the reality is that they are iterative and totally interdependent upon each other; the stages unfold in 'waves' to produce an ultimate interpretation of great value to decision-makers.

The use of computers in the stages of qualitative data collection and analyses is described. To be able to cope with the great amount of data generated from qualitative techniques, a great variety of software packages are available. Examples of analysis software are briefly described followed by Internet addresses that allow demonstration versions of the software to be downloaded and explored. There are many distinct advantages to the use of qualitative data analysis software, but many researchers contend that it should be a 'hands-on' process that cannot be mechanised. The arguments from both of these perspectives are presented. The social and cultural values of qualitative researchers affect how qualitative researchers gather and analyse data. Understanding the social and cultural norms of respondents in international environments is discussed. The chapter concludes by examining how the social and cultural values of researchers affect their interpretation of qualitative data and the ethical implications of not reflecting upon these values.

The qualitative researcher

Self-reflection of social and cultural values

In Chapter 2 when discussing the diagnosis of research problems on page 49 we stated:

> *A major problem for researchers is that their perception of problems may be reflected through their own social and cultural development. Before defining the problem, researchers should reflect upon their unconscious reference to cultural values ... The unconscious reference to cultural values can be seen to account for these differences.*

This implies that the marketing researcher needs to reflect upon their own values and attitudes, the factors that may bias the way they perceive and what they observe. This reflection is just as important in the analysis of qualitative data as it is in the diagnosis

of research problems. To illustrate why researchers need to reflect upon what may bias the way they perceive and what they observe, we start this chapter with an example from the world of literature and the treatment of narrative. The example is a précis of an English translation of a Japanese novel; the example could be derived from any novel.

South of the Border, West of the Sun[1]

This novel tells the story of an only child, Hajime, growing up in the suburbs of post-war Japan. His childhood sweetheart and sole companion in childhood was Shimamoto, also an only child. As children they spent long afternoons listening to her father's record collection. When Hajime's family moved away, the childhood sweethearts lost touch. The story moves to Hajime in his thirties. After a decade of drifting he has found happiness with his loving wife and two daughters, and success in running a jazz bar. Then Shimamoto reappears. She is beautiful, intense, enveloped in mystery. Hajime is catapulted into the past, putting at risk all he has at the present. ■

Imagine that you had been asked to read this novel, but before you read it you were expected to prepare by reading a description of conditions in post-war Japan. From that you may appreciate the significance of a record collection of 15 albums, and how privileged a family may be to own a record player and to have this collection. Imagine someone else being asked to prepare by reading a biography of the author. From that you may appreciate the social and economic conditions of his upbringing, the literature, music and education that he enjoyed. Preparing to read the novel in these two ways may mean that the reader sees very different things in the story. They may interpret passages differently, have a different emotional attachment with the conditions and behaviour of the characters, and appreciate the effect of quite subtle events upon the characters.

Put aside any prior reading and imagine a female reader enjoying the book. She may empathise with the main female character Shimamoto and understand her attitudes, values and behaviour in the way that male readers may not be able to comprehend. In the story, Shimamoto suffered from polio as a child, which made her drag her left leg. Imagine a reader who has had to cope with a disability and who may appreciate how as a child one copes with the teasing of young children. The two main characters were 'only children'; imagine the reader who was an only child and who can recall how they would view large families and appreciate the emotions of the only child. The list could go on of the different perspectives of the story that may be seen. The reader, with their inherent values and attitudes, may perceive many different things happening in the story. The reader does not normally reflect upon their unconscious values and attitudes; they just enjoy the story. In talking to others about the story, they may be surprised about how others see it. In watching the film version of the book, they may be shocked at how different an image the film director presents, an image that is very different from the one that resides in their head. Now consider whether there is one ultimate interpretation of the novel, one ultimate 'truth'. It is very difficult to conceive that there is one ultimate interpretation. One may question why anyone would want to achieve such a thing; surely the fun of literature is the ability to have multiple interpretations and 'truths' of a novel.

Narrative for the qualitative researcher

What is the link from the interpretation of a novel to qualitative data analysis in marketing research? Quite simply, the qualitative marketing researcher builds up a narrative and creates a story of the consumers whom decision-makers wish to understand. Imagine yourself as a qualitative marketing researcher, supporting

decision-makers who wish to develop advertisements for an expensive ride-on lawn-mower. The key target market they wish to understand is 'wealthy men, over the age of 60 who own a home(s) with at least 1 hectare of garden'. The decisions they face may include the understanding of:

1 What gardening and cutting grass mean to target consumers.
2 How they feel about the process of using a lawnmower.
3 What relative values (tangible and intangible) are inherent in different brands of lawnmower.
4 What satisfaction they get from the completed job of mowing a large lawn.
5 The nature and qualities of celebrities they admire (who may be used to endorse and use the product in an advertisement).

These questions may be tackled through the use of focus groups. Imagine yourself running these groups. What could you bring to the groups if you have personally gone through the experience of buying an expensive ride-on lawnmower and have gardening and lawnmowing experiences? You may have an empathy with the respondents in the same manner as the 'only child' reading of the experiences and emotions of an only child in a story. From this empathy, you may be able to question, probe and interpret the respondents' answers really well, drawing an enormous amount from them. Without those experiences you may have to devise ways to 'step into the shoes' of the respondents. You may look to the attitudes, values and behaviour of your parents, grandparents or friends for a start, looking for reference points that you are comfortable with, that make sense to you. As you go through a pilot or experimental focus group, you may be surprised by certain respondents talking about their lawnmowers as 'friends', giving them pet names and devoting lavish care and attention upon them. Getting an insight into this may mean looking at cases from past research projects or literature from analo-gous situations such as descriptions of men forming a 'bond' with their cars.

The direction that the qualitative marketing researcher takes in building up their understanding and ultimately their narrative is shaped by two factors. The first factor is the *theoretical understanding* of the researcher as they collect and analyse the data. This theoretical understanding can be viewed from two perspectives. The first is the use of theory published in secondary data, intelligence and literature. The use of theory from these sources may help the researcher to understand what they should focus their attention upon, in their questioning, probing, observations and interpreta-tions. The second is the use of theory from a grounded theory perspective. The researcher may see limitations in existing theory that do not match the observations they are making. These limitations help the researcher to form the focus of their ques-tioning, probing, observations and interpretations.

The second factor that shapes the direction that the researcher takes is a *marketing understanding*. In the case of understanding the wealthy male lawnmower owner, the researcher needs to understand what marketing decision-makers are going to do with the story they create. The researcher needs to appreciate the decisions faced in creat-ing an advertisement, building a communications campaign or perhaps changing features of the product. Reference to theoretical and marketing understanding in the researcher helps them to present the most valid interpretation of their story to decision-makers. Unlike writing a novel, where the author is happy for the reader to take their own 'truth', the marketing researcher is seeking an ultimate interpretation and validity in their story. Achieving a valid interpretation enables the researcher to convey to decision-makers a vision or picture of a target market that they can quickly 'step into'. A marketing decision-maker, for example, may wish to include a passage of music in an advertisement that the target market has an emotional attachment to, that they find positive and uplifting. With a rich picture or vision of this target market they may be able to choose the right piece of music. The decision-maker's cultural

Be prepared to get lost and to change direction as a qualitative researcher.

and social development may mean that the piece of music is meaningless to them, but you as a researcher have given them the confidence and enabled them to step into the world of the target market. The following example represents the views of an advertising practitioner praising the support gained from qualitative researchers through their understanding of the issues faced by decision-makers.

e x a m p l e

Cinderella's getting ready for the ball[2]

I think a lot of researchers have worked hard at understanding advertising and the advertising process and are now regarded as quite critical contributors. What has allowed them to become 'wise' advisors rather than providers of fuel for literal and obedient fools? I think the following:

1 *There has been more crossover between planners and qualitative researchers*. Qualitative researchers understand their role in the process of developing advertisements and building brands. They want to know what the creative brief is and what responses or feelings an advert is meant to elicit. For example, when developing a current and famous premium package lager, we purposely discussed what levels of initial bewilderment or mild alienation were acceptable. The campaign had to challenge consumers – hopefully the brightest ones would let their mates in on the secret jokes down the pub!

2 *Good researchers work to more 'holistic' views of how ads work and thus avoid mechanistic and simplistic diagnoses*. We may therefore use projective techniques to help consumers express feelings. Whatever it is, good researchers try and look at advertisements in totality, not as a disaggregated series of frames.

3 *Good qualitative researchers, planners and clients are much more conscious about stimuli in advertising research*. People put more effort into defining stimuli – e.g. using film clips to get across production values or special effects, mood boards to give texture alongside traditional animatics, narratives or key frames. We have showed *Terminator 2* to help consumers imagine products 'metamorphosing' from one thing to another, and music snippets to conjure up different feelings of sensuality and sexuality.

4 *Interpretation has improved*. Good qualitative researchers know that consumers in group discussions are likely to be evaluative/judgemental, cynical and literal. It is the researcher's job to get them to 'imagine' beyond that and to interpret responses using one's cumulative knowledge of advertising chemistry. ■

The researcher's learning as qualitative data

Qualitative marketing researchers have to reflect upon their own social and cultural development, their own attitudes and values to see how these have shaped the narrative and how they shape their interpretation of the narrative. The researcher should recognise their own limitations and the need to develop and learn; in the case above, this means learning about wealthy men and their relationship with lawnmowers. Ultimately they wish to present the most valid story that they see, to have examined the story from many perspectives, to have immersed themselves in the world of their target markets.

If you are reading a novel, you may not be inclined to make notes as your reading progresses. You may not make notes of other books to read that may help you to understand the condition of particular characters, or to understand the environment in which they behave. You may not wish to write down the way that you change and learn as you read through the story. A reflection of your unconscious social and cultural values as revealed through your interpretation of the novel may be the last thing you want to do.

As a qualitative researcher you need to do all the above as you build and interpret your story of target consumers. A notebook should be on hand to note new question areas or probes you wish to tackle, and to reflect upon how they have worked. As interviews unfold and you feel your own development and understanding progress, a note of these feelings should be made. As you seek out specific secondary data, intelligence or theory to develop your understanding, you should note why. If you see limitations in existing theories or ideas, you should note why. As an understanding of how decision-makers can use the observations that are being made, these should be recorded. Ultimately the story that emerges in your own notebook should be a revelation of your own social and cultural values. There should be an explicit desire to develop this self-awareness and understand how it has shaped the direction of an investigation and the ultimate story that emerges.

The creation and development of the researcher's notebook is a major part of the narrative that is vital to the successful interpretation of questions and observations of consumers. The key lesson that emerges from the creation and development of the researcher's notebook is that qualitative data analysis is an ongoing process through all stages of data collection, not just when the data have been collected.

> *Analysis is a pervasive activity throughout the life of a research project. Analysis is not simply one of the later stages of research, to be followed by an equally separate phase of 'writing up results'.*[3]

The evolution of questions, probes and even deciding who should be targeted for questions or observations means that analysis takes place as data are being gathered.

The process of qualitative data analysis

The process of analysing qualitative data can be encapsulated in the four stages outlined in Figure 9.1.

Data assembly

Data assembly means the gathering of data from a variety of sources. These would include:

Data assembly
The gathering of data from a variety of disparate sources.

1 Notes taken during or after interviewing or observations.
2 Reflections of researchers, moderators or observers involved in the data collection process.
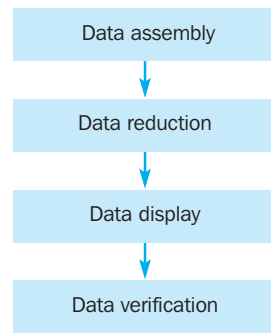
Figure 9.1
Stages of qualitative data analysis

3 Theoretical support – from secondary data, intelligence or literature sources.
4 Documents produced by or sourced from respondents.
5 Photographs, drawings, diagrams, i.e. still visual images.
6 Audiotape recordings and transcripts of those recordings.
7 Videotape recordings.
8 Records made by respondents such as mood boards or collages.

Field notes
A log or diary of observations, events and reflections made by a researcher as a study is planned, implemented and analysed.

As discussed in the previous section, the researcher should get into the habit of maintaining a notebook or **field notes**. As a qualitative investigation evolves in terms of the issues to explore and the respondents to target, the researcher goes through a learning process. This learning process means that the researcher may see things differently as interviews or observations progress. Keeping field notes aids the researcher's memory when it comes to the formal process of data analysis and helps enormously in categorising and interpreting collected data. It ultimately helps to generate a 'deeper and more general sense of what is happening'.

In order to make 'deeper and more general sense of what is happening', it is suggested[4] that researchers keep four separate sets of notes:

1 Short notes made at the time of observation or interview.
2 Expanded notes made as soon as possible after each session of interviews or observations.
3 A fieldwork journal to record problems and ideas that arise during each stage of fieldwork.
4 A provisional running record of analysis and interpretation.

These suggestions help to systematise field notes and thus improve their reliability. Data assembly also includes deciding lines of enquiry which should be developed and those that should be dropped. Given that qualitative research is primarily exploratory in nature, questions and probes are not fixed. As an interview or observation takes place, the researcher learns more about an issue and can develop a new question or probe and decide that a question, initially thought to be vital, is no longer relevant. There may be issues that can be compared over a series of interviews or observations, but the whole data collection and data assembly can evolve. Keeping notes is vital as memory alone is fallible, unreliable and potentially biased. Being able to recall, for example, the hesitation in replying to a question displayed by a focus group respondent, may upon reflection be seen as someone evading the issue. After all the group discussions have been completed and the same question has been posed to others, the interpretation may change to the individual being embarrassed about an issue, primarily through becoming aware of his or her own ignorance. The researcher's notes help them to recall how they were feeling at the point of setting the question, and recall the situation in other groups that gives meaning to *a pause that shows up as a quiet spot in an audio or video recording.*

Data reduction

Data reduction
The organising and structuring of qualitative data.

Data reduction involves handling the data. This process involves organising and structuring the data. It means having to throw some data away! Imagine a series of 10 focus group discussions and the amount of data that can be collected. There are the memories and notes of the moderator and any other observers who took part, there are the **transcripts** of what was actually said in interviews, and there may be contributions from respondents in the form of mood boards. The transcripts are a vital data source in qualitative data analysis and much care should be taken in typing them up. Tape recordings are notoriously 'unclear'. Imagine a focus group in full swing: not every respondent takes their turn to speak without talking over other respondents, and then they may not speak clearly and loudly enough. As a result it can take a great deal of time to work out what respondents actually said and how the questions, responses and ideas connect together. In producing transcripts, it is much better for the researcher to work through the tape recordings and piece together the components using their notes and memory of events. This is very time-consuming, so many researchers use typists to transcribe their tape recordings of interviews, arguing that their time is better spent reading through and editing transcripts produced in this manner. The use of the Internet in depth interviews and focus groups means that this time-consuming task is eliminated as the transcript is built up as the interview progresses.

Transcripts
'Hard copies' of the questions and probes and the corresponding answers and responses in focus group or depth interviews.

The researcher with their transcripts, notes and other supporting material has to decide what is relevant in all these data. Reducing the data involves a process of **coding data**, which means breaking down the data into discrete chunks and attaching a reference to those chunks of data. Coding is a vital part of coping with qualitative data analysis, especially in the context of developing theory with a grounded theory approach. Given this importance, the process is discussed in some detail.

Coding data
Breaking down qualitative data into discrete chunks and attaching a reference to those chunks of data.

Coding data. Researchers need to be able to organise, manage and retrieve the most meaningful bits of qualitative data that they collect.[5] This is normally done by assigning 'labels' or codes to the data, based upon what the researcher sees as a meaningful categorisation. What happens is that the researcher condenses the great mass of data from a qualitative study into analysable units by creating categories from the data. This process is termed the coding of data.

Table 9.1 presents the verbatim responses from an open-ended question in a self-completion survey targeted at 12–14-year-olds. The question asked respondents what facilities they would like in a planned new community centre. Though the technique

Table 9.1 Teenager requests for facilities at a planned community centre

Requested feature of new community centre	Gender
Skate park, death slide, basketball courts, swimming pool	Male
Computer room	Male
Stuff for all ages	Male
Swimming pool	Male
Computers, snooker room	Male
A space for computers, tuck shop	Male
Music, television, up-to-date magazines, pool tables	Female
Music, discos	Female
Swimming pool	Female
Music, pool/snooker, discos	Female
What people will enjoy	Female
All the things people enjoy	Female

used was quantitative, the survey generated qualitative data and in its much shortened format illustrates the process that qualitative researchers must go through.

In categorising the responses, the researcher could create codes of '**swimming pool**' or '**disco**' and count the times that these were literally expressed. Alternatively, the researcher could code on the basis of '**sports activities**' or '**recreational activities**' and group together activities such as 'swimming, basketball and snooker' for sports and 'computers, television, discos and tuck shop' for recreational activities. They could code '**indoor activities**' and '**outdoor activities**', or activities that would need supervision and those that would need no supervision. There are many ways that the researcher can categorise the data – it is their choice. Consider how the researcher may cope with the requests for a 'computer room' and 'computers'. Could these be combined under one heading of '**computing**' or would this lose the meaning of having a devoted space, away from other activities that could be noisy and distracting? Consider also how the researcher would cope with the requests for 'stuff for all ages', 'what people will enjoy' and 'all the things people enjoy'. It may seem obvious that a new leisure centre needs to develop facilities that people enjoy and that these may be discarded, but there may be a hint in the first statement of 'stuff for all ages' that may link to the word 'people' used in the two other statements. If the researcher interprets the statements in this way, a category of '**activities to draw in all ages**' could be created; these responses may be seen as tapping into a notion of a leisure centre that is welcoming and not exclusive.

Table 9.2 presents a small selection of the verbatim responses from the same open-ended question in a self-completion survey, this time targeted at adults.

Table 9.2 Adult requests for facilities at a planned community centre

Requested feature of new community centre	Gender
Regards for residents living nearby, special car parking area to avoid streets nearby being jammed	Male
New centre would soon bring the wrong sort of people; it could form a centre for thugs and crime	Male
Strict rules so as to inconvenience local people living close as little as possible, e.g. noise	Male
Run and organised well to run functions at affordable prices with dress rules for the lounge and bar	Male
Membership should be given on signature of applicants to a strict set of rules	Male
Emphasis on youth on the estate and run in a way to encourage rather than regiment them	Male
Supervised youth activities, daytime crèche, dance floor, serve coffee/soft drinks for youths	Female
Should be very welcoming and developed for all kinds of people	Female
Active participation by those using the facilities which should give opportunities for the young	Female
To make a safe place for all people of all ages to enjoy	Female
Exterior should be modern. Inside decorated tastefully with nice seats and tables, plenty of hall space	Female
Youth club with a youth leader. Luncheon club for older groups and gentle keep-fit for the older.	Female

The interesting feature in comparing the statements from the adults with those from the teenagers is how they express themselves in more detail and how they thought beyond specific facilities that make up the leisure centre. These statements were unprompted, so one can imagine how much richer the explanations and justifications would be with a face-to-face interview. Again, there are many ways that the researcher can categorise the data, perhaps even more than with the teenagers. Categorising these adult statements is not as straightforward as for the teenagers. The researcher could draw out the words '**youth**' or '**rules**' and set these as categories. They could pull out named '**facilities**' such as 'dance floor' and 'nice seats and tables' or '**activities**' such as 'youth club' and 'luncheon club'. What becomes apparent in reading through the statements (especially with the full set of responses) are implied

problems related to issues of parking, the types of people that are attracted or could be attracted, and how 'regimented' or not the centre should be. These are themes or patterns that may be apparent to a reader, though not explicitly expressed. There may be words expressed that make up the themes, but the words broken down and taken in isolation may lose their impact, if they are just counted.

Table 9.2 illustrates that categorisation into the component words may mean that the contextual material that gives these words meaning can be lost. From the above example, coding can be thought of as a means to:

1 *Retrieve data*, i.e. from the whole mass of data, particular words or statements can be searched for and retrieved to examine the 'fit' with other words or statements.
2 *Organise the data*, i.e. words or statements can be reordered, put alongside each other and similarities and differences evaluated.
3 *Interpret data,* i.e. as words or statements are retrieved and organised in different ways, different interpretations of the similarities and differences can be made.

Coding is a process that enables the researcher to identify what they see as meaningful and to set the stage to draw conclusions and interpret the meaning. Codes are essentially labels to assign meaning to the data compiled during a study. Codes usually are attached to 'chunks' of varying size – words, phrases, sentences or whole paragraphs, connected or unconnected to a specific setting.[6] Coding can be examined from two perspectives. First, it can be thought of as a means to simplify or reduce the mass of data. If the codes are kept to a general level and their number is relatively small, then the data can be 'stripped down' to a simple general form. This coding approach can be compared directly to simple forms of content analysis.[7] Second, it can be thought of as a means to expand, transform and reconceptualise data, opening up more diverse analytical possibilities. The general analytical approach is to open up the categories in order to interrogate them further, to try to identify and speculate about further features. Coding here is about going beyond the data, thinking creatively with the data, asking the data questions, and generating theories and frameworks.[8] This perspective of coding matches the process of grounded theory introduced in Chapter 6. We describe the coding process in grounded theory to illustrate how theory may develop from expanding, transforming and reconceptualising data.

Coding in grounded theory. This aims to organise data into a set of themes or codes in three phases. These phases should be understood as different ways of handling the data in which the researcher moves back and forth in iterations to develop each phase. They should not be seen either as distinguishable procedures nor as temporarily separated phases in the process:[9]

1 *Open coding* aims at expressing the data in the form of 'concepts'. For this purpose, data are first 'disentangled'. The individual words or short sequences of words that make up an expression are drawn out and the meanings that could be attached to them explored. An example of this is presented from an expression taken from Table 9.2: *Exterior/ should/ be modern/. Inside/ decorated/ tastefully/ with nice/ seats and tables/, plenty of/hall space/.* Examining each of these chunks in detail helps to develop a deeper understanding of the statement, e.g. consider the many interpretations of 'modern' when referred to a building. Examining the chunks also opens up possibilities in interpreting other parts of the text or other statements in different ways, e.g. linking 'modern' to 'seats and tables' could open up the possibilities of interpreting 'modern' as *modernist* 1930s style; a contemporary *reproduction* of that period; or a style that is truly contemporary, *21ˢᵗ century* modern or even

futuristic. Open coding may be applied in varying degrees. The data can be coded line by line, sentence by sentence or paragraph by paragraph. Which of these alternatives is chosen depends upon the research question, on the amount and complexity of the qualitative data, on the style of the researcher and on the stage that the analysis has reached. The result of open coding should be a list of codes that were attached to the data. These are complemented by the field notes made by the researcher as they collected the data and as they go through the coding process. The open coding process is summarised by Strauss and Corbin as:[10]

> *Concepts are the basic building blocks of theory. Open coding in grounded theory is the analytic process by which concepts are identified and developed in terms of their properties and dimensions. The basic analytic procedures by which this is accomplished are the asking of questions about the data, and the making of comparisons for similarities and differences between each incident, event and other instances of phenomena. Similar events and incidents are labelled and grouped to form categories.*

2 *Axial coding.* This step helps to refine and differentiate the categories that result from open coding. From the potential multitude of categories that can be originated, certain ones are selected that the researcher sees as being the most promising and worthy of further development. The significant element of this stage is to establish relationships between categories and sub-categories; these relationships are clarified or established. The axial coding process is summarised by Strauss and Corbin as:[11]

> *Axial coding is the process of relating sub-categories to a category. It is a complex process of inductive and deductive thinking* [see Chapter 6, page 141] *involving several steps. These are accomplished as with open coding, by making comparisons and asking questions. However, in axial coding the use of these procedures is more focused and geared towards discovering and relating categories.*

How would your hairdresser interpret your request for a 'modern' hairstyle?

3 *Selective coding.* This step continues axial coding at a higher level of abstraction. The aim of this step is to develop and elaborate a core category around which the other developed categories can be grouped and by which they are integrated.[12] The researcher must decide between equally important phenomena and evaluate them, so that one central category results. The core category is developed in its features and dimensions and linked to other categories. The analysis and the development of the theory built around this core category aim at discovering patterns in the data as well as the conditions under which these apply. Grouping the data in this way should enable the researcher to say '*under these conditions* [*list and describe*] *this happens; whereas under those conditions* [*list and describe*] *this is what occurs.* Theory is formulated in detail and checked against the data. The procedure of interpreting the data, like the integration of additional data, ends at a point known as theoretical saturation, i.e. a point where further coding, enrichment of categories and evaluation of connections no longer provides any new knowledge.

Throughout the process of coding in grounded theory, the researcher should regularly address the data, as it is built up and in its entirety, with the following list of questions. Attempting to answer these facilitates the questioning of how codes are defined, connected and subsumed into broader categories:[13]

- *What?* What is it about here? Which phenomenon is mentioned?
- *Who?* Which persons are involved? What roles do they play? How do they interact?
- *How?* Which aspects of the phenomenon are mentioned (or not mentioned)?
- *When? How long? Where?* Time, course, location?
- *How much? How strong?* Aspects of intensity.
- *Why?* Which reasons are given or can be reconstructed?
- *What for?* With what intention, to what purpose?
- *By which?* Means, tactics and strategies for reaching that goal.

Coding is a major process involved in data reduction. The process forces the researcher to focus upon what they believe to be the most valid meaning held in the data. In order to develop that meaning further, the researcher needs to communicate their vision to others, to evaluate their interpretations of the data and to reflect upon their own vision. The stage of data display is the means by which researchers communicate their vision of meaning in the data.

Data display

Data display
Involves summarising and presenting the structure that is seen in collected qualitative data.

Data display involves summarising and presenting the structure that is seen in the collected data. The display allows a 'public' view of how the researcher has made connections between the different 'data chunks'. Even if others may not have made the same connections and interpret the data in exactly the same manner, the logic of connections should be clear. The display may be in a graphical format, with boxes summarising issues that have emerged and connecting arrows showing the interconnection between issues. Verbatim quotes can be used to illustrate the issues or the interconnections. Pictures, drawings, music or advertisements can also be used to illustrate issues or interconnections. The overall structure allows the marketer to see the general meaning in the collected data. The illustration of issues or interconnections brings that meaning to life.

One of the simplest means to display data is through the use of a spreadsheet. This can be built up and displayed in a manual or electronic format. Table 9.3 presents an example of how a spreadsheet may be set out. This spreadsheet is a sample of all the interviews that may be conducted and the number of issues that may be tackled. The example relates to a bus and tram operator who wishes to understand the attitudes

Table 9.3 Spreadsheet data display of focus group discourse

Interviews? Issues, e.g.?	Group 1 – 18–25 yr old male car drivers	Group 2 – 18–25 yr old female car drivers	Group 3 – 18–25 yr old male bus and tram users	Group 4 – 18–25 yr old female bus and tram users	Notes on the similarities and diffferences between groups on issues?
Evening travel	Verbatim discourse taken from the interview that relates to this issue				
Commuting					
Freedom					
Friends					
Notes on the dynamics of individual groups?					

and behaviour of 18–25-year-olds related to using public transport. In the columns, details of each interview are presented, and in the final column, notes are made of observations between interviews with a focus on each issue. In the rows, the issues that were discussed in the interviews are presented. These issues may be generated from the topic guide used and/or from the notes of the researcher related to what they see as the emerging issues. The final row details notes of the dynamics of the group, explaining why particular exchanges may be interpreted in a particular way. The analyst cuts and pastes extracts from the transcripts into the relevant cells. With the spreadsheet built up of the reordered transcripts (each focus group may tackle the issues in a different order and with different emphases), comparisons can be made across the columns on particular issues, looking for similarities and differences. Connections between issues can be mapped out with the use of arrows to show the flow of dialogue. The responses from types of respondents such as 'city-dwellers' or 'suburb-dwellers' can be colour coded to examine differences. Different notes, images or any other supplementary material can be pasted on to the spreadsheet to help in the interpretation; all the assembled data can be displayed.

Such a spreadsheet can be built up manually using large sheets of paper from, e.g. flip charts, divided into a grid, and the evidence such as chunks of the transcript physically pasted in. The big advantage of this approach is being able to visualise the whole body of data and to move around the data to 'play' with ideas and connections. This works particularly well when there is more than one person working on the analysis and they are drawing ideas and questions out of each other as they relate to the data. The disadvantage is that editing, moving data around and re-categorising data can become very cumbersome and messy. This is where electronic means of displaying the data work well. With electronic means, images and notes can be scanned in and added to the transcripts. Changes can be made very easily and quickly in moving data around, re-categorising and incorporating new material. Different versions can be easily stored to allow an evaluation of how the thought processes of the researcher have developed. The disadvantage of the approach is that, when attempting to view the data in its entirety, the entire dataset is there but in effect is viewed
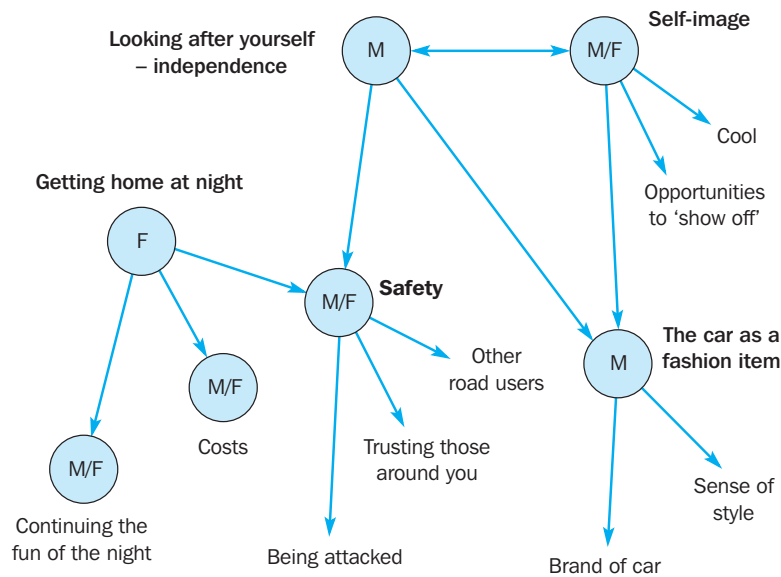
Figure 9.2
Flow chart depicting how 18–25-year-olds view public transport

through a 'window' with a limited field of vision. The 'window' can be readily moved about but the overall perspective is limited.

The other major means of displaying data is to use flow charts. Figure 9.2 displays a very basic structure of the issues or major categories and sub-categories related to how 18–25-year-olds view the use of public transport after an evening out.

Visualising the data in this matter can allow the researcher to dip back into the transcripts and their notes to seek alternative ways of connecting evidence and justifying connections. This means that this form of graphic can play a vital role in data reduction and coding, i.e. in making sense of the data, as well as in portraying a final interpretation of the data. Most proprietary qualitative analysis software packages allow data structures to be displayed as in Figure 9.2 but with far more sophisticated features to display structure, differences and supporting evidence. A simple illustration of this in Figure 9.2 is the 'M/F' label attached to categories, used to display behavioural tendencies of male or female respondents. With an analysis package, quite distinctive structures for respondent types may be mapped, with the ability to tap into supporting evidence in actual categories or in the links between categories.

Once the researcher has displayed what they see as the meaning in the data, they need to demonstrate the credibility of their vision. This involves data verification.

Data verification

Data verification
Involves seeking alternative explanations of the interpretations of qualitative data, through other data sources.

Data verification involves seeking alternative explanations through other data sources and theories. Researchers need to demonstrate that they have presented a valid meaning of the data that they have collected. They need to show that the structure or meaning they see is not just a reflection of their own views. This is where the concept of theoretical understanding as discussed at the start of this chapter can help. It is also where the use of the researcher's field notes proves to be invaluable. The use of theory from secondary data, intelligence and literature can help to guide what may be reasonably expected as a meaning. Other means to verify the data can be through seeking 'similar' research findings and explanations taken from different contexts, different time frames and different researchers.[14] Though the findings from these different scenarios will not be the same, there can be themes that give qualitative researchers the confidence that they are representing a valid view of their respondents.

If the researcher is using a grounded research approach and is seeking to develop new theory, then the approach they use to validate their theory is through theoretical sampling (introduced in Chapter 6). Applying theoretical sampling in this context means seeking out different situations and learning from the comparisons that can be made from these situations. The purpose is to go to places, people or events that will maximise opportunities to discover variations among concepts. The grounded theorist uses theory initially (even if it is in a highly critical and dismissive manner) to help guide which cases they should focus upon, the issues they should observe and the context of their investigation. As their research design evolves they seek to develop new theory and do not wish to be 'blinkered' or too focused on existing ideas. They will seek multiple explanations of the phenomena they observe and will create what they see as the most valid relationship of concepts and, ultimately, theory. The procedure of interpreting the data, like the integration of additional data, ends at a point known as theoretical saturation, i.e. a point where further coding, enrichment of categories and evaluation of connections no longer provide any new knowledge.

The concept of validity will be examined in more detail in Chapter 12.[15] At this stage it is worth noting that the qualitative researcher should not just present an interpretation and then seek validation or verification of that perspective. The search for verification is a quest that permeates the whole research process. At face value, data assembly, reduction, display and verification appear to be quite distinct and consecutive stages of data analysis. The reality is that they are iterative and totally interdependent upon each other. As the researcher assembles new data, they should be thinking of means to validate their views, asking questions in different ways of different individuals and recording these thoughts in their field notes. As data are being reduced and coded, the researcher seeks different possible explanations and evidence to support categorising, naming and connecting views in a particular manner. The researcher will question their interpretations of words and gestures and their own ways of seeing. This questioning process adds to the verification. The use of data display is a means to communicate to others the meaning and structure that a researcher 'sees' in qualitative data. The display allows others to understand that vision, to question and evaluate it. The exposure and critique of the vision by other researchers and decision-makers further verify the data. Ultimately, such critique can direct the researcher to further data assembly, reduction and display; the stages may unfold in 'waves' to produce an ultimate interpretation of great value to decision-makers.

Using computers in qualitative research and analysis

A major problem for the qualitative researcher is the sheer volume of data that they may collect. In an attempt to 'step into the shoes' of target consumers, a whole array of questions, probes, observations, answers and personal notes have to be analysed. As with quantitative data analyses, it is possible to complete analyses without the aid of a computer. Using the computer should provide speed, memory, ease of data access and the ability to transform and manipulate data in many ways. Overall it allows a much more efficient and ultimately effective process, as the researcher's effort may be focused upon generating the most effective support for decision-makers as quickly as possible rather than upon laborious administrative tasks. The following list summarises qualitative research activities that may be supported by the use of computers:

1 *Field notes*. Making notes before, during and after interviews and observations. Writing and editing these notes if needed as part of a data display to justify a particular interpretation.

2 *Transcripts.* Building up transcripts to represent the discourse in interviews.

3 *Coding.* Attaching key words to chunks of data or text.

4 *Storage, Search and Retrieval.* Keeping data in an organised manner, so that relevant segments of data or text can be located, pulled out and evaluated.

5 *Connection.* Linking relevant data segments with each other.

6 *Memoing.* Writing up reflective comments that can be 'pasted' on to relevant codes and connections.

7 *Data display.* Placing selected or reduced data in a condensed and organised format using a spreadsheet matrix or network. The display can be part of the development of the analysis or in the final vision produced by the researcher.

8 *Drawing conclusions and verification.* Aiding the researcher to interpret the data display and to test or confirm findings.

9 *Theory building.* Developing systematic and conceptually coherent explanations of findings that are meaningful to marketing decision-makers.

10 *Reporting.* Presenting interim and final reports of the findings in a written and oral manner.

Many of these tasks can be performed with readily available word-processing, spreadsheet and presentation packages. Many researchers may be very comfortable using such packages to gather and record data and to present findings. What may be new to many researchers is the use of proprietary software to help with the technical integration of data assembly, reduction, display and verification. Improvements in the functions and power of software that copes with this technical integration occur at a rapid pace. To see an array of different qualitative data analysis packages, download demo discs and evaluate how applicable they may be to a particular qualitative technique, visit **www.scolari.co.uk**. Three analysis packages presented on this site are briefly described below.

QSR Nvivo can cope with qualitative data that emerges from ethnographic approaches, focus groups, depth interviews, open-ended questions in surveys, and even from Web searches. It enables the researcher to combine the analysis of text or multimedia images. It allows for the detailed annotation of researcher notes, enabling links to memos that may have been built up in a researcher's field notebook. Coding can be visually presented with the ability to filter data or to ask questions that may reveal the existence of corroborating or contradictory evidence. It allows one to display data in matrices or in a rich text format and ultimately to present graphical models of the analysis and research findings. In essence the technical integration of data assembly, reduction, display and verification can be fully supported by this software.

BEST (Behavioural Evaluation Strategy and Taxonomy) is a software package that is more geared to ethnographic observational data. Using BEST, the researcher can record a series of live behavioural events that may occur in a mutually exclusive manner or overlap. The researcher can create categories of event that they feel should be focused upon. These categories can be observed and automatically recorded to allow quantitative assessment of frequency, duration and intervals. It can also allow qualitative assessment by being able to process and compare multiple data files that may be visual or in memos and text. It can graphically present the quantitative and qualitative interpretations and interface with other statistical and visual editing packages.

Atlas.ti: The Knowledge Workbench is similar to Nvivo in that it can cope with large bodies of textual, graphical and audio or video data generated from ethnographic approaches, focus groups and depth interviews. The overall feel of the package tries to recreate the best features of manual analysis, such as a feel for the whole dataset and a 'paper and pencil' feel of being able to make notes in margin areas

as the analysis unfolds. It enables the researcher to build networks of connected passages of text, audio and visual files with their notes or memos. The ability to maintain an overview of the whole dataset with the ability to quickly search, retrieve and browse chunks of data facilitates a clear coding process.

These descriptions do little justice to the power and array of features that these software packages possess. A full description of these would require a great deal of space and would struggle to compete with the illustrative power of the demo disk that we recommend you download and explore.

Qualitative data analysis packages do not automate the analysis process, nor is that their purpose. The process of coding, as described in the data verification section, depends upon the interpretations made by the researcher. The overall description, model or theory that emerges from the analysis also depends upon interpretations made by the researcher. No analysis package can perform such interpretations.

Qualitative data analysis is not formulaic; it requires an approach that gives quick feedback to the researcher on the results of emergent questions. This involves an iterative cycle of reflection and innovation, which means total interaction between the researcher and the computer. So, rather than seeing analysis as an automated process, the purpose of software is to aid the researcher to analyse data in a systematic and thorough manner. The researcher seeks patterns, meanings and interconnections in their qualitative data. This can be conducted manually, but by using software they can manipulate the data far more efficiently to help them see patterns, meaning and interconnections and ultimately to develop theory. In summary, software packages offer the qualitative researcher the following advantages.

Advantages of computer-assisted qualitative data analysis

1 *Speed.* The speed at which programs can carry out sorting procedures on large volumes of data is remarkable, and continues to get faster. It gives the data analyst more time to think about the meaning of data, enabling rapid feedback of the results of particular analytic ideas so that new ones can be formulated. Analysis becomes more devoted to creative and intellectual tasks, less immersed in routine.

2 *Rigour.* Rigour adds to the trust placed in research findings. In this context it means counting the number of times things occur as well as demonstrating that negative incidences have been located rather than selecting anecdotes that support a particular interpretation.

3 *Team.* In collaborative research projects where researchers need to agree on the meaning of codes, a check can easily be made of whether team members are interpreting segments in the same way. This is particularly useful as coding moves from the more descriptive and mundane codes to ones that reflect broader theoretical concerns. Researchers can pass coded interviews between them, and compare the results.

4 *Sampling.* It is easy to keep track of who *has* been interviewed, compared with the intentions of who *should* be interviewed. Beyond the sampling of individuals is the concept of theoretical sampling, i.e. the inclusion of events that corroborate or contradict developing theory. As the researcher has more time to spend on creative and intellectual tasks, they can develop stronger descriptions and theories and strengthen the validity of their views by ensuring they have sampled sufficient incidences. [16]

It must be reinforced, however, that software packages cannot interpret and find meaning in qualitative data. The programs do facilitate, and in some cases automate, the identification and coding of text. But there is sometimes a false assumption that identification and coding are simple and unproblematic, and critical evaluation and scrutiny of coded segments and code counts are not needed. By facilitating quick

analyses, which focus on quantitative category relationships, the software may discourage more time-consuming, in-depth interpretations. Thus while the programs are intended as a means of allowing the researcher to stay close to the data, their misuse can have the unintended result of distancing the researcher from the data. As discussed earlier, many decision-makers who use qualitative marketing research do not question how analysis is completed, or indeed why it should be completed. The following arguments illustrate the nature of their concerns. [17]

Disadvantages of computer-assisted qualitative data analysis

1 *Mechanistic data analysis.* The computer cannot replace the creative process expected of the qualitative researcher. The researcher can evaluate the interrelated play on particular words, the tone of voice or the gestures of a particular respondent. The sensitivity towards these relationships and connections can be lost in a mechanistic search for statements.

2 *Loss of the overview.* The researcher may be seduced into concentrating on the detail of individual chunks of data and assigning codes to the data. This focus may detract from the overall context that is so vital to identify and name chunks of data. Making sense of codes can be greatly facilitated by an ability to visualise the data in its entirety.

3 *Obsession with volume.* Given the ability to manipulate large amounts of data, there may be a push to increase the number of interviews. This may be counter-productive in that the emphasis should be on the interrelated *qualities* of:

- individual respondents
- the interview process.

4 *Exclusion of non-text data.* As noted earlier, qualitative 'text' can include notes, observations, pictures and music that make up the total 'picture' or representation of individuals. Many programs can only cope with the narrative of questions and answers recorded in transcripts.

Many software developers recognise these limitations and have gone to great pains to overcome them. One of the trade-offs faced by software developers in overcoming these limitations is the user-friendliness of their programs compared with the sophistication of being able to manipulate and represent the structure that may lie in multifarious data. As qualitative researchers use and learn how to generate the most from the software, user-friendliness may take a lesser though not ignored role. Experienced qualitative researchers can demand more sophistication to match the realities of coping with qualitative data. For the novice qualitative researcher the packages may seem daunting, but this problem is analogous to an initial exposure to sophisticated survey design packages such as SNAP or statistical packages such as SPSS. In all cases the researcher needs to appreciate how the software may serve them and work through the examples and cases, to experiment and to build up their knowledge and confidence.

➤➤➤
See Professional Perspective 20.

Go to the Companion Website and read Professional Perspective 20 'Apply yourself' by Tim Macer. Tim evaluates the effectiveness of qualitative analysis software applications with summaries of the most popular analysis software.

International marketing research

If one were to take the very naive view of qualitative data analysis being to feed data into an analysis package, and to wait for processed data to emerge, then international analysis would focus purely upon issues of language and translation. Such a perspective of qualitative analysis would ignore the context and process of collecting data and the role that context and process play in interpreting the meaning that emerges from interviews and observations. As discussed at the start of this chapter, qualitative marketing researchers need an acute self-awareness of how they 'see' – which affects the way they pose questions and interpret answers.

Consumers in any country use their social and cultural frames of reference to interpret questions posed to them by qualitative researchers and to present a response. Likewise, qualitative researchers use their social and cultural frames to present questions and interpret answers. If the researcher and the respondent share the same or similar social and cultural frames of reference, the analysis and interpretation of the data can be relatively straightforward. If the qualitative researcher goes into an international market, there is the potential for big differences in social and cultural frames between the researcher and researched. The qualitative researcher needs to develop an understanding of the social and cultural frames of the types of respondent in an international market. At the same time, they must have a strong awareness of their own social and cultural frames. Only when they have examined both perspectives can they start to interpret consumer responses.

The process is summarised by leading qualitative researchers Virginia Valentine and Malcolm Evans as:[18]

> *Consumers give a 'coded' version of the social and cultural relationship with products and brands that drive their 'feelings'. Because language (and language systems) are the medium of culture, the rules of language become the rules of the code. Qualitative research then becomes a matter of working with the code through understanding the rules of language.*

Thus, simple literal translations of transcripts of interviews from international markets entered into a qualitative data analysis package are doomed to failure. Understanding the rules of language and understanding oneself are vital for the qualitative researcher to interpret interviews and observations. As the rules of language, with the social and cultural forces that shape those rules, become more alien to the researcher in international markets, the task of analysis and meaningful interpretation becomes more difficult.

Ethics in marketing research

It is interesting to note that within the ESOMAR code of conduct, little reference is made to what is deemed as the ethical practice of data analysis, be that quantitative or qualitative. This is understandable, as the chief concern for the marketing research industry is how respondents are handled, i.e. the process of eliciting data from them. Care must be taken to ensure that the precious resource of respondents is not misled or manipulated.

With quantitative data, as will be seen in Chapters 18 to 24, there are many established and consistent procedures of analysis. With qualitative data, even though there exists a broad framework to manage analysis procedures, there does not exist a body of consistent and established procedures of analysis. The difficulty in establishing consistent procedures lies primarily in the great diversity of data that can be included in the analysis procedure. Go back to the 'Data assembly' subsection to see the list of types of qualitative data and it is easy to see why this is so. Combining researcher's notes, transcripts of interviews, pictures, audio and video recordings and mood boards does not lead to a structured process. It is a messy process that owes much to individual patience, creativity and vision.

In searching for support of ethical practice to cope with such a 'messy process', there is one area of support that comes from the Code of Conduct of the Market Research Society in Britain. In their section Mutual Rights and Responsibilities of Researchers and Clients, Rule B26 states:

> *When reporting on the results of a marketing research project the Researcher must make a clear distinction between the findings as such, the Researcher's interpretation of these and any recommendations based on them.*

The key element of this rule is that the researcher should be explicit about their interpretation of the data they have collected. This takes us back to the start when we discussed the self-reflection of the social and cultural values of the researcher. If the qualitative researcher fails or cannot be bothered to reflect upon their own values and cultural norms, their interpretation of qualitative data may be extremely biased. It therefore follows that, for the most valid as well as the most ethical interpretation of qualitative data, the researcher must continually reflect and test the extent and effect of their social and cultural values.

Summary

Qualitative marketing researchers should reflect upon how their social and cultural values affect the way they perceive and observe target respondents. These reflections should be built up as field notes as the whole process of data gathering develops and evolves. These notes form a key source of qualitative data to complement the broad array of qualitative data generated from interviews and observations. To successfully draw together a valid interpretation, qualitative data analysis must be set in the context of a theoretical understanding of the issue being researched, and an understanding of the marketing decision-makers' use of the findings.

The first stage of the process of analysing qualitative data involves assembling data in its rich and varying formats. The second stage involves reducing the data, i.e. selecting, classifying and connecting data that is believed to be of the greatest significance. A key element of this stage is the concept of coding. The third stage involves displaying data, i.e. using graphical means to display the meaning and structure that a researcher sees in the data they have collected. The final stage involves verifying the data. The marketing researcher aims to generate the most valid interpretation of the data they collect, which may be supported by existing theories or through the concept of theoretical sampling. The stages of analysis seem quite distinct but in reality they are totally interdependent upon each other.

To be able to cope with the great amount of data generated from qualitative techniques, a great variety of software packages are available. Used correctly, they can

facilitate a speedy and rigorous exploration of qualitative data, allowing teams of researchers to perform creative and incisive analyses and interpretation. The main concern with the use of qualitative data analysis packages lies in the potential for them to be mechanistic and to encourage yet more interviews to be completed, sacrificing the quality of data capture. The qualitative researcher needs to develop an understanding of the social and cultural frames of target respondents in international markets. At the same time, they must have a strong awareness of their own social and cultural frames. Only when they have examined both perspectives can they effectively interpret consumer responses. There are ethical implications of the extent to which researchers seek the valid interpretation they can make of the qualitative data they have gathered.

Questions

1. How may the social and cultural background of a researcher affect they way they:
   - gather qualitative data?
   - interpret the whole array of qualitative data they have gathered?

2. What is the significance of a qualitative researcher having a theoretical and marketing understanding of the subject they are researching?

3. Why should a qualitative researcher maintain a field notebook?

4. What should be recorded in a field notebook?

5. What may be classified as 'data' when assembling data as part of the data analysis process?

6. What does the word 'coding' mean in the context of qualitative data analysis? What problems do you see associated with the process of coding?

7. What are the advantages and disadvantages of handing over audiotapes of qualitative interviews to a typist who has taken no part in the interviews?

8. Evaluate the purpose of displaying qualitative data.

9. What advantages and disadvantages do you see in displaying qualitative data in a spreadsheet format?

10. Evaluate 'when' the stage of data verification should occur.

11. How may theoretical sampling aid the process of verification?

12. How may computers help in the whole process of qualitative data gathering and analysis?

13. Evaluate the main concerns that exist with the use of computers in qualitative data analysis.

14. Why is the researcher's understanding of their social and cultural values particularly important in international marketing research?

15. Why does the interpretation of qualitative findings have ethical implications?

Notes

1 Marakami, H., *South of the Border, West of the Sun* (London: Harvill, 1999).

2 Payne, G., 'Cinderella's getting ready for the ball', *Research Plus* (February 1993), 7.

3 Coffey, A. and Atkinson, P., *Making Sense of Qualitative Data* (London: Sage, 1996), 10–11.

4 Spradley, J.P., *The Ethnographic Interview* (New York: Holt, Rinehart & Winston, 1979).

5 Coffey, A. and Atkinson, P., *Making Sense of Qualitative Data: Complementary Research Strategie*s (Thousand Oaks, CA: Sage, 1996), 26.

6 Hubermann, A.M. and Miles, M.B., 'Data management and analysis methods' in Denzin, N.K. and Lincoln, Y.S. (eds), *Handbook of Qualitative Research* (Thousand Oaks, CA: Sage, 1994), 428–44.

7 Krippendorf, K., *Content Analysis: An Introduction to its Methodology* (Beverly Hills, CA: Sage), 1980.

8 Coffey, A. and Atkinson, P., *Making Sense of Qualitative Data: Complementary Research Strategies* (Thousand Oaks, CA: Sage, 1996), 29–30.

9 Flick, U., *An Introduction to Qualitative Research* (London: Sage, 1998), 180.

10 Strauss, A. and Corbin, J.M., *Basics of Qualitative Research: Techniques and Procedures for Dveloping Grounded Theory*, 1st edn (London: Sage, 1990), 74.

11 Strauss, A. and Corbin, J.M., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 1st edn (London: Sage, 1990), 114.

12 Flick, U., *An Introduction to Qualitative Research* (London: Sage, 1998,) 184–5.

13 Böhm, A., Legewie, H. and Muhr, T., *Kursus Textinterpretation: Grounded Theory* (Berlin: Technische Universität, Bericht aus dem IfP Atlas), 92–3, MS.

14 This process is known as triangulation. For a fuller description of this topic, set in the context of researcher bias, see: Griseri, P., *Management Knowledge: A Critical View* (Basingstoke: Palgrave, 2002), 60–78.

15 For a fuller explanation of validity in qualitative research, see: Kirk, J. and Miller, M.L., *Reliability and Validity in Qualitative Research* (Newbury Park, CA: Sage), 1986.

16 Seale, C., in Silverman, D., *Doing Qualitative Research: A Practical Handbook* (London: Sage, 2000), 161.

17 Wolfe, R.A., Gephart, R.P. and Johnson, T.E., 'Computer-facilitated qualitative data analysis: potential contributions to management research', *Journal of Management* 19(3) (Fall 1993), 637; Dembrowski, S. and Hanmer-Lloyd, S., 'Computer applications – a new road to qualitative data analysis?', *European Journal of Marketing* 29(11) (November 1995), 50.

18 Valentine, V. and Evans, M., 'The dark side of the onion: rethinking the meanings of "rational" and "emotional" responses', *Journal of the Market Research Society* 35(2) (April 1993), 127.

Survey and quantitative observation techniques

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

**Stage 4
Fieldwork or data collection**

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Objectives

After reading this chapter, you should be able to:

1 discuss and classify survey techniques available to marketing researchers, and describe various survey techniques;

2 identify the criteria for evaluating survey techniques, compare the different techniques and evaluate which is the best for a particular research project;

3 explain and classify the different quantitative observation techniques;

4 identify the criteria for evaluating observation techniques, compare the different techniques, and evaluate which are suited for a particular research project;

5 describe the relative advantages and disadvantages of observation techniques and compare them with survey techniques;

6 discuss the considerations involved in implementing surveys and observation techniques in an international setting;

7 understand the ethical issues involved in conducting survey and observational research.

Know exactly what you want to measure – and then select a survey or observation technique that creates cooperative respondents, willing to think and be honest.

Overview

In this chapter, we focus on the major techniques employed in descriptive research designs: surveys and quantitative observation. As explained in Chapter 3, descriptive research has as its prime objective the description of something, usually consumer or market characteristics. Survey and quantitative observation techniques are vital techniques in descriptive research designs. Survey techniques may be classified by mode of administration as traditional telephone interviews, computer-assisted telephone interviews, personal in-home or office interviews, street interviews, computer-assisted personal interviews, postal surveys, electronic surveys and mail panels. We describe each of these techniques and present a comparative evaluation of all the survey techniques. Then we consider the major observational techniques: personal observation including mystery shopping research, electronic observation, audit, content analysis and trace analysis. The relative advantages and disadvantages of observation over survey techniques and the considerations involved in conducting survey and observation research when researching international markets are also discussed. Several ethical issues that arise in survey research and observation techniques are identified.

To begin our discussion, we present an example of how the survey process can benefit from technological advances. Computer-assisted personal interviewing can make the interview process interesting for respondents and at the same time elicit more knowledge from them to support an array of marketing decisions.

example

More muscle from microchips[1]

Multimedia CAPI (Computer Assisted Personal Interviewing) builds on the technology used in conventional CAPI research but uses multimedia notebook computers, not only to collect data, but also to present planned marketing campaigns to respondents in their homes. Clips from TV advertisements can be shown on the screen and followed up with on-screen questionnaires, which the respondent answers by clicking on the relevant box or, where an opinion is being sought, by speaking into the notebook's microphone. The spoken answers are saved and can be downloaded on to an audiotape, so that other researchers can listen to them and analyse them at a later date.

The system is not limited to showing TV ad clips. Respondents can be shown different packaging treatments, or a piece of the client company's marketing communications, alongside competitors' material. Notebook computers can also help interviewers overcome language barriers. The questionnaire can be translated beforehand into any language and played back to the respondents. ■

Survey techniques

Survey techniques
Techniques based upon the use of structured questionnaires given to a sample of a population.

Survey techniques are based upon the use of structured questionnaires given to a sample of a population. Respondents may be asked a variety of questions regarding their behaviour, intentions, attitudes, awareness, motivations, and demographic and lifestyle characteristics. These questions may be asked verbally, in writing or via a computer, and the responses may be obtained in any of these forms. 'Structured' here refers to the degree of standardisation imposed on the data collection process. In **structured data collection**, a formal questionnaire is prepared and the questions are asked in a prearranged order; thus, the process is also direct. Whether research is classified as direct or indirect is based on whether the true purpose is known to the respondents. As explained in Chapter 7, a direct approach is undisguised in that the purpose of the project is disclosed to the respondents or is otherwise obvious to them from the questions asked.

Structured data collection
Use of a formal questionnaire that presents questions in a prearranged order.

In a typical questionnaire, most questions are fixed-response alternative questions that require the respondent to select from a predetermined set of responses. Consider, for example, the following question, designed to measure a dimension of students' attitudes towards the way they are assessed in marketing research classes:

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
I prefer written examinations compared with continual assessment	☐	☐	☐	☐	☐

The survey method has several advantages. First, the questionnaire is simple to administer. Second, the data obtained are consistent because the responses are limited to the alternatives stated. The use of fixed-response questions reduces the variability in the results that may be caused by differences in interviewers. Finally, coding, analysis and interpretation of data are relatively simple.

Disadvantages are that respondents may be unable or unwilling to provide the desired information. For example, consider questions about motivational factors. Respondents may not be consciously aware of their motives for choosing specific brands or shopping at particular stores. Therefore, they may be unable to provide accurate answers to questions about their motives. Respondents may be unwilling to respond if the information requested is sensitive or personal. In addition, structured questions and **fixed-response alternative questions** may result in loss of validity for certain types of data such as beliefs and feelings. Finally, wording questions properly is not easy (see Chapter 13 on questionnaire design). In other words, the survey imposes the language and logic of the researcher on to questionnaire respondents. Given this core characteristic of survey techniques, great care must be taken to ensure that the language and logic used in questionnaires are meaningful and valid to potential respondents. Despite the above disadvantages, the survey approach is by far the most common method of primary data collection in marketing research.

Survey questionnaires may be administered in three major modes: (1) telephone interviews, (2) personal interviews, and (3) mail interviews (see Figure 10.1). Telephone interviews may be further classified as traditional telephone interviews and computer-assisted telephone interviews. Personal interviews may be conducted in the home or office, as street interviews, or as computer-assisted personal interviews. The third major method, mail interviews, takes the form of traditional mail surveys, electronic mail surveys or surveys conducted using mail panels. We now describe each method.

Fixed-response alternative questions
Questions that require respondents to choose from a set of predetermined answers.
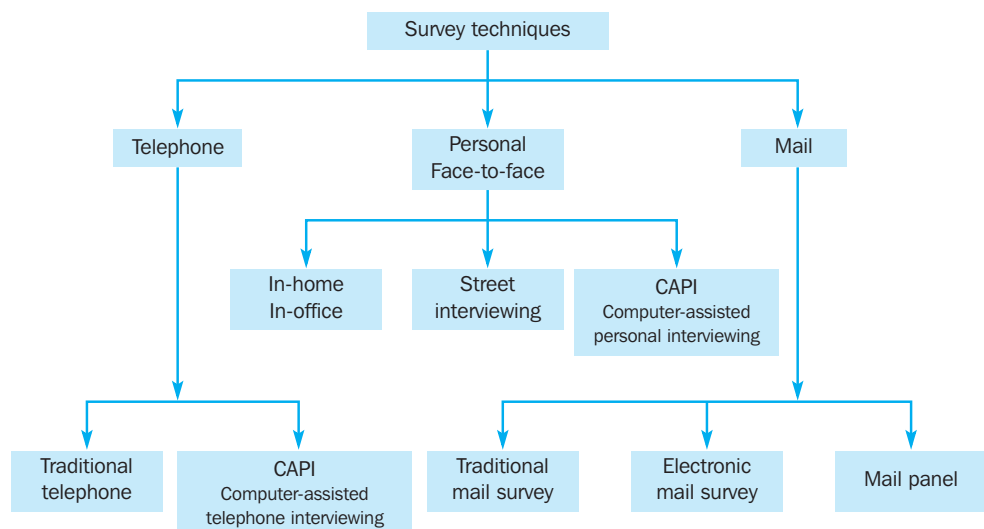

Figure 10.1
A classification of survey techniques

Telephone interviews

As stated earlier, telephone interviews may be categorised as traditional or computer-assisted.

Traditional telephone interviews

Traditional telephone interviews involve phoning a sample of respondents and asking them a series of questions. The interviewer uses a paper questionnaire and records the responses with a pencil. From a central location, a wide geographical area can be covered, including international markets. Given that telephone interviewers cannot give respondents any visual prompts, they have to write down answers to any open-ended questions and may have to flick through the questionnaire to find appropriate questions for a particular respondent (filtering). These interviews tend to be short in duration and have questions with few options as answers. Today, this approach is rarely used in commercial marketing research, the common approach being a computer-assisted telephone interview.

Computer-assisted telephone interviews (CATI)

Computer-assisted telephone interviewing (CATI) uses a computerised questionnaire administered to respondents over the telephone. A computerised questionnaire may be administered using a networked computer or a personal computer. The interviewer sits in front of a terminal and wears a small headset. The terminal replaces a paper and pencil questionnaire, and the headset substitutes for a telephone. Upon command, the computer dials the telephone number to be called. When contact is made, the interviewer reads questions posed on the screen and records the respondent's answers directly into the computer memory bank, ready for immediate analysis. The following example illustrates how the technique has developed over the years. The main benefit of the technique illustrated here is the speed of collecting data and analyses. Speed is of the essence in subjects where respondent attitudes and behaviour can change quickly, an example being toys that parents may wish to buy for their children. One of the most widespread uses of CATI is for political opinion polls.

e x a m p l e

The fans of the phone pump the numbers up [2]

In the 1970s, the telephone simply was not adequate as a major research tool: there just weren't enough names in the phone book to guarantee workable results. Now it couldn't be more different: year on year, telephone proliferates. The areas of market research for which it is deemed unsuitable grow smaller.

The commercial world conducts business over the phone (no accident that telephone research thrives in business-to-business research); we are phone fluent now, at ease communicating down a phone line. By contrast, it is also an age of no-go areas, security fences and entryphones, a fragmented culture where people are loath to open doors of their homes to a stranger. Small wonder that traditional face-to-face interviewing is losing out.

Technology's most precious gift to the telephone, with features like auto dialling, is speed of delivery. The pace of collection and the immediacy of the data is more than a selling point. 'It's going to happen now, tonight and that's it,' enthuses Virginia Monk, who manages Network Research's telephone department. 'That's what I love about it: it's so instantaneous, so quick. I love the speed of it, to have the data come in, to play with it.'

Clients at Network are allowed in to witness the research at first hand. Potential problems are dealt with almost straight away. Quotas and standards are checked there and then; call rates and refusal rates are constantly monitored. In comparison, traditional street and home interviewing seems fretful and wearing: it might be hours before a problem becomes known; in a telephone centre they can spot it the minute it pokes its head above ground. ■

Go to the Companion Website and read Professional Perspective 10 by Virginia Monk. Virginia's article 'Ringing home' details the applications, benefits and limitations of telephone interviewing.

The computer systematically guides the interviewer. Only one question at a time appears on the screen. The computer checks the responses for appropriateness and consistency. It uses the responses as they are obtained to personalise the questionnaire. The data collection flows naturally and smoothly. Interviewing time is reduced, data quality is enhanced, and the laborious steps in the data collection process, coding questionnaires and entering data into the computer are eliminated. Because the responses are entered directly into the computer, interim and update reports on data collection or results can be provided almost instantaneously. The following example shows how technological developments allow flexibility in interviewing different types of respondent throughout Europe. Flexibility means that different responses to a particular question can lead to different routes and sets of questions for particular types of respondent.

| example |

A pan-European view of the executive at lunch[3]

The general perception of CATI is, in the main, one of large sample size and reasonably short, simple, pre-coded questionnaires, conducted during a short fieldwork period. As computers become much more powerful and updated software becomes available, it is apparent that CATI surveys can be approached in new ways. Its abilities to deal with large sample sizes, complicated quota structures and complex, long, modular questionnaires enables us to deal with surveys which once would have caused untold sleepless nights.

In a survey of European executives, CATI was used and coped well with a long and complex questionnaire. The interview covered respondents' exposure to newspapers, magazines, television (terrestrial, satellite and cable) and radio, at any time on any day. To cover the array of media exposure options throughout Europe, the questionnaire needed a comprehensive, pre-coded list of in excess of 100 possibilities. On the first fieldwork day it became apparent that in mainland Europe the borders were already down. Many Belgian, Dutch, French and German businessmen regularly listened to and watched each other's channels. This meant that a European media list was built up and was added to each national list. The CATI system also had to select three publications – the sponsoring organisation's and two others – to be rated on 14 rotated statements according to frequency of readership and whether they were international or not.

The survey not only collected comprehensive and accurate data, it was greatly enjoyed by respondents and interviewers alike, even though the interview length ranged from 20 minutes to a full hour, depending upon the executive's habits. The respondents could see the point and therefore the interviewers could probe and delve while enjoying the responses. 'High Fliers', after all, are not often quizzed about their breakfast, lunch and evening activities – and coding departments see what they get up to even less! ▪

The biggest drawback of CATI lies in respondents' willingness to be interviewed by telephone. At times the technique may be confused in the minds of respondents with cold-call selling techniques. To overcome this problem requires interviewers who are trained to reassure respondents and to make the interviewing experience worthwhile. In most businesses, phone calls are screened; how else could managers avoid calls from companies and individuals they do not want to talk to? Many businesses have set up formal procedures to minimise unnecessary calls to management. This is a hurdle well-trained interviewers are taught to circumvent. Experienced interviewers can 'run the gauntlet' of three or four screenings and still secure an interview without in any way upsetting the interviewee.

The usual blocking responses of '(s)he's busy', '(s)he wouldn't be interested', '(s)he doesn't deal in that area', or '(s)he's in a meeting', have been joined by 'we don't allow

marketing research interviews – it's company policy'. Even this will not deter telephone interviewers from seeking interviews but it is a continuing and worrying trend.[4]

So why do companies which commission marketing research themselves restrict their own employees from participating? The reasons given are:

■ The confidentiality of information divulged to interviewers
■ The length of interviews
■ The frequent apparent or real irrelevance of questions
■ The number of requests
■ *Taking part is seen as providing no direct benefit to the company.*

The last point is set in italic as a reminder to all researchers to place themselves in the shoes of potential respondents. The marketing researcher, and the decision-maker they support, may have distinct information requirements, but why should respondents supply this information? What benefits do they gain from giving this information? This issue will be dealt with in examining questionnaire design in Chapter 13, but here we have a reminder of the impact of ignoring the benefits that respondents gain from participating in a survey.

Personal interviews

Omnibus survey
A distinctive form of survey that serves the needs of a syndicated group. The omnibus survey targets particular types of respondents such as those in specific geographic locations, e.g. Luxembourg residents, or consumers of particular types of products, e.g. business air travellers. With that target group of respondents, a core set of questions can be asked, with other questions added as syndicate members wish.

Personal interviewing techniques may be categorised as in-home, in-office, street or computer-assisted.

Personal in-home and in-office interviews

In personal in-home interviews, respondents are interviewed face-to-face in their homes or in their workplace. The interviewer's task is to contact the respondents, ask the questions and record the responses. In recent years, the use of personal in-home interviews has declined due to their high cost. Nevertheless, they are still used, particularly by syndicated firms, as illustrated in the following example. The use of **omnibus surveys** for marketing decision-makers is shown with examples of the types of marketing research projects that are supported by this technique.

example

Omnibus – the flexible research tool for many jobs[5]

The tracking aspect of an omnibus is important for a number of clients, not least because it enables the following:

■ A continuous assessment of awareness and effectiveness of marketing/advertising activity
■ A monitor of opinion/behaviour in constantly changing competitive markets, which in addition may be highly fragmented
■ The accumulation of ongoing data to provide robust bases and thus build, for example, profiles of customers that would otherwise be inaccessible for small brands
■ Access to competitors' performance, without introducing overt bias among respondents' answers
■ Detection of early signs of change in a defined marketplace
■ Opportunities for syndication (with additional benefits, such as shared costs).

Omnimas offers omnibus facilities, each week interviewing 2,000 adults face-to-face in their homes. A number of continuous tracking projects are conducted that range through the following:

■ Advertising awareness tracking (which, depending on the client and subject matter, is conducted on a weekly, monthly or less frequent basis), monitoring whether the advertisement has been seen or heard and how it is communicated or interpreted

- Regular (weekly) behavioural monitors (the Eating Out Of Home Monitor examines consumers' eating habits outside their own homes in restaurants and similar outlets)
- Ongoing tracking of people's awareness, usage and participation in loyalty schemes across a number of retail sectors
- Using ongoing information about consumers' intentions to buy video titles to predict volume sales (data are supplied for Buena Vista Entertainment, the Disney Video Company). ■

In-office research is used extensively in business-to-business research to research subjects who cannot be effectively interviewed by telephone or mail. Managers being interviewed have the comfort and security of their office and can control the timing and pace of the interview. For the researcher, the big benefit of meeting managers in their office is the ability to build up a rapport, probe and gain the full attention of the manager.

Street interviews

In street interviews, respondents are intercepted while they are shopping in town centres or shopping centres. They may be questioned there and then in the street or taken to a specific test facility. In the testing of new product formulations, test facilities are ideal to allow respondents the time and context to sample and evaluate products. The technique can also be used to test merchandising ideas, advertisements and other forms of marketing communications. The big advantage of the street interview is that it is more efficient for the respondent to come to the interviewer than for the interviewer to go to the respondent.[6]

Computer-assisted personal interviews (CAPI)

In computer-assisted personal interviewing, the third form of personal interviewing, the respondent sits in front of a computer terminal and answers a questionnaire on the screen by using the keyboard or a mouse. There are several user-friendly electronic packages that design relatively simple questions for the respondent to understand. Help screens and courteous error messages are also provided. The
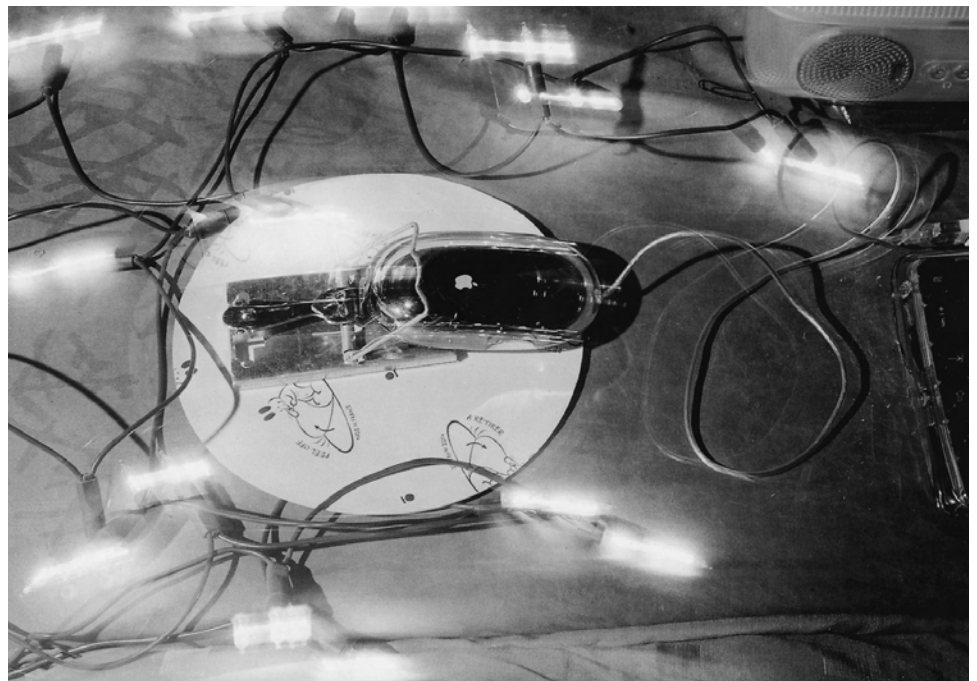
A courteous error message lets the CAPI respondent know that they've made a mistake.

colourful screens and on- and off-screen stimuli add to the respondent's interest and involvement in the task. This method has been classified as a personal interview technique because an interviewer is usually present to serve as a host or hostess and to guide the respondent as needed.

CAPI has been used to collect data at test facilities from street interviews, product clinics, conferences and trade shows. It may also be used for in-home or in-office interviews. You may wonder, however, how CAPI compares with the traditional method of conducting personal interviews, using paper-and-pencil questionnaires. The following example illustrates the benefits of CAPI compared with traditional interviewing.

<table>
<tr><td>

e x a m p l e

</td><td>

How the omnibus hit the fast track – and now, hold very tight . . .[7]

The introduction of the CAPI overnight transformed the capabilities of the personal omnibus survey. The speed of delivering results improved dramatically, from three to four weeks, which was the norm for a paper survey, down to seven to eight days for a CAPI omnibus. There were also considerable improvements in data quality obtained by eliminating routing errors, providing the ability to build edits into questionnaire scripts and introducing more sophisticated field management systems.

The major impact of CAPI lay in the facility to produce and store complex and highly filtered questionnaire scripts, which transformed the scope of the omnibus as a data collection vehicle. For example, it became possible and cost-efficient to hold complex batteries of questions, applying only to a small minority group, which are triggered when a target informant is identified. ◼

</td></tr>
</table>

A major development for marketers, especially in financial services, has been the use of customer satisfaction surveys to guide strategic and operational decisions. With traditional interview techniques, the interviewer may have to carry a huge questionnaire to cope with questions that measure attitudes to a range of banks and a range of services taken from those banks. With CAPI, when a particular bank is chosen, particular questions may be filtered out, and choosing a particular service from that bank can filter out further questions. Questions specific to the respondent may then be asked, in all making the interview process far more efficient.

Mail interviews

Mail interviews, the third major form of survey administration, can be conducted via traditional mail, mail panel or electronic mail.

Traditional mail interviews

In the traditional mail interview, questionnaires are mailed to preselected potential respondents. A typical mail interview package consists of the outgoing envelope, cover letter, questionnaire, return envelope, and possibly an incentive.[8] The respondents complete and return the questionnaires. There is no verbal interaction between the researcher and the respondent in the interview process.[9] There may be an initial contact with potential respondents, to establish who is the correct person to send the questionnaire to, and to motivate them before they receive the survey.

Before data collection can begin, however, the respondents need to be at least broadly identified. Therefore, an initial task is to obtain a valid mailing list. Mailing lists can be compiled from telephone directories, customer databases or association membership databases, or can be purchased from publication subscription lists or commercial mailing list companies.[10] To experience an online selection and purchase of lists, allowing very specific demographic and lifestyle characteristics of target

respondents to be selected, see **www.prospectlocator.com.** Regardless of its source, a mailing list should be current and closely related to the population of interest. (Chapters 13 and 14 will detail the full questionnaire design and sampling implications of this approach.) With an understanding of characteristics of target respondents and what will motivate them to respond honestly, fully and as quickly as possible, the researcher must also make decisions about the various elements of the mail interview package (see Table 10.1).

Table 10.1 **Some decisions related to the mail survey package**

Outgoing envelope	Method of addressing	Envelope size, colour	Postage	
Covering letter	Personalisation	Sponsorship	Type of appeal	Signature
Questionnaire	Size, length and content	Colour and layout	Format and reproduction	Respondent anonymity
Instructions	As part of covering letter	A separate sheet	Alongside individual questions	
Return envelope	Whether to include one	Type of envelope	Postage	
Incentives	Feedback of findings	Monetary vs. non-monetary	Prepaid vs. promised amount	

The GlobalCash Project had a long and complex questionnaire which was administered by mail throughout Europe.

A postal survey throughout Europe

The GlobalCash questionnaire was sent by mail to Treasury and Cash Managers in Europe's largest corporations. It covered many complex and technical issues, which had to be carefully translated into 10 European languages. In Switzerland, for example, French, German, Italian or English versions were available, depending upon the company and region targeted. Before questionnaires were sent out, telephone calls were made to every company in the sampling frame. These ensured that targeted respondents were correctly identified, all other postal details were correct and the purpose and benefits to the respondent were clear. A second telephone call was made, two weeks after the questionnaire was sent out, to further encourage a positive response. Given the wide geographic spread, the long and technical nature of the subject and little pressure to quickly generate analyses for the whole of Europe, the mail survey was the ideal survey technique. ■

Mail panels

Mail panels
A large and nationally representative sample of households that have agreed to participate in periodic mail questionnaires, product tests and telephone surveys.

Mail panels were introduced in Chapters 3 and 4. A **mail panel** consists of a large, nationally representative sample of households that have agreed to participate in periodic mail questionnaires, product tests and telephone surveys. The households are compensated with various incentives. Mail panels can be used to obtain information from the same respondents repeatedly. Thus, they can be used to implement a longitudinal design.

Electronic mail

Electronic mail can be broken down into email and Internet interviews. To conduct a survey by email, a list of email addresses needs to be obtained. The survey is written within the body of the email message and sent to respondents. Email surveys use pure text (ASCII) to represent questionnaires, and can be received and responded to by anyone with an email address, whether or not they have access to the Web.

Respondents type their answers to either closed-ended or open-ended questions at designated places, and click on 'reply'. Responses are then entered into an analysis package and tabulated. Alternatively, a program can be written that interprets the emailed responses and reads the answers directly into a format compatible with the requirements of an analysis package.

Text-based email surveys are convenient for the respondent because they require no facilities or expertise beyond those that they use in their day-to-day email communication. Email surveys do have several limitations, chief of which is that they can appear dry and uninteresting.[11] Given the technical limitations of most email systems, questionnaires cannot utilise programmed filters or skip-patterns, logic checks or randomisation. The limited intelligence of ASCII text cannot keep a respondent from, say, choosing both 'yes' and 'no' to a question where only one response is meaningful. Filter or skipping instructions (e.g. 'If the answer to question 8 is 'Yes', go to question 16) must appear explicitly, just as on paper. These factors can reduce the quality of data from an email survey and can require post-survey data cleaning.[12]

Hypertext markup language (HTML)
The language of the Web.

In contrast to email surveys, Internet or Web surveys use **hypertext markup language** (HTML), the language of the Web, and are posted on a Website. Respondents may be recruited online from potential respondent databases maintained by the marketing research firm or they can be recruited by conventional techniques (mail, telephone). Respondents are asked to go to a particular Web location to complete the survey. Frequently, respondents are not recruited, but those who happen to be visiting the Website on which the survey is posted (or other popular Websites) are invited to participate in the survey. Either all or every ith Web visitor is allowed to participate. Web surveys offer several advantages over email surveys. It is possible in HTML, but not in ASCII text, to construct buttons, check boxes and data entry fields that prevent respondents from selecting more than one response where only one is intended, or from otherwise typing where no response is required. Skip patterns can be programmed and performed automatically as in CATI or CAPI. It is possible to validate responses as they are entered. Finally, additional survey stimuli such as graphs, images, animations and links to other Web pages may be integrated into or around the survey. The responses are collected in an adjoining database. The data require some processing before they can be tabulated or used in a statistical package. All these factors contribute to a higher-quality experience and elicitation process. With the growth of the use of email and the Internet as means of communication, electronic mail is rapidly becoming the most feasible and popular means of conducting surveys. Most survey analysis packages that include the function of questionnaire design have the ability to formulate the questionnaire into email and Internet formats. Big sampling problems exist with electronic mail, which will be tackled in Chapter 14. The speed of administering the survey, collecting and analysing the data, linked with very low costs, make electronic mail very attractive.

The following lists summarise the advantages and disadvantages of Internet surveys:[13]

Advantages

■ *Speed*. Compared with a postal survey, especially on an international basis, the time taken can be reduced to a matter of days rather than weeks. Even if one includes the time taken to contact respondents by email, to establish their willingness to take part in a survey, for them to reply, for the survey to be sent, for it to be completed and then emailed back (the procedure adopted to reduce the perception of 'junk' email), such a survey can be completed far more quickly than a postal procedure. There can be a 'seamless' international coordination that using traditional mail would be very time-consuming to organise.

- *Cost.* Once the electronic questionnaire is set up, it is almost as easy to mail it to 10,000 people as to 10, since there are no printing, stationery and postage costs.
- *Quality of response.* Quality can be measured by the number and clarity of responses to open-ended questions. Graphics can be displayed to maintain the interest of respondents but also to put them in a frame of mind that elicits more from them.
- *Interviewer bias removed.* The method maintains the key advantage of mail surveys in being able to present a consistent form of measurement.
- *Data quality.* Logic and validity checks can be built in, and in areas where open-ended or 'other – please state' responses are required, the respondent types in answers ready for analysis.
- *Contacting certain target groups.* Such groups are mostly those who are regular users of the Internet, and certain business target markets.

Disadvantages

- *Sampling frames.* Where are the lists of email numbers? Unlike telephone numbers, there is no published list of email addresses. This is being slowly rectified by the addition of email addresses to established directories used as sampling frames. Another sampling issue is that the questions are self-completion based and the researcher does not know whether those who choose to take part are really representative of a target population. Being able to generalise findings to a particular population becomes difficult with a self-selecting sample.
- *Access to the Web.* At present the penetration of households and businesses is low, though it is growing rapidly, and in businesses the Web is rapidly being adopted as the medium of communication.
- *Technical problems.* Depending upon the hardware and software that respondents use, the questionnaire may 'work' or not as was intended by the designer.

➤➤➤
See Professional Perspective 6.

Go to the Companion Website and read Professional Perspective 6 by Jeff Miller. Jeff's article 'Net vs. phone: the great debate' presents research conducted by his company, Burke Interactive. It describes the relative benefits and applications of Internet surveys compared with telephone interviews.

A comparative evaluation of survey techniques

Not all survey techniques are appropriate in a given situation. Therefore, the researcher should conduct a comparative evaluation to determine which techniques are appropriate. Table 10.2 compares the different survey techniques through a range of criteria. For any particular research project, the relative importance attached to these criteria will vary. The criteria consist of flexibility of data collection, diversity of questions, use of physical stimuli, sample control, control of the data collection environment, control of field force, quantity of data, response rate, perceived respondent anonymity, social desirability, obtaining sensitive information, potential for interviewer bias, potential to probe respondents, potential to build rapport, speed and cost.

Flexibility of data collection

The flexibility of data collection is determined primarily by the extent to which the respondent can interact with the interviewer and the survey questionnaire. The personal interview, whether conducted as an in-home, in-office or street interview, allows the highest flexibility of data collection. Because the respondent and the interviewer meet face to face, the interviewer can administer complex questionnaires, explain and clarify difficult questions, and even use unstructured techniques.

Table 10.2 A comparative evaluation of survey techniques

	Telephone CATI	In-home In-office interviews	Street interviews	CAPI	Traditional mail surveys	Mail panels	Email	Internet
Flexibility of data collection	Moderate to high	High	High	Moderate to high	Low	Low	Low	Moderate to high
Diversity of questions	Low	High	High	High	Moderate	Moderate	Moderate	Moderate to high
Use of physical stimuli	Low	Moderate to high	High	High	Moderate	Moderate	Low	Moderate
Sample control	Moderate to high	Potentially high	Moderate	Moderate	Low	Moderate to high	Low	Low to moderate
Control of data collection environment	Moderate	Moderate to high	High	High	Low	Low	Low	Low
Control of field force	Moderate	Low	Moderate	Moderate	High	High	High	High
Quantity of data	Low	High	Moderate	Moderate	Moderate	High	Moderate	Moderate
Response rate	Moderate	High	High	High	Low	Moderate	Low	Low
Perceived respondent anonymity	Moderate	Low	Low	Low	High	High	Moderate	High
Social desirability	Moderate	Low to moderate	Low	Low to high	High	High	High	High
Obtaining sensitive information	Low	High	Low	Moderate	Moderate	Moderate	Moderate	Moderate
Potential for interviewer bias	Moderate	High	High	Low	None	None	None	None
Potential to probe respondents	Low	High	Moderate	Moderate	Low	Low	Low	Low
Potential to build rapport	Moderate	High	Moderate to high	Moderate to high	Low	Low	Low	Low
Speed	High	Moderate	Moderate to high	Moderate to high	Low	Low to high	High	High
Cost	Moderate	High	Moderate to high	Moderate to high	Low	Low to moderate	Low	Low

By contrast, the traditional telephone interview allows only moderate flexibility because it is more difficult to use unstructured techniques, ask complex questions, or obtain in-depth answers to open-ended questions over the telephone. CATI and CAPI and Internet surveys allow somewhat greater flexibility because the researcher can use various question formats, can personalise the questionnaire, and can handle complex skip or filter patterns (directions for skipping questions in the questionnaire based on

the subject's responses). Because the electronic and mail questionnaire allows for no interaction between the interviewer and the respondent, mail surveys, mail panels and email surveys have low flexibility.

An often-overlooked benefit of Internet survey research is the ease with which an Internet survey can be quickly modified. For example, early data returns may suggest additional questions that should be asked. Changing or adding questions as the need becomes apparent would be almost impossible with a traditional mail survey, possible but difficult with personal or telephone questionnaires, but achievable in a matter of minutes with some Internet survey systems.

Diversity of questions

The diversity of questions that can be asked in a survey depends on the degree of interaction the respondent has with the interviewer and the questionnaire, as well as the respondent's ability to actually see the questions. A variety of questions can be asked in a personal interview because the respondent can see the questionnaire and the interviewer is present to clarify ambiguities. Thus in-home and in-office interviews, street interviews and CAPI allow for diversity. In Internet surveys, multimedia capabilities can be utilised and so the ability to ask a diversity of questions is moderate to high, despite the absence of an interviewer. In mail surveys, mail panels and email surveys, less diversity is possible. In traditional telephone interviews and CATI, the respondent cannot see the questions while answering, and this limits the diversity of questions. For example, in a telephone interview or CATI, it would be very difficult to ask respondents to rank 15 television programmes in terms of preference.

Use of physical stimuli

Often it is helpful or necessary to use physical stimuli such as products, product prototypes, commercials or promotional displays during an interview. For the most basic example, a taste test involves tasting the product and evaluating the taste. In other cases, photographs, maps or other audio-visual cues are helpful. In these cases, personal interviews conducted at central locations (guided through street interviews and CAPI) are more preferable to in-home interviews. In the central location, many intricate visual stimuli can be set up prior to the actual interview. Mail surveys and mail panels are moderate on this dimension, because sometimes it is possible to mail the facilitating aids or even product samples. Internet surveys are also moderately suitable. Because they are Web-based, the questionnaires can include multimedia elements such as prototype Web pages and advertisements. The use of physical stimuli is limited in traditional telephone interviews and CATI, as well as in email surveys (depending upon the respondent's ability to open attachments).

Sample control

Sample control
The ability of the survey mode to reach the units specified in the sample effectively and efficiently.

Sample control is the ability of the survey mode to reach the units specified in the sample effectively and efficiently.[14] At least in principle, in-home and in-office personal interviews offer the best sample control. It is possible to control which sampling units are interviewed, who is interviewed, the degree of participation of other members of the household, and many other aspects of data collection. In practice, to achieve a high degree of control, the researcher has to overcome several problems. It is difficult to find respondents at home during the day because many people work outside the home. Also, for safety reasons, interviewers are reluctant to venture into certain neighbourhoods and people have become cautious of responding to strangers at their door. Street interviews allow only a moderate degree of sample control. Although the interviewer has control over which respondents to intercept, the choice

is limited to individuals who are walking down a street or through a shopping centre, and frequent shoppers have a greater probability of being included. Also, potential respondents can intentionally avoid or initiate contact with the interviewer. Compared with street interviews, CAPI offers slightly better control, as sampling quotas can be set and respondents randomised automatically.

Moderate to high sampling control can be achieved with traditional telephone interviews and CATI. Telephones offer access to geographically dispersed respondents and hard-to-reach areas. These procedures depend upon a **sampling frame**, a list of population units with their telephone numbers.[15] The sampling frames normally used are telephone directories, but telephone directories are limited in that:

Sampling frame
A representation of the elements of the target population that consists of a list or set of directions for identifying the target population.

1 Not everyone has a phone, while some individuals have several phone numbers partly due to the growing use of mobile phones.
2 Some individuals have unlisted phones or are ex-directory.
3 Directories do not reflect new phones in service or recently disconnected phones.

Internet, email and traditional mail surveys require a list of addresses of individuals or households eligible for inclusion in the sample. Mail surveys can reach geographically dispersed respondents and hard-to-reach areas. However, mailing lists are sometimes unavailable, outdated or incomplete, especially for electronic addresses. Another factor outside the researcher's control is whether the questionnaire is answered and who answers it. Some subjects refuse to respond because of lack of interest or motivation; others cannot respond because they are illiterate.[16] Given these reasons, the degree of sample control in electronic and mail surveys is low.[17]

Mail panels, on the other hand, provide moderate to high control over the sample. They can provide samples matched to national census statistics on key demographic variables. It is also possible to identify specific user groups within a panel and to direct the survey to households with specific characteristics. Specific members of households in the panel can be questioned. Finally, low-incidence groups, groups that occur infrequently in the population, can be reached with panels, but there is a question of the extent to which a panel can be considered representative of the entire population.

Not all populations are candidates for Internet survey research. Although respondents can be screened to meet qualifying criteria and quotas imposed, the ability to meet quotas is limited by the number and characteristics of respondents who visit the Website. However, there are some exceptions to this broad statement. For example, computer products purchasers and users of Internet services are ideal populations. Business and professional users of Internet services are also an excellent population to reach with Internet surveys. Sample control is low to moderate for Internet surveys, whilst email surveys suffer from many of the limitations of mail surveys and thus offer low sample control.

Control of the data collection environment

The context in which a questionnaire is completed can affect the way that a respondent answers questions. An example of this would be the amount of distraction from other people around, noise and temperature. The degree of control a researcher has over the context or environment in which the respondent answers the questionnaire differentiates the various survey modes. Personal interviews conducted at central locations (from street interviews and CAPI) offer the greatest degree of environmental control. For example, the researcher can set up a special facility for demonstrating a product upon which a survey is based. In-home and in-office personal interviews offer moderate to high control because the interviewer is present. Traditional tele-

phone interviews and CATI offer moderate control. The interviewer cannot see the environment in which the interview is being conducted, but he or she can sense the background conditions and encourage the respondent to be attentive and involved. In mail surveys, panels, email and Internet surveys, the researcher has little or no control over the environment.

Control of field force

The **field force** consists of interviewers and supervisors involved in data collection. Because they require no such personnel, mail surveys, mail panels, email and Internet surveys eliminate field force problems. Traditional telephone interviews, CATI, street interviews and CAPI all offer moderate degrees of control because the interviews are conducted at a central location, making supervision relatively simple. In-home and in-office personal interviews are problematic in this respect. Because many interviewers work in many different locations, continual supervision is impractical.[18]

Quantity of data

In-home and in-office personal interviews allow the researcher to collect large amounts of data. The social relationship between the interviewer and the respondent, as well as the home or office environment, motivates the respondent to spend more time in the interview. Less effort is required of the respondent in a personal interview than in a telephone or mail interview. The interviewer records answers to open-ended questions and provides visual aids to help with lengthy and complex scales. Some personal interviews last for as long as 75 minutes. In contrast to in-home and in-office interviews, street interviews and CAPI provide only moderate amounts of data. Because these interviews are conducted in shopping centres and other central locations, a respondent's time is more limited. Typically, the interview time is 20 minutes or less.

Mail surveys also yield moderate amounts of data. Fairly long questionnaires can be used because short questionnaires do not necessarily generate higher response rates than long ones. The same is true for email and Internet surveys, although the Internet is a better medium in this respect. Mail panels, on the other hand, can generate large amounts of data because of the special relationship between the panel members and the sponsoring organisation.

Traditional telephone interviews and CATI result in the most limited quantities of data. They tend to be shorter than other surveys because respondents can easily terminate the telephone conversation at their own discretion. These interviews commonly last about 15 minutes, although longer interviews may be conducted when the subject matter is of interest to the respondents.[19]

Response rate

Survey **response rate** is broadly defined as the percentage of the total attempted interviews that are completed. Personal, in-home and in-office, street and CAPI yield the highest response rates. Problems caused by 'not-at-homes' can often be resolved by calling back at different times. Telephone interviews, traditional and CATI, also suffer from not-at-homes or no-answers. Higher response rates are obtained by call-backs. Many telephone surveys attempt to call back at least three times.

Mail surveys have the poorest response rate. In a mail survey of randomly selected respondents, without any pre- or postmailing contact, response rates can be less than 15%. Such low response rates can lead to serious bias (**non-response bias**). This is because whether a person responds to a mail survey is related to how well the benefits of taking part in the survey are meaningful to the respondent and are clearly communicated to them. The magnitude of non-response bias increases as the response rate

decreases. Response rates in mail panels are much higher than in traditional mail techniques because of assured respondent cooperation.

Internet surveys can have very poor response rates, much depending upon how comprehensive and current the sampling frame is. With a well-constructed sampling frame that is relevant to the survey topic, response rates on the Internet can be relatively high.

A comprehensive review of the literature covering 497 response rates in 93 journal articles found weighted average response rates of 81.7%, 72.3% and 47.3% for, respectively, personal, telephone and mail surveys.[20] The same review also found that response rates increase with the following:

- Either prepaid or promised monetary incentives
- An increase in the amount of monetary incentive
- Non-monetary premiums and rewards (pens, pencils, books)
- Preliminary notification
- Foot-in-the door techniques. These are multiple request strategies. The first request is relatively small, and all or most people agree to comply. The small request is followed by a larger request, called the **critical request**, which is actually the target behaviour being researched
- Personalisation (sending letters addressed to specific individuals)
- Follow-up letters.

A further discussion of improving response rates is found in Chapter 15.

Perceived respondent anonymity

Perceived respondent anonymity refers to the respondents' perceptions that their identities will not be discerned by the interviewer or the researcher. Perceived anonymity of the respondent is high in mail surveys, mail panels and Internet surveys because there is no contact with an interviewer while responding. It is low in personal interviews (in-home, street and CAPI) due to face-to-face contact with the interviewer. Traditional telephone interviews and CATI fall in the middle. It is also moderate with email; while there is no contact with the interviewer, respondents know that their names can be located on the return email.

Social desirability

Social desirability is the tendency of respondents to give answers that they feel to be acceptable in front of others, including interviewers. When a respondent is questioned face to face by an interviewer, they may give an answer that they feel to be 'acceptable' rather than how they really feel or behave. An example of this phenomenon occurred during the Opinion Polls that led up to the British General Election in 1992. The Opinion Polls placed the Labour Party a clear 8% ahead of the Conservative Party. In the actual election, the Conservative Party won, confounding all of the Opinion Poll predictions. This outcome resulted in an investigation by the Royal Statistical Society and the Market Research Society into why all the Opinion Polls had consistently predicted that the Labour Party would win. The investigation revealed that many respondents (who primarily had been interviewed face to face and by telephone) said they would vote Labour, as this was seen to be the socially acceptable thing to do. In the privacy of the polling booth, they voted Conservative.

The best techniques to avoid respondents distorting their views are those where face-to-face contact is avoided; this is the case with all mail methods. Traditional telephone interviews and CATI are moderately good at handling socially desirable responses, as there is an amount of anonymity afforded by not meeting face to face.

Critical request
The target behaviour being researched.

Perceived respondent anonymity
The respondents' perceptions that their identities will not be discerned by the interviewer or researcher.

Social desirability
The tendency of respondents to give answers that may not be accurate but may be desirable from a social standpoint.

The weakest techniques are face-to-face interviews, though in the case of in-home and in-office interviews, the chance to build up a rapport with respondents may nurture the respondent to reveal how they really feel. [21]

Obtaining sensitive information

Sensitive information may mean an issue that is personally sensitive, such as the way in which a respondent may be classified or the use of hygiene products. What may be deemed 'sensitive' varies enormously between different types of respondent. For some respondents, asking questions about the type and amount of household cleaning products may be seen as revealing characteristics of their personal cleanliness; they see it as a sensitive issue and would need a lot of reassurance before revealing the truth. Many classification questions in a survey, such as the respondent's age, gender, educational or income level, can also be seen as highly sensitive. In business research, characteristics of an organisation's activities may be seen as commercially sensitive.

In some situations, the interviewer plays a very important role in explaining to respondents why they are being asked such a question and that their views will be handled in a confidential and proper manner. In-home and in-office interviews allow the time and context to build up such explanations and reassure respondents. Interviews conducted in this way may be seen as the best means to handle certain sensitive topics. For some issues, respondents may not wish to face any interviewer and would like to complete the survey alone. A CAPI can be set up so that the interviewer introduces the respondent to a terminal and then leaves the respondent to get on with the on-screen interview.[22] All mail surveys can be seen as moderately successful in handling sensitive questions. Their success depends upon how well the purpose of

Questions about drinking habits may be perceived as a sensitive issue by some respondents.

sensitive questions is introduced. Handled correctly, respondents can take their time to answer questions without any embarrassing contact. For telephone and street interviews, the interviewer can reassure the respondent about the questions being asked, but may not have the time and context to really relax respondents and overcome any embarrassment.

Potential for interviewer bias

An interviewer can bias the results of a survey by the manner in which he or she:

1 selects respondents (e.g. interviewing a male aged 34 when required to interview a male aged between 36 and 45);
2 asks research questions (omitting questions);
3 poses questions in another way when respondents do not understand the question as presented on the questionnaire;
4 probes (e.g. by offering examples to encourage respondents);
5 records answers (recording an answer incorrectly or incompletely).

The extent of the interviewer's role determines the potential for bias.[23] In-home, in-office and street personal interviews are highly susceptible to interviewer bias. Traditional telephone interviews and CATI are less susceptible, although the potential is still there. For example, with inflection and tone of voice, interviewers can convey their own attitudes and thereby suggest answers. CAPI has a low potential for bias. Mail surveys, mail panels, email and Internet surveys are free of it.

Potential to probe respondents

Though the interviewer has the potential to create bias in the responses elicited from respondents, it is balanced somewhat by the amount of probing that can be done. For example, a survey may ask respondents which brands of beer they have seen advertised on television over the past month. A list of brands could be presented to respondents and they could simply look at the list and call out the names. What may be important in the survey is what brands they could remember. There may be a first response, of a brand that could be remembered unaided. A simple probe such as 'any others?' or 'any involving sports personalities?' could be recorded as a second response.

Much deeper prompts and probes can be conducted, within limits. The intention is not to turn the interview into a qualitative interview but, for example, some of the reasons why a respondent has chosen a brand may be revealed. In-home, in-office and street personal interviews have great potential for probing respondents. Traditional telephone interviews and CATI can also probe but not to the same extent as being face to face. CAPI has limited potential to probe, though particular routines can be built into a survey to ask for further details. Mail surveys, mail panels, email and Internet surveys have very limited means to probe respondents.

Potential to build rapport

Another counter to the bias in the responses elicited from respondents by personal interviews is the amount of rapport that can be built up with respondents. Rapport may be vital to communicate why the survey is being conducted, with a corresponding rationale for the respondent to spend time answering the questions. Beyond motivating respondents to take part in a survey is the need for the respondent to answer truthfully, to reflect upon the questions properly and not to rush through the questionnaire. Building up a good rapport with respondents can be vital to gain a full and honest response to a survey.

In-home, in-office and street personal interviews have great potential to build up rapport with respondents. Traditional telephone interviews and CATI can also

develop rapport but not to the same extent as being face to face. CAPI has limited potential to build up rapport through particular graphics and messages that can be built into a survey. When recruiting respondents to mail panels, an amount of rapport can be built up. Mail surveys, email and Internet surveys have very limited means to build up a rapport with respondents.

Speed

First, there is the speed with which a questionnaire can be created, distributed to respondents, and the data returned. Because printing, mailing and data keying delays are eliminated, data can be in hand within hours of writing an Internet or telephone questionnaire. Data are obtained in electronic form, so statistical analysis software can be programmed to process standard questionnaires and return statistical summaries and charts automatically. Thus, the Internet can be an extremely fast method of obtaining data from a large number of respondents. The email survey is also fast, although slower than the Internet, since more time is needed to compile an email list and data entry is also required.

Traditional telephone interviews and CATI are also fast means of obtaining information. When a central telephone facility is used, several hundred telephone interviews can be done per day. Data for even large national surveys can be collected in a matter of days or even within a day. Next in speed are street and CAPI interviews that reach potential respondents in central locations. In-home personal interviews are slower because there is dead time between interviews while the interviewer travels to the next respondent. To expedite data collection, interviews can be conducted in different markets or regions simultaneously. Mail surveys are typically the slowest. It usually takes several weeks to receive completed questionnaires; follow-up mailings to boost response rates take even longer. Mail panels are faster than mail surveys because little follow-up is required.

Cost

For large samples, the cost of Internet surveys is the lowest. Printing, mailing, keying and interviewer costs are eliminated, and the incremental costs per respondent are typically low, so studies with large numbers of respondents can be done at substantial savings compared with mail, telephone or personal surveys. Personal interviews tend to be the most expensive mode of data collection per completed response, whereas mail surveys tend to be the least expensive. In general, Internet, email, mail surveys, mail panel, traditional telephone, CATI, CAPI, street and personal in-home interviews require progressively larger field staff and greater supervision and control. Hence, the cost increases in this order. Relative costs, however, depend on the subject of enquiry and the procedures adopted.[24]

Selection of survey method(s)

As is evident from Table 10.2 and the preceding discussion, no survey method is superior in all situations. Depending on such factors as information requirements, budgetary constraints (time and money) and respondent characteristics, no, one, two or even all techniques may be appropriate. Remember that the various data collection modes are not mutually exclusive. Rather, they can be employed in a complementary fashion to build on each other's strengths and compensate for each other's weaknesses. The researcher can employ these techniques in combination and develop creative twists within each technique. To illustrate, in a classic project, interviewers

distributed the product, self-administered questionnaires and return envelopes to respondents. Traditional telephone interviews were used for follow-up. Combining the data collection modes resulted in telephone cooperation from 97% of the respondents. Furthermore, 82% of the questionnaires were returned by mail.[25]

Observation techniques

Quantitative observation techniques are extensively used in descriptive research. Observation involves recording the behavioural patterns of people, objects and events in a systematic manner to obtain information about the phenomenon of interest. The observer does not question or communicate with the people being observed unless he or she takes the role of a **mystery shopper**. Information may be recorded as the events occur or from records of past events. Observational techniques may be structured or unstructured, disguised or undisguised. Furthermore, observation may be conducted in a natural or a contrived environment.[26]

Structured versus unstructured observation

For **structured observation**, the researcher specifies in detail what is to be observed and how the measurements are to be recorded, such as when an auditor performs a stock or inventory analysis in a store. This reduces the potential for observer bias and enhances the reliability of the data. Structured observation is appropriate when the phenomena under study can be clearly defined and counted. For example, suppose that the researcher wished to measure the ratio of visitors to buyers in a store. The reason for such observations could be to understand the amount of browsing that occurs in a store. They could observe and count the number of individuals who enter the store and the number who make a purchase. Counting people who enter a shop could be a manual observation, and could have a rule that the store visitor is counted 'if they actually look at any of the products on display'. Counting the number of transactions through the till may be a simpler electronic observation. With these two counts they could simply calculate the required ratio. Structured observation is suitable for use in conclusive research.

In **unstructured observation** the observer monitors all aspects of the phenomenon that seem relevant to the problem at hand, such as observing children playing with new toys and trying to understand what activities they enjoy the most. This form of observation can be used when a research problem has yet to be formulated precisely and when flexibility is needed in observation to identify essential components of the problem and to develop hypotheses. Unstructured observation is most appropriate for exploratory research and as such was discussed in detail in Chapter 6 under the heading of **Ethnography**. Ethnographic techniques require a researcher to spend a large amount of time observing a particular group of people, by sharing their way of life.[27]

Disguised versus undisguised observation

In disguised observation, the respondents are unaware that they are being observed. Disguise enables respondents to behave naturally because people tend to behave differently when they know they are being observed. Disguise may be accomplished by using two-way mirrors, hidden cameras or inconspicuous electronic devices. Observers may be disguised as shoppers, sales assistants or other appropriate roles. One of the most widespread techniques of observation is through the use of mystery shoppers. The following example illustrates what a mystery shopper may observe in a bank service delivery.

Quantitative observation
The recording and counting of behavioural patterns of people, objects and events in a systematic manner to obtain information about the phenomenon of interest.

Mystery shopper
An observer visiting providers of goods and services as if they were really a customer, and recording characteristics of the service delivery.

Structured observation
Observation where the researcher clearly defines the behaviours to be observed and the techniques by which they will be measured.

Unstructured observation
Observation that involves a researcher monitoring all relevant phenomena, without specifying the details in advance.

Ethnography
A research approach based upon the observation of the customs, habits and differences between people in everyday situations.

The mystery squad's tougher challenge[28]

Typically a mystery shopper would go into a bank, note practical things such as the number of counter positions open, the number of people queuing, or the availability of specific leaflets, and then ask a number of specific questions. The mystery shopper takes the role of the ordinary 'man or woman in the street', behaves just as a normal customer would, asks the same sort of questions a customer would, leaves, and fills in a questionnaire detailing the various components observed in their visit. ■

Mystery shopping differs from conventional survey research in that it aims to collect facts rather than perceptions. Conventional customer service research is all about customer perceptions. Mystery shopping, on the other hand, aims to be as objective as possible and to record as accurately as possible what actually happened in encounters such as the following.

Personal visits
- How long were you in the queue?
- How many tills were open?
- Did the counter clerk apologise if you were kept waiting?
- What form of greeting or farewell was given?

Telephone calls
- How many rings were there before the phone was answered?
- Did the person who answered the phone go on to answer all your questions?
- Were you asked a password?
- How many times during the conversation was your name used?

In undisguised observation, respondents are aware that they are under observation. Respondents may be aware of the situation either by being told that an observer is in their presence or by its being obvious that someone is recording their behaviour. Researchers disagree on how much effect the presence of an observer has on behaviour. One viewpoint is that the observer effect is minor and short-lived.[29] The other position is that the observer can seriously bias the behaviour patterns.[30] There are ethical considerations to disguised versus undisguised observations that will be tackled at the end of this chapter.

Go to the Companion Website and read Professional Perspective 11 by David Backinsell. David's article 'The new management tool that's no mystery' details characteristics of this method of observation. It is set very clearly in a context of the process of consumption and service delivery.

Natural versus contrived observation

Natural observation involves observing behaviour as it takes place in the environment. For example, one could observe the behaviour of respondents eating a new menu option in Burger King. In **contrived observation**, respondents' behaviour is observed in an artificial environment, such as a test kitchen.

The advantage of natural observation is that the observed phenomenon will more accurately reflect the true phenomenon, as the behaviour occurs in a context that feels natural to the respondent. The disadvantages are the cost of waiting for the phenomenon to occur and the difficulty of measuring the phenomenon in a natural setting.

Natural observation
Observing behaviour as it takes place in the environment.

Contrived observation
Observing behaviour in an artificial environment.

Observation techniques classified by mode of administration

As shown in Figure 10.2, observation techniques may be classified by mode of administration as personal observation, electronic observation, audit, content analysis and trace analysis.

Personal observation

Personal observation
An observational research strategy in which human observers record the phenomenon being observed as it occurs.

In **personal observation**, a researcher observes actual behaviour as it occurs. The observer does not attempt to control or manipulate the phenomenon being observed but merely records what takes place. For example, a researcher might record the time, day and number of shoppers who enter a shop and observe where those shoppers 'flow' once they are in the shop. This information could aid in designing a store's layout and determining the location of individual departments, shelf locations and merchandise displays.

Electronic observation

Electronic observation
An observational research strategy in which electronic devices, rather than human observers, record the phenomena being observed.

In **electronic observation**, electronic devices rather than human observers record the phenomenon being observed. The devices may or may not require the respondents' direct participation. They are used for continuously recording ongoing behaviour for later analysis.

Of the electronic devices that do not require respondents' direct participation, the A.C. Nielsen audimeter is best known. The audimeter is attached to a television set to record continually the channel to which a set is tuned. Another way to monitor viewers is through the people meter. People meters attempt to measure not only the channels to which a set is tuned but also who is watching.[31] Other common examples include turnstiles that record the number of people entering or leaving a building and traffic counters placed across streets to count the number of vehicles passing certain locations.

The most significant electronic observation form as detailed in Chapter 5 is through the use of the bar code on products. As goods are sold, optical scanners can determine which products have been sold. With a link to a 'loyalty card', electronic observation links the whole array of purchases made by a consumer to the actual identity of that consumer. In this example, electronic observation does not require direct involvement of the participants.

In contrast, there are many electronic observation devices that do require participant involvement. These electronic devices may be classified into five groups: (1) eye tracking monitors, (2) pupilometers, (3) psycho-galvanometers, (4) voice pitch analysers, and (5) devices measuring response latency. **Eye tracking equipment** – such as oculometers, eye cameras or eye view minuters – records the gaze movements of the eye. These devices can be used to determine how a respondent reads an advertisement or views a TV commercial and for how long the respondent looks at various parts of the stimulus. Such information is directly relevant to assessing advertising effective-

Eye tracking equipment
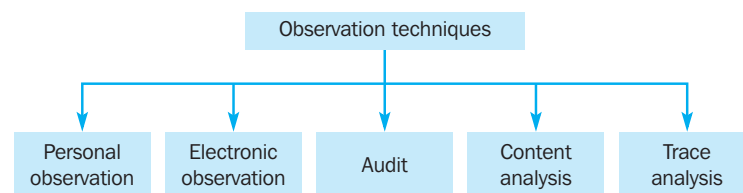Instruments that record the gaze movements of the eye.

Figure 10.2
A classification of observation techniques

Observation techniques

Personal observation | Electronic observation | Audit | Content analysis | Trace analysis

Pupilometer
An instrument that measures changes in the eye pupil diameter.

Psycho-galvanometer
An instrument that measures a respondent's galvanic skin response.

Galvanic skin response
Changes in the electrical resistance of the skin that relate to a respondent's affective state.

Voice pitch analysis
Measurement of emotional reactions through changes in the respondent's voice.

Response latency
The amount of time it takes to respond to a question.

Pantry audit
A type of audit where the researcher inventories the brands, quantities and package sizes of products in a consumer's home.

ness. The **pupilometer** measures changes in the diameter of the pupils of the respondent's eyes. The respondent is asked to look at a screen on which an advertisement or other stimulus is projected. Image brightness and distance from the respondent's eyes are held constant. Changes in pupil size are interpreted as changes in cognitive (thinking) activity resulting from exposure to the stimulus. The underlying assumption is that increased pupil size reflects interest and positive attitudes towards the stimulus.[32]

The **psycho-galvanometer** measures **galvanic skin response** (GSR) or changes in the electrical resistance of the skin.[33] The respondent is fitted with small electrodes that monitor electrical resistance and is shown stimuli such as advertisements, packages and slogans. The theory behind this device is that physiological changes such as increased perspiration accompany emotional reactions. Excitement leads to increased perspiration, which increases the electrical resistance of the skin. From the strength of the response, the researcher infers the respondent's interest level and attitudes toward the stimuli.

Voice pitch analysis measures emotional reactions through changes in the respondent's voice. Changes in the relative vibration frequency of the human voice that accompany emotional reaction are measured with audio-adapted computer equipment.[34]

Response latency is the time a respondent takes before answering a question. It is used as a measure of the relative preference for various alternatives.[35] Response time is thought to be directly related to uncertainty. Therefore, the longer a respondent takes to choose between two alternatives, the closer the alternatives are in terms of preference. On the other hand, if the respondent makes a quick decision, one alternative is clearly preferred. With the increased popularity of computer-assisted data collection, response latency can be recorded accurately and without the respondent's awareness.

Use of eye-tracking monitors, pupilometers, psycho-galvanometers and voice pitch analysers assumes that physiological reactions are associated with specific cognitive and affective responses. This has yet to be clearly demonstrated.[36] Furthermore, calibration of these devices to measure physiological arousal is difficult, and they are expensive to use. Another limitation is that respondents are placed in an artificial environment and know that they are being observed.

Audit

In an audit, the researcher collects data by examining physical records or performing inventory analysis. Audits have two distinguishing features. First, data are collected personally by the researcher. Second, the data are based upon counts, usually of physical objects. Retail and wholesale audits conducted by marketing research suppliers were discussed in the context of secondary data (see Chapter 4). Here we focus on the role of audits in collecting primary data. In this respect, an important audit conducted at the consumer level, generally in conjunction with one of the survey techniques, is the **pantry audit**. In a pantry audit, the researcher takes an inventory of brands, quantities and package sizes in a consumer's home, perhaps in the course of a personal interview. Pantry audits greatly reduce the problem of untruthfulness or other forms of response bias. Obtaining permission to examine consumers' pantries can be difficult, however, and the fieldwork is expensive. Furthermore, the brands in the pantry may not reflect the most preferred brands or the brands purchased most often. Moreover, similar data can be obtained from scanned data more efficiently. For these reasons, pantry audits are no longer commonly used; audits are more common at the retail and wholesale level.

Content analysis

Content analysis is an appropriate method when the phenomenon to be observed is communication, rather than behaviour or physical objects. It is defined as the objective, systematic and quantitative description of the manifest content of a communication.[37] It includes observation as well as analysis. The unit of analysis may be words (different words or types of words in the message), characters (individuals or objects), themes (propositions), space and time measures (length or duration of the message) or topics (subject of the message). Analytical categories for classifying the units are developed, and the communication is broken down according to prescribed rules. Marketing research applications involve observing and analysing the content or message of advertisements, newspaper articles, television and radio programmes and the like. For example, the frequency of appearance of ethnic minorities and women has been studied using content analysis. In the GlobalCash Project, content analysis may be used to analyse magazine advertisements of the sponsoring and competing banks to compare their projected images. A crucial requirement to the success of content analysis is that the categories are sufficiently precise to enable different coders to arrive at the same results when the same body of material (e.g. advertising copy) is examined.[38]

In qualitative research, content analysis is one of the classical procedures for analysing textual material. The text being analysed may come from the narrative held in brochures or advertising copy through to dialogues held in interview data. Primarily the objective of content analysis is to 'reduce' the data, to simplify by summarising and structuring the data according to rules derived from existing theory. In effect, therefore, content analysis should be classified as a quantitative technique based upon classifying and 'counting'.

Trace analysis

An observation method that can be inexpensive if used creatively is **trace analysis**. In trace analysis, data collection is based on physical traces, or evidence, of past behaviour. These traces may be left by the respondents intentionally or unintentionally. Several innovative applications of trace analysis have been made in marketing research.

- The selective erosion of tiles in a museum indexed by the replacement rate was used to determine the relative popularity of exhibits.
- The number of different fingerprints on a page was used to gauge the readership of various advertisements in a magazine.
- The position of the radio dials in cars brought in for service was used to estimate share of listening audience of various radio stations. Advertisers used the estimates to decide on which stations to advertise.
- The age and condition of cars in a car park were used to assess the affluence of customers.
- The magazines people donated to charity were used to determine people's favourite magazines.
- Internet visitors leave traces that can be analysed to examine browsing and usage behaviour through **cookie technology**.

A comparative evaluation of observation techniques

A comparative evaluation of the observation techniques is given in Table 10.3. The different observation techniques are evaluated in terms of the degree of structure, degree of disguise, ability to observe in a natural setting, observation bias, measurement and analysis bias, and additional general factors.

Table 10.3 A comparative evaluation of observation techniques

Criteria	Personal observation	Electronic observation	Audit	Content analysis	Trace analysis
Degree of structure	Low	Low to high	High	High	Medium
Degree of disguise	Medium	Low to high	Low	High	High
Natural setting	High	Low to high	High	Medium	Low
Observation bias	High	Low	Low	Medium	Medium
Analysis bias	High	Low to medium	Low	Low	Medium
General remarks	Most flexible	Can be intrusive	Expensive	Limited to communications	Limited traces available

Structure relates to the specification of what is to be observed and how the measurements are to be recorded. As can be seen from Table 10.3, personal observation is low, trace analysis is medium, and audit and content analysis are high on the degree of structure. Electronic observation can vary widely from low to high, depending on the techniques used. Techniques such as optical scanners are very structured in that the characteristics to be measured – for example, characteristics of items purchased scanned in supermarket checkouts – are precisely defined. In contrast, electronic techniques, such as the use of hidden cameras to observe children at play with toys, tend to be unstructured.

The degree of disguise is low in the case of audits as it is difficult to conceal the identity of auditors. Personal observation offers a medium degree of disguise because there are limitations on the extent to which the observer can be disguised as a shopper, sales assistant, employee and so forth. Trace analysis and content analysis offer a high degree of disguise because the data are collected 'after the fact', that is, after the phenomenon to be observed has taken place. Some electronic observations such as hidden cameras offer excellent disguise, whereas others, such as the use of psycho-galvanometers, are very difficult to disguise.

The ability to observe in a natural setting is low in trace analysis because the observation takes place after the behaviour has occurred. It is medium in the case of content analysis because the communication being analysed is only a limited representation of the natural phenomenon. Personal observation and audits are excellent on this score because human observers can observe people or objects in a variety of natural settings. Electronic observation techniques vary from low (e.g. use of psycho-galvanometers) to high (e.g. use of turnstiles).

Observation bias is low in the case of electronic observation because a human observer is not involved. It is also low for audits. Although the auditors are humans, the observation usually takes place on objects and the characteristics to be observed are well defined, leading to low observation bias. Observation bias is medium for trace analysis and content analysis. In both these techniques, human observers are involved and the characteristics to be observed are not very well defined. The observers typically do not interact with human respondents during the observation process, however, thus lessening the degree of bias. It is high for personal observation due to the use of human observers who interact with the phenomenon being observed.

Data analysis bias is low for audits and content analysis because the variables are precisely defined, the data are quantitative, and statistical analysis can be conducted. Trace analysis has a medium degree of bias as the definition of variables is not very precise. Electronic observation techniques can have a low (e.g. scanner data) to

medium (e.g. hidden camera) degree of analysis bias. Unlike personal observation, the bias in electronic observation is limited to the medium level due to improved measurement and classification, because the phenomenon to be observed can be recorded continuously using electronic devices.

In addition, personal observation is the most flexible, because human observers can observe a wide variety of phenomena in a wide variety of settings. Some electronic observation techniques, such as the use of psycho-galvanometers, can be very intrusive, leading to artificiality and bias. Audits using human auditors tend to be expensive. Content analysis is well suited for and limited to the observation of communications. As mentioned earlier, trace analysis is a method that is limited to where consumers actually leave 'traces'. This occurs infrequently and very creative approaches are needed to capture these traces.

Evaluating the criteria presented in Table 10.3 helps to identify the most appropriate observation technique, given the phenomena to be observed, the nature of respondents being observed and the context in which the observation occurs. To strengthen the choice of a particular observation technique, it is also helpful to compare the relative advantages and disadvantages of observation versus survey techniques.

Advantages and disadvantages of observation techniques

Other than the use of scanner data, few marketing research projects rely solely on observational techniques to obtain primary data.[39] This implies that observational techniques have some major disadvantages compared with survey techniques. Yet these techniques offer some advantages that make their use in conjunction with survey techniques quite fruitful.

Relative advantages of observation techniques

The greatest advantage of observational techniques is that they permit measurement of actual behaviour rather than reports of intended or preferred behaviour. There is no reporting bias, and potential bias caused by the interviewer and the interviewing process is eliminated or reduced. Certain types of data can be collected only by observation. These include behaviour patterns which the respondent is unaware of or unable to communicate. For example, information on babies' toy preferences is best obtained by observing babies at play, because they are unable to express themselves adequately. Moreover, if the observed phenomenon occurs frequently or is of short duration, observational techniques may cost less and be faster than survey techniques.

Relative disadvantages of observation techniques

The biggest disadvantage of observation is that the reasons for the observed behaviour may be difficult to determine because little is known about the underlying motives, beliefs, attitudes and preferences. For example, people observed buying a brand of cereal may or may not like it themselves; they may be purchasing that brand for someone else in the household. Another limitation of observation is the extent to which the researcher is prepared to evaluate the extent of their own bias, and how this can affect what they observe. In addition, observational data can be time-consuming and expensive to collect. It is also difficult to observe certain forms of behaviour such as personal activities that occur in the privacy of the consumer's home. Finally, in some cases such as in the use of hidden cameras, the use of observational techniques may border on being or may actually be unethical. It can be argued that individuals being observed should be made aware of the situation, but this may cause them to behave in a contrived manner.

International marketing research

The selection of appropriate interviewing techniques is much more difficult in foreign countries because of the challenges of conducting research there. Given the differences in the economic, structural, informational, technological and sociocultural environment, the feasibility and popularity of the different interviewing techniques vary widely. In the United States and Canada, for example, the telephone has achieved almost total penetration of households. Consequently, telephone interviewing is a dominant mode of questionnaire administration. The same situation exists in some European countries, such as Sweden. In many other European countries, however, the telephone interview gets confused with telephone sales. This results in high refusal rates and scepticism of what the purpose of a survey is. In developing countries, the problem with the telephone is the low number of households that have telephones.

Because of the low cost, mail interviews continue to be used in most developed countries where literacy is high and the postal system is well developed, for example in Britain, Canada, Denmark, Finland, Iceland, the Netherlands, Norway, Sweden and the United States. In many parts of Africa, Asia and South America, however, the use of mail surveys and mail panels is low because of illiteracy and the large proportion of the population living in rural areas.

The following example illustrates how CAPI technology has developed to allow consistent approaches to survey techniques across Europe. It shows that, by developing an appreciation of the cultural differences between countries, a research design can be built that allows accurate and comparable surveys to be conducted.

example

CAPIBUS Europe [40]

The concept of the International Omnibus is not new; many research groups offer an international service. In reality, however, these have been little more than brokering services, bolting together whatever omnibus is available in each country, with little real standardisation.

Research Services Limited and the IPSOS group have tackled this problem by calling on the benefits of CAPI technology to introduce CAPIBUS Europe, a weekly omnibus survey covering the six major markets of Europe (Britain, France, Germany, Italy, Netherlands and Spain).

The use of computer technology means that questionnaires can be scripted in one location and transmitted electronically to other countries. While the need to ensure accurate translation remains, the problems involved in having different questionnaire formats, classification systems and data maps are minimised. At the end of a project, data are again transmitted electronically, to be aggregated by the lead agency in a standardised format for all markets. This can then be weighted to provide information on the European market as well as for each local market. ■

Selection of survey techniques

No questionnaire administration method is superior in all situations. Table 10.4 presents a comparative evaluation of the major modes of collecting quantitative data in the context of international marketing research. In this table, the survey techniques are discussed only under the broad headings of telephone, personal, mail and electronic interviews. The use of CATI, CAPI, electronic survey and mail panels depends heavily on the state of technological development in the country. Likewise, the use of street interviewing is contingent upon the dominance of shopping centres in the retailing environment.

Table 10.4 A comparative evaluation of survey techniques for international marketing research

Criteria	Telephone	Personal	Mail	Electronic
High sample control	+	+	–	–
Difficulty in locating respondents at home	+	–	+	+
Inaccessibility of homes	+	–	+	+
Unavailability of a large pool of trained interviewers	+	–	+	+
Large population in rural areas	–	+	–	–
Unavailability of maps	+	–	+	+
Unavailability of current telephone directory	–	+	–	+
Unavailability of mailing lists	+	+	–	+
Low penetration of telephones	–	+	+	–
Lack of an efficient postal system	+	+	–	+
Low level of literacy	–	+	–	–
Face-to-face communication culture	–	+	–	–
Poor access to computers and Internet	+	+	+	–

Note: A + denotes an advantage, and a – denotes a disadvantage

Ethics in marketing research

Respondent anonymity was discussed in the context of qualitative research in Chapter 8. It was argued then that large samples could 'hide' the specific answers of individual respondents. This is true to some extent but respondents may still be identifiable. It is up to the marketing researcher to protect a respondent's identity and not disclose it to anyone outside the research organisation, including the client. The client is not entitled to see the names and contact details of respondents. The only instance where respondents' identity can be revealed to the client is when respondents are notified in advance and their consent is obtained prior to administering the survey. Even in such situations, the researcher should have the assurance that the respondents' trust will be kept by the client and their identities will not be used in a sales effort or misused in other ways.

Special care must be taken to ensure that any record which contains a reference to the identity of an informant is securely and confidentially stored during any period before such reference is separated from that record and/or destroyed.[41] Ethical lapses in this respect by unscrupulous researchers and marketers have resulted in a serious backlash for marketing research. The result has been a consistent fall in the levels of response rate, to all forms of survey method. This reinforces the message that considering the needs of survey respondents makes sound research sense as well as being ethically sound. Dubious practices may generate a response for a single survey but may create long-term damage to the marketing research industry.

Another issue facing the marketing research industry is image, as the public may not distinguish between telephone research and telemarketing. This identity crisis is exacerbated by the action of some firms to commit 'sugging and frugging', industry terms for selling or fund-raising under the guise of a survey. The damage done to the marketing research industry by the unethical use of survey research as a guise for targeting sales effort was discussed in Chapter 1. The overall effect of these

activities has given a poor image in particular to telephone research, raising the cost and making it difficult for researchers to obtain full and representative samples.

Although concerns for the respondents' psychological well-being are mild in survey data collection when compared with either qualitative or experimental research, researchers should not place respondents in stressful situations. Disclaimers such as 'there are no correct responses; we are only interested in your opinion' can relieve much of the stress inherent in a survey. In many face-to-face interview situations, respondents are given a 'thank you booklet' at the end of the interview. As well as saying a genuine thank you for taking part in a survey, the booklet briefly sets out the purpose and benefits of bona fide marketing research. The use of the 'thank you booklet' helps to educate the public to distinguish between genuine, professionally conducted marketing research and 'research' conducted as a front for generating sales leads.

Observation of people's behaviour without their consent is often done because informing the respondents may alter their behaviour.[42] But this can compromise the privacy of the respondents. One guideline is that people should not be observed for research in situations where they would not expect to be observed by the public. Therefore, public places like a shopping centre or a grocery aisle are fair game. These are places where people observe other people routinely. However, notices should be posted in these areas stating that they are under observation by marketing researchers.

With the growth of mystery shopping, where the essence of the technique is that service deliverers cannot spot the observer, the debate over what is ethical practice has intensified. The following example taken from a seminar on mystery shopping illustrates some of the issues that the marketing research industry considers vital to debate for the future 'health' of the technique.

e x a m p l e   **The ethical puzzle of mystery shopping**[43]

'For us to believe we can do this without some form of public outrage would make us very naive indeed.' This was the statement of Professor Roger Jowell to delegates attending the joint Social Research Association/Market Research Society Seminar on mystery shopping.

Mystery shopping has been going on for years in one form or another but was formally originated in the USA in the early 1970s. The definition of mystery consumer research used at the seminar states: 'The use of individuals *trained* to observe . . . by acting as a prospective customer.' Whether the individuals should be trained or not brought heated discussion, some suggesting that training would not encourage the shopper to act 'naturally' while others felt training was necessary in some form or another.

The contentious issue of competitive videoing was debated. One agency director said 'this is fly on the wall journalism' and clients should go to a production company rather than a marketing research agency. One agency had been asked by a client to mystery shop competitors only and, on turning to their code of conduct, found that there was nothing to stop them going ahead with this course of action.

Professor Jowell likened competitive shopping to industrial espionage, at which point Shirley Featherstone, Field Director of ACE Fieldwork, suggested that unlike industrial espionage, mystery shopping is information that is in the public domain.

Colin Brown, an independent consultant, explained that traditionally mystery shopping has been used:

■ to campaign on public issues with which they have been able to make changes,
■ as a monitor within companies to assess staff performance.

He then pointed to the latest tradition, that of enforcement, which results in people losing their jobs. ■

The final point in the above example graphically illustrates how important ethical issues are in conducting marketing research. If potential respondents perceive marketing research as a means to generate sales and 'snoop' on service delivery practices, then the goodwill needed to generate full and honest responses will wither. ESOMAR have specific guidelines to help with issues in mystery shopping, which can be seen on **www.esomar.nl/guidelines/mysteryshopping.htm**.

Internet and computer applications

We begin this section with a description of a technological development that may have profound effects upon how CAPI is undertaken. We then move on to broader Internet and computer issues.

A major technological development that has affected the conduct of CAPI has been Personal Digital Assistants (PDAs). PDA is the generic title for a range of devices, from the ubiquitous PalmPilot from 3Com to the Pocket PCs from manufacturers such as Hewlett Packard and Compaq. Essentially they are computers with a small touch-screen that fit in your hand.

Recent developments have included colour screens, and increasingly telecommunications are being added to new models being released. 2002 has seen the first appearances of PDAs that merge both a mobile phone and a mobile computer. This will extend their functionality and take-up even further, and might see the end of the mobile phone as we know it today.

Marketing research applications for PDAs started to appear in 2001, with Techneos (**www.techneos.com**) producing a specific product for the PalmPilot. Other software developers have opted for the Pocket PCs using variations of Windows CE from Microsoft. These have included SNAP, **www.snapsurveys.com**. The logic of basing mobile interviewing on Windows-based products has been that software authors are less dependent on the fortunes of a single vendor of hardware.

The main benefits of using PDAs over the more traditional CAPI laptop or even the traditional paper and pencil clipboard solution are:

- Low cost
- Portability
- The fact that they are seen as a technology-based solution and in certain market sectors this is important.

Moving to broader issues, two factors dominate the computer-related aspects of the survey process:

1 It is now likely that 99% of all quantitative surveys today include an element of IT at some stage in their life cycle. This may be simply for the questionnaire design stage or for the presentation of results. However, the computer is increasingly involved in many stages of the survey process – questionnaire design and printing, preparation of CATI scripts, publication of online surveys, collecting of email responses, scanning of paper questionnaires, keying of replies, data validation, cleaning and verification, data analysis and presentation, electronic distribution of survey results, and integration of results into office products, to name but a few.

2 Surveys increasingly involve more than one method of data collection. A survey might be designed for the Web, but respondents may be emailed a paper version to view the questions, and asked to complete offline if access to the Web is not possible. Surveys are increasingly being designed to incorporate alternative scenarios, perhaps using interviewers with PDAs alongside mailed questionnaires for self-completion.

The result is that researchers are becoming more sophisticated in their IT requirements. They are now less likely to outsource the entire process to the DP department or an outside agency, and will need to be more heavily involved in the IT process than they were perhaps 10 years ago. This is certainly a beneficial development, as the researcher will acquire a far better understanding of the principles of survey design and analysis.

The increased use of multiple data collection methods may, however, present a further dilemma. On the one hand, an array of different software packages can be used. This approach is likely to generate surveys that are rich in their list of features, but possibly a nightmare to integrate. On the other hand, a single software package may be used that is capable of handling the range of options being requested. This approach may generate a solution that is easier to manage and operate, but possibly less rich in its presentation.

Internet surveys are gaining in popularity and have great potential for the marketing research industry. One reason is that the cost in most cases is less than that of personal, telephone or mail surveys. The cost of designing online surveys (with graphics, colour, buttons, routing, etc.) is likely to be higher than for a standard paper questionnaire, but the cost of data collection and preparation will be greatly reduced. Also, an Internet survey is not as intrusive as a phone call in the middle of dinner or at other inconvenient times. The online survey can be completed in one's own time, place and pace. Quick response time is another advantage cited by those producing online surveys. It is not true, however, that response rates are increased by using online surveys. Indeed, they have been shown to be no more successful than any other method if no follow-up reminders are included.[44]

Internet surveys can be used to target specific populations or potential target markets. There is a growth of specialist directories that include email and Internet addresses, but these are likely to be even less reliable than telephone directories, as online users frequently change their email addresses. Unlike telephone numbers, which do at least result in somebody answering, routing to a voicemail or simply responding as 'unobtainable', invalid email addresses simply bounce back. There is no indication as to whether the address has been misspelt, whether that person still exists at that email address, or whether the email address itself any longer exists. With no way to speak to a human at the end of an email address, the researcher is faced with a lost opportunity.

It has been argued that the use of the Internet to conduct marketing research is more revolutionary than other technological developments such as CATI or CAPI. Practitioners argue that the Internet has fundamentally changed the way that marketing researchers design questionnaires, collect data and analyse it. [45] The Internet is a visual medium: it allows respondents to see images, long text messages, long lists of response options and, as bandwidth grows, video images. It captures the unedited view of the respondent, eliciting responses to open-ended questions that are long, rich and revealing. It may be more effective in addressing sensitive issues; adults may be more willing to reveal information about their experiences of sensitive medical conditions, for example.

The downside of the Internet is its representativeness of broader populations. It is extremely difficult to conduct probability sampling, and the respondent population depends mostly on 'volunteer' or convenience sampling (see Chapter 14 for a full discussion of the sampling implications). Another limitation is the verification of who is actually responding to the survey. The absence of a human facilitator to motivate participants, security and privacy are also areas of concern.

There are also ethical considerations in conducting surveys on the Internet that primarily concern the fact that it is so inexpensive to use and difficult to regulate. This means that it can be open to misuse by less experienced or less scrupulous organisations, often based outside the research industry. Any Internet surveys which fall seriously below the high standards promoted by ESOMAR and other leading professional bodies make it more difficult to use the medium for research and could seriously damage the credibility of such research, as well as being an abuse of the goodwill of Internet users generally. ESOMAR have specific guidelines that address the use of the Internet to protect the interests both of Internet respondents and of the users of Internet research findings. These can be accessed on **www.esomar.nl/guidelines/internet_guidelines.htm.**

Visit **www.mercator.co.uk** to review the SNAP survey software, **www.camsp.com** to see Keypoint2, **www.sawtooth.com** (click on the PC-based interviewing mode Ci3), **www.voxco.com** for Web, CATI and CAPI software in English and French formats, and **www.scolari.co.uk** and click on SphinxSurvey and **www.askia.com** for the Askia survey software. On all these sites you will be able to see an array of applications of survey software and to download demo versions.

Summary

The two basic means of obtaining primary quantitative data in descriptive research are through survey and observation techniques. Survey involves the direct questioning of respondents, while observation entails recording respondent behaviour or the behaviour of service deliverers.

Surveys involve the administration of a questionnaire and may be classified, based on the method or mode of administration, as (1) traditional telephone interviews, (2) Computer Assisted Telephone Interviews (CATI), (3) in-home or in-office personal interviews, (4) street interviews, (5) Computer Assisted Personal Interviews (CAPI), (6) traditional mail surveys, (7) mail panels, (8) email surveys, and (9) Internet surveys. Of these techniques, CATI, CAPI and Internet surveys have grown enormously in their use in developed Western economies. Each method has some general advantages and disadvantages, however. Although these data collection techniques are usually thought of as distinct and 'competitive', they should not be considered to be mutually exclusive in much the same manner as using quantitative and qualitative techniques should not be considered to be mutually exclusive. It is possible to employ them productively in combination.

Quantitative observational techniques may be classified as structured or unstructured, disguised or undisguised, and natural or contrived. The major techniques are personal observation (including mystery shopping), electronic observation, audit, content analysis and trace analysis. Compared with surveys, the relative advantages of observational techniques are that they permit measurement of actual behaviour, there is no reporting bias, and there is less potential for interviewer bias. Also, certain types

of data can best, or only, be obtained by observation. The relative disadvantages of observation are that very little can be inferred about motives, beliefs, attitudes and preferences, there is a potential for observer bias, most techniques are time-consuming and expensive, it is difficult to observe some forms of behaviour, and questions of ethical techniques of observation are far more contentious. Observation is rarely used as the sole method of obtaining primary data, but it can be usefully employed in conjunction with other marketing research techniques.

In collecting data from different countries, it is desirable to use survey techniques with equivalent levels of reliability rather than to use the same method. Respondents' anonymity should be protected, and their names should not be turned over to the clients. People should not be observed without consent for research in situations where they would not expect to be observed by the public. Practitioners argue that the Internet has fundamentally changed the way that marketing researchers design questionnaires, collect survey data and analyse it.

Questions

1 With a context of the survey researcher imposing their language and logic upon potential respondents, what do you see as being the advantages and disadvantages of conducting surveys?

2 Discuss the dilemma faced by the survey designer who wishes to develop a survey that is not prone to interviewer bias but also sees that interviewer rapport with respondents is vital to the success of the survey.

3 Evaluate the reasons why response rates to industrial surveys are declining.

4 Why do interviewers need to probe respondents in surveys? What distinguishes survey probing from probing conducted in qualitative interviews?

5 What are the relevant factors for evaluating which survey method is best suited to a particular research project?

6 What are the distinct advantages of conducting a survey using CAPI technology compared to traditional paper questionnaires?

7 What are the key advantages of conducting interviews on the Internet? Evaluate the potential that this technique holds for the future.

8 How would you classify mystery shopping as an observation technique? Why would you classify it in this way?

9 How may electronic observation techniques be used in supermarkets?

10 Explain, using examples, where content analysis may be used.

11 Describe the criteria by which you would evaluate the relative benefits of different observation techniques.

12 What is the difference between qualitative and quantitative observation?

13 Describe the relative advantages and disadvantages of observation.

14 Describe a marketing research problem in which both survey and observation techniques could be used for obtaining the information needed.

15 What do you see as being the main ethical problems of mystery shopping?

Notes

1 *Marketing* (9 October 1997), 33.

2 Savage, M., 'The fans of the phone pump the numbers up', *Research* (September 1997), 30.

3 Kirby, R., 'A pan-European view of the executive at lunch', *Research Plus* (October 1992), 7.

4 Burrows, S., 'Verbatim', *Research* (September 1995), 20.

5 Luker, J., 'Omnibus – the flexible research tool for many jobs', *ResearchPlus* (February 1996), 12.

6 'Auto-mania at the mall', *American Demographics* (Marketing Tools Supplement) (June 1997), 6; Bush, A.J. and Hair Jr, J.E., 'An assessment of the mall-intercept as a data collection method', *Journal of Marketing Research* (May 1985), 158–67.

7 Denny, M., 'How the omnibus hit the fast track – and now, hold very tight', *Research Plus* (February 1996), 4.

8 Mail surveys are common in institutional and industrial marketing research. See, for example, Brossard, H.L., 'Information sources used by an organisation during a complex decision process: an exploratory study', *Industrial Marketing Management* 27(1) (January 1998), 41–50.

9 Ganesan, S., 'Determinants of long-term orientation in buyer seller relationships', *Journal of Marketing* 58 (April 1994), 1–19.

10 Yoegei, R., 'List marketers head to cyberspace', *Target Marketing* 20(8) (August 1997), 54–5.

11 Mann, C. and Stewart, F., *Internet Communication and Qualitative Research: A Handbook for Researching Online* (London: Sage, 2000), 67.

12 Stevens, J. and Chisholm, J., 'An integrated approach: technology firm conducts worldwide satisfaction research survey via email, Internet', *Quirk's Marketing Research Review* 11(8) (October 1997): 12–13, 64–5.

13 Johnston, A., 'Welcome to the wired world', *Research* (November 1999), 22–5; Comley, P., 'Will working the web provide a net gain?', *Research* (December 1996), 16.

14 'Demonstrating control: more retailers are calling the shots in choosing sampling and demonstration firms', *Supermarket News* (Brand Marketing Supplement) (4 March 1996), 1; Childers, T.L. and Skinner, S.J., 'Theoretical and empirical issues in the identification of survey respondents', *Journal of the Market Research Society* 27 (January 1985), 39–53.

15 Smith, W., Mitchell, P., Attebo, K. and Leeder, S., 'Selection bias from sampling frames: telephone directory and electoral rolls compared to door-to-door population census: results from the Blue Mountain eye study', *Australian & New Zealand Journal of Public Health* 21(2) (April 1997) 127–33; Czaja, R., Blair, J. and Sebestik, J.P., 'Respondent selection in a telephone survey: a comparison of three techniques', *Journal of Marketing Research* 19 (August 1982), 381–5; O'Rourke, D. and Blair, J., 'Improving random respondent selection in telephone interviews', *Journal of Marketing Research* 20 (November 1983), 428–32.

16 Printz, J. and Maltby, D., 'Beyond personalisation: when handwriting makes a difference', *Fund Raising Management* 28(3) (May 1997), 16–19; Conant, J.S., Smart, D.T. and Walker, B.J., 'Mail survey facilitation techniques: an assessment and proposal regarding reporting practices', *Journal of the Market Research Society* 32 (October 1990), 569–80.

17 Edmonston, J., 'Why response rates are declining', *Advertising Age's Business Marketing* 82(8) (September 1997) 12; Hubbard, R. and Little, E.L., 'Promised contributions to charity and mail survey responses: replications with extension', *Public Opinion Quarterly* 52 (Summer 1988), 223–30; Erdos,

P.L. and Ferber, R. (eds), 'Data collection methods: mail surveys', *Handbook of Marketing Research* (New York: McGraw-Hill, 1974), 102.

18 Grey, R., 'Speeding up the process', *Campaign-London* (Information Interpretation) (18 October 1996), 7; Guengel, P.C., Berchman, T.R. and Cannell, C.E., *General Interviewing Techniques: A Self Instructional Workbook for Telephone and Personal Interviewer Training* (Ann Arbor, MI: Survey Research Center, University of Michigan, 1983).

19 Fletcher, K., 'Jump on the omnibus', *Marketing* (15 June 1995), 25–8.

20 Nicholls II, W.L., 'Highest response', *Marketing Research: A Magazine of Management and Applications* 8(1) (Spring 1996), 5–7; Yu, J. and Cooper, H., 'A quantitative review of research design effects on response rates to questionnaires', *Journal of Marketing Research* 20 (February 1983), 36–44. See also James, J.M. and Bolstein, R., 'The effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys', *Public Opinion Quarterly* 54 (Fall 1990), 346–61.

21 Ones, D., Reiss, A.D. and Viswesvaran, C., 'Role of social desirability in personality testing for personnel selection: the red herring', *Journal of Applied Psychology* 81(6) (December 1996), 660–79.

22 Vinten, G., 'The threat in the question', Credit Control 18(1) (1997), 25–31; Raghubir, P. and Menon, G., 'Asking sensitive questions: the effects of type of referent and frequency wording in counterbiasing method', *Psychology & Marketing* 13(7) (October 1996), 633–52.

23 Rubel, C., 'Researcher praises online methodology', *Marketing News* 30(12) (3 June 1996), H18; Cannell, C.E., Miller, P.U., Oksenberg, L. and Leinhardt, S. (eds), 'Research on interviewing techniques', in *Sociological Methodology* (San Francisco, CA: Jossey-Bass, 1981); Miller, P.U. and Cannell, C.E., 'A study of experimental techniques for telephone interviewing', *Public Opinion Quarterly* 46 (Summer 1982), 250–69.

24 'How the survey was designed', *Management Accounting* 79(8) (February 1998), 48; Fink, A., *A Survey Handbook* (Thousand Oaks, CA: Sage, 1995).

25 Bolton, R.N., 'Covering the market', *Marketing Research: A Magazine of Management and Applications* 6(3) (Summer 1994), 30–5; Payne, S.L., 'Combination of survey methods', *Journal of Marketing Research* (May 1964), 62.

26 Wilcox, S.B., 'Trust, but verify', *Appliance Manufacturer* 46(1) (January 1998), 8, 87; Rust, L., 'How to reach children in stores: marketing tactics grounded in observational research', *Journal of Advertising Research* 33 (November–December 1993), 67–72.

27 Travers, M., *Qualitative Research through Case Studies* (London: Sage, 2001), 4.

28 McNeil, R., 'The mystery squad's tougher challenge', *ResearchPlus* (April 1994), 13.

29 'Keeping track of the customer', *Retail Business-Market Surveys* (475) (September 1997), 1–11; Scott, C., Klein, D.M. and Bryant, J., 'Consumer response to humour in advertising: a series of field studies using behavioural observation', *Journal of Consumer Research* 16 (March 1990), 498–501; Kerlinger, F.N., *Foundations of Behavioural Research*, 3rd edn (New York: Holt, Rinehart & Winston, 1986), 538.

30 Seaton, A.V., 'Unobtrusive observational measures as a qualitative extension of visitor surveys at festivals and events: mass observation revisited', *Journal of Travel Research* 35(4) (Spring

1997), 25–30; Webb, E.J., Campbell, D.T., Schwarts, K.D. and Sechrest, L., *Unobtrusive Measures: Non-reactive Research in the Social Sciences* (Chicago, IL: Rand McNally, 1966), 113–14.

31 Gold, L.N., 'Technology in television research: the meter', *Marketing Research: A Magazine of Management and Applications* 6(1) (Winter 1994) 57–8.

32 Russo, J.E. and Leclerc, F., 'An eye-fixation analysis of choice processes for consumer non-durables', *Journal of Consumer Research* 21 (September 1994), 274–90.

33 For examples of an application of GSR, see LaBarbera, P.A. and Tucciarone, J.D., 'GSR reconsidered: a behaviour based approach to evaluating and improving the sales potency of advertising', *Journal of Advertising Research* 35(5) (September/October 1995), 33–53; Abeele, P.V. and Maclachlan, D.L., 'Process tracing of emotional responses to TV ads: revisiting the warmth monitor', *Journal of Consumer Research* 20 (March 1994), 586–600.

34 Gregory, S., Webster, S. and Huang. G., 'Voice pitch and amplitude convergence as a metric of quality in dyadic interviews', *Language and Communication* 13(3) (July 1993), 195–217; Buckman, G.A., 'Uses of voice-pitch analysis', *Journal of Advertising Research* 20 (April 1980), 69–73.

35 Bassili, J.N. and Scott, B.S., 'Response latency as a signal to question problems in survey research', *Public Opinion Quarterly* 60(3) (Fall 1996), 390–9; Aaker, D.A., Bagozzi, R.P., Carman, J.M. and MacLachlan, J.M., 'On using response latency to measure preference', *Journal of Marketing Research* 17 (May 1980), 237–44.

36 Rachlin, H., 'Can we leave cognition to cognitive psychologists? Comments on an article by George Loewenstein', *Organisational Behaviour & Human Decision Processes* 65(3) (March 1996), 300–4; Stewart, D.W., 'Physiological measurement of advertising effects', *Psychology and Marketing* (Spring 1984), 43–8; Stewart, D.W. and Furse, D.H., 'Applying psychological measures to marketing and advertising research

problems', in Leigh, J.H. and Martin, C.R. (eds), *Current Issues in Advertising, 1982* (Ann Arbor, MI: University of Michigan Press, 1982), 1–38.

37 Manickas, P.A. and Shea, L.J., 'Hotel complaint behaviour and resolution: a content analysis', *Journal of Travel Research* 36(2) (Fall 1997), 68–73; Morris, R., 'Computerised content analysis in management research: a demonstration of advantages and limitations', *Journal of Management* 20(4) (Winter 1994), 903–31; Kolbe, R.H. and Burnett, M.S., 'Content-analysis research: an examination of applications with directives for improving research reliability and objectivity', *Journal of Consumer Research* 18 (September 1991), 243–50.

38 Berelson, B., *Content Analysis in Communicative Research* (New York: Free Press, 1952).

39 Berstell, G. and Nitterhouse, D., 'Looking outside the box', *Marketing Research: A Magazine of Management and Applications* 9(2) (Summer 1997), 4–13.

40 Denny, M., 'How the omnibus hit the fast track – and now, hold very tight', *ResearchPlus* (February 1996), 4.

41 Worcester, R.M. and Downham, J. (eds), 'ICC/ESOMAR International Code of Marketing and Social Research Practice', *Consumer Market Research Handbook* (Amsterdam: North-Holland, 1986), 813–26.

42 Stafford, M.R. and Stafford, T.E., 'Participant observation and the pursuit of truth: methodological and ethical considerations', *Journal of the Market Research Society* 35 (January 1993), 63–76.

43 Social Research Association and Market Research Society Joint Seminar, 11 September 1997.

44 Chan, H., Lee, R., Dillon, T. and Chang, E., *Electronic Commerce: Fundamentals and Applications* (New York: Wiley, 2001).

45 Taylor, H., 'Does Internet research work?', *International Journal of Market Research* 42(1) (2000), 51–63.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Causal research design: experimentation

Objectives

After reading this chapter, you should be able to:

1 explain the concept of causality as defined in marketing research and distinguish between the ordinary meaning and the scientific meaning of causality;

2 define and differentiate two types of validity: internal validity and external validity;

3 discuss the various extraneous variables that can affect the validity of results obtained through experimentation and explain how the researcher can control extraneous variables;

4 describe and evaluate experimental designs and the differences among pre-experimental, true experimental, quasi-experimental and statistical designs;

5 compare and contrast the use of laboratory versus field experimentation and experimental versus non-experimental designs in marketing research;

6 describe test marketing and its various forms: standard test market, controlled test market and simulated test market;

7 understand the problems of internal and external validity of field experiments when conducted in international markets;

8 describe the ethical issues involved in conducting causal research and the role of debriefing in addressing some of these issues.

HERE NOW

Causality can never be proved; in other words, it can never be demonstrated decisively. Inferences of cause-and-effect relationships are the best that can be achieved.

Overview

We introduced causal designs in Chapter 3, where we discussed their relationship to exploratory and descriptive designs and defined experimentation as the primary method employed in causal designs. This chapter explores the concept of causality further. We identify the necessary conditions for causality, examine the role of validity in experimentation, and consider the extraneous variables and procedures for controlling them. We present a classification of experimental designs and consider specific designs, along with the relative merits of laboratory and field experiments. An application in the area of test marketing is discussed in detail. The considerations involved in conducting experimental research when researching international markets are discussed. Several ethical issues, which arise in experimentation, are identified. We begin with an example that encapsulates the application and process of experimentation.

example

POP buys [1]

The Eckerd Drug Company conducted an experiment to examine the effectiveness of in-store radio advertisements to induce point-of-purchase (POP) buys. Twenty statistically compatible stores were selected based on store size, geographical location, traffic flow count and age. Half of these were randomly selected as test stores, and the other half served as control stores. The test stores aired the radio advertisements, whereas the control stores' POP radio systems were removed. Tracking data in the form of unit sales and turnover were obtained for the following three periods: seven days before the experiment, during the course of the four-week experiment, and seven days after the experiment. The products monitored varied from inexpensive items to small kitchen appliances. Results indicated that sales of the advertised products in the test stores at least doubled. Based on this evidence, Eckerd concluded that in-store radio advertising was highly effective in inducing POP buys, and they decided to continue it. ■

Concept of causality

Experimentation is commonly used to infer causal relationships. The concept of **causality** requires some explanation. The scientific concept of causality is complex. 'Causality' means something very different to the average person on the street than to a scientist.[2] A statement such as 'X causes Y' will have the following meaning to an ordinary person and to a scientist.

Causality
Causality applies when the occurrence of X increases the probability of the occurrence of Y.

Ordinary meaning	Scientific meaning
X is the only cause of Y.	X is only one of a number of possible causes of Y.
X must always lead to Y.	The occurrence of X makes the occurrence of Y more probable (X is a probabilistic cause of Y).
It is possible to prove that X is a cause of Y.	We can never prove that X is a cause of Y. At best, we can infer that X is a cause of Y.

The scientific meaning of causality is more appropriate to marketing research than is the everyday meaning.[3] Marketing effects are caused by multiple variables and the relationship between cause and effect tends to be probabilistic. Moreover, we can never prove causality (i.e. demonstrate it conclusively); we can only infer a cause-and-effect relationship. In other words, it is possible that the true causal relation, if one exists, will not have been identified. We further clarify the concept of causality by discussing the conditions for causality.

Conditions for causality

Before making causal inferences, or assuming causality, three conditions must be satisfied: (1) concomitant variation, (2) time order of occurrence of variables, and (3) elimination of other possible causal factors. These conditions are necessary but not sufficient to demonstrate causality. No one of these three conditions, nor all three conditions combined, can demonstrate decisively that a causal relationship exists.[4] These conditions are explained in more detail in the following sections.

Concomitant variation

Concomitant variation
A condition for inferring causality that requires that the extent to which a cause, X, and an effect, Y, occur together or vary together is predicted by the hypothesis under consideration.

Concomitant variation is the extent to which a cause, X, and an effect, Y, occur together or vary together in the way predicted by the hypothesis under consideration. Evidence pertaining to concomitant variation can be obtained in a qualitative or quantitative manner.

For example, in the qualitative case, the management of a bank may believe that the retention of customers is highly dependent on the quality of service in bank branches. This hypothesis could be examined by assessing concomitant variation. Here, the causal factor X is branch service and the effect factor Y is retention level. A concomitant variation supporting the hypothesis would imply that banks with satisfactory levels of service would also have a satisfactory retention of customers. Likewise, banks with unsatisfactory service would exhibit unsatisfactory retention of customers. If, on the other hand, the opposite pattern was found, we would conclude that the hypothesis was untenable.

For a quantitative example, consider a random survey of 1,000 respondents questioned on the purchase of shares from a bank branch. This survey yields the data in Table 11.1. The respondents have been classified into high- and low-education groups based on a median or even split. This table suggests that the purchase of shares is influenced by education level. Respondents with high education are likely to purchase more shares. Seventy-three per cent of the respondents with high education have a high purchase level, whereas only 64% of those with low education have a high purchase level. Furthermore, this is based on a relatively large sample of 1,000 people.

Table 11.1 Evidence of concomitant variation between purchase of shares and education

		Purchase of shares from a bank, Y		
		High	*Low*	*Total*
Education, X	*High*	363 (73%)	137 (27%)	500 (100%)
	Low	322 (64%)	178 (36%)	500 (100%)

Based on this evidence, can we conclude that high education causes high purchase of shares? Certainly not! All that can be said is that association makes the hypothesis more tenable; it does not prove it. What about the effect of other possible causal factors such as income? Shares can be expensive, so people with higher incomes may buy more of them. Table 11.2 shows the relationship between purchase of shares and education for different income segments. This is equivalent to holding the effect of income constant. Here again, the sample has been split at the median to produce high- and low-income groups of equal size. Table 11.2 shows that the difference in purchase of shares between high- and low-education respondents has been reduced considerably. This suggests that the association indicated by Table 11.1 may be spurious.

Table 11.2 Purchase of shares by income and education

		Low-income purchase		
		High	*Low*	*Total*
Education	*High*	122 (61%)	78 (39%)	200 (100%)
	Low	171 (57%)	129 (43%)	300 (100%)

		High-income purchase		
		High	*Low*	*Total*
Education	*High*	241 (80%)	59 (20%)	300 (100%)
	Low	151 (76%)	49 (24%)	200 (100%)

We could give similar examples to show why the absence of initial evidence of concomitant variation does not imply that there is no causation. It is possible that considering a third variable will crystallise an association that was originally obscure. The time order of the occurrence of variables provides additional insights into causality.

Time order of occurrence of variables

The time order of occurrence condition states that the causing event must occur either before or simultaneously with the effect; it cannot occur afterwards. By definition, an effect cannot be produced by an event that occurs after the effect has taken place. It is possible, however, for each event in a relationship to be both a cause and an effect of the other event. In other words, a variable can be both a cause and an effect in the same causal relationship. To illustrate, customers who shop frequently in a particular supermarket are more likely to have a loyalty card for that supermarket. In addition, customers who have a loyalty card for a supermarket are likely to shop there frequently.

Consider banks and the retention of customers. If in-bank service is the cause of retention, then improvements in service must be made before, or at least simultaneously with, an increase in retention. These improvements might consist of training or hiring more counter staff. Then, in subsequent months, the retention of bank customers should increase. Alternatively, retention may increase simultaneously with the training or hiring of additional counter staff. On the other hand, suppose that a bank experienced an appreciable increase in the level of retaining customers and then decided to use some of that money generated to retrain its counter staff, leading to an improvement in service. In this case, bank service cannot be a cause of increased retention; rather, just the opposite hypothesis might be plausible.

Elimination of other possible causal factors

The absence of other possible causal factors means that the factor or variable being investigated should be the only possible causal explanation. Bank service may be a cause of retention if we can be sure that changes in all other factors affecting retention: pricing, advertising, promotional offers, product characteristics, competition and so forth, were held constant or were otherwise controlled.

In an after-the-fact examination of a situation, we can never confidently rule out all other causal factors. In contrast, with experimental designs it is possible to control some of the other causal factors. It is also possible to balance the effects of some of the uncontrolled variables so that only random variations resulting from these uncontrolled variables will be measured. These aspects are discussed in more detail later in this chapter.

The difficulty of establishing a causal relationship is illustrated by the following example.

example	

Which comes first?[5]

There are studies that contend that consumers increasingly make buying decisions in the store while they are shopping. Some studies indicate that as much as 80% of buying decisions are made at point-of-purchase (POP). POP buying decisions have increased concurrently with increased advertising efforts in the stores. These include radio advertisements, ads on shopping trolleys and grocery bags, ceiling signs and shelf displays. It is difficult to ascertain from these data whether the increased POP decision-making is the result of increased advertising efforts in the store or whether the increase in store advertising results from attempts to capture changing consumer attitudes towards purchasing and to capture sales from the increase in POP decision-making. It is also possible that both variables may be both causes and effects in this relationship. ■

If, as this example indicates, it is difficult to establish cause-and-effect relationships, what is the role of evidence obtained in experimentation? Evidence of concomitant variation, time order of occurrence of variables, and elimination of other possible causal factors, even if combined, still does not demonstrate conclusively that a causal relationship exists. If all the evidence is strong and consistent, however, it may be reasonable to conclude that there is a causal relationship. Accumulated evidence from several investigations increases our confidence that a causal relationship exists. Confidence is further enhanced if the evidence is interpreted in light of intimate conceptual knowledge of the problem situation. Controlled experiments can provide strong evidence on all three conditions.

Definitions and concepts

In this section, we define some basic concepts and illustrate them using examples.

Independent variables
Variables that are manipulated by the researcher and whose effects are measured and compared.

■ **Independent variables**. Independent variables are variables or alternatives that are manipulated (i.e. the levels of these variables are changed by the researcher) and whose effects are measured and compared. These variables, also known as treatments, may include price levels, package designs and advertising themes. In the Eckerd Drug Company example at the beginning of this chapter, the independent variable was 'in-store radio advertising' (present versus absent).

Test units
Individuals, organisations or other entities whose responses to independent variables or treatments are being studied.

■ **Test units**. Test units are individuals, organisations or other entities whose response to the independent variables or treatments is being examined. Test units may include consumers, stores or geographical areas. The test units were stores in the Eckerd example.

Dependent variables
Variables that measure the effect of the independent variables on the test units.

■ **Dependent variables**. Dependent variables are the variables that measure the effect of the independent variables on the test units. These variables may include sales, profits and market shares. In the Eckerd example, the dependent variable was the sales level of advertised point-of-purchase products.

Extraneous variables
Variables, other than dependent and independent variables, which may influence the results of the experiment.

■ **Extraneous variables**. Extraneous variables are all variables other than the independent variables that affect the response of the test units. These variables can confound the dependent variable measures in a way that weakens or invalidates the results of the experiment. In the Eckerd example, store size, geographical location, traffic flow count and age of the stores were extraneous variables that had to be controlled.

Experiment
The process of manipulating one or more independent variables and measuring their effect on one or more dependent variables, while controlling for the extraneous variables.

■ **Experiment**. An experiment is formed when the researcher manipulates one or more independent variables and measures their effect on one or more dependent variables, while controlling for the effect of extraneous variables.[6] The Eckerd research project qualifies as an experiment based on this definition.

■ **Experimental design**. An experimental design is a set of procedures specifying: (1) the test units and how these units are to be divided into homogeneous sub-samples, (2) what independent variables or treatments are to be manipulated, (3) what dependent variables are to be measured, and (4) how the extraneous variables are to be controlled.[7]

As a further illustration of these definitions, consider the following example.

e x a m p l e

Taking coupons at face value[8]

An experiment was conducted to test the effects of the face value of sales promotion coupons (i.e. the amount saved when a consumer next buys the product) on the likelihood of consumers redeeming those coupons, controlling for the frequency of brand usage. Personal interviews were conducted with 280 shoppers who were entering or leaving a supermarket. Subjects were randomly assigned to two treatment groups. One offered low-value coupons and the other high-value coupons for four products: Tide detergent, Kellogg's Cornflakes, Aim toothpaste, and Joy liquid detergent. During the interviews, the respondents answered questions about which brands they used and how likely they were to cash the coupons of the given face value the next time they shopped. An interesting finding was that higher face-value coupons produced a greater likelihood of redemption among infrequent or non-buyers of the promoted brand but had little effect on regular buyers. ■

In the preceding experiment, the independent variable that was manipulated was the value of the coupon. The dependent variable was the likelihood of cashing the coupon. The extraneous variable that was controlled was brand usage. The test units were individual shoppers. The experimental design required the random assignment of test units (shoppers) to treatment groups (low or high value coupon).

Definition of symbols

To facilitate our discussion of extraneous variables and specific experimental designs, we define a set of symbols now commonly used in marketing research.[9]

X = the exposure of a group to an independent variable, treatment or event, the effects of which are to be determined

O = the process of observation or measurement of the dependent variable on the test units or group of units

R = the random assignment of test units or groups to separate treatments

In addition, the following conventions are adopted:

■ Movement from left to right indicates movement through time.
■ Horizontal alignment of symbols implies that all those symbols refer to a specific treatment group.
■ Vertical alignment of symbols implies that those symbols refer to activities or events that occur simultaneously.

For example, the symbolic arrangement

$$X \quad O_1 \quad O_2$$

means that a given group of test units was exposed to the treatment variable (X) and the response was measured at two different points in time O_1 and O_2.

Likewise, the symbolic arrangement

$$R \quad X_1 \quad O_1$$
$$R \quad X_2 \quad O_2$$

means that two groups of test units were randomly assigned to two different treatment groups at the same time, and the dependent variable was measured in the two groups simultaneously.

Validity in experimentation

When conducting an experiment, a researcher has two goals: (1) to draw valid conclusions about the effects of independent variables on the study group, and (2) to make valid generalisations to a larger population of interest. The first goal concerns internal validity, the second external validity.[10]

Internal validity

Internal validity
A measure of accuracy of an experiment. It measures whether the manipulation of the independent variables, or treatments, actually caused the effects on the dependent variable(s).

Internal validity refers to whether the manipulation of the independent variables or treatments actually caused the observed effects on the dependent variables. Thus, internal validity refers to whether the observed effects on the test units could have been caused by variables other than the treatment. If the observed effects are influenced or confounded by extraneous variables, it is difficult to draw valid inferences about the causal relationship between the independent and dependent variables. Internal validity is the basic minimum that must be present in an experiment before any conclusion about treatment effects can be made. Without internal validity, the experimental results are confounded. Control of extraneous variables is a necessary condition for establishing internal validity.

External validity

External validity
A determination of whether the cause-and-effect relationships found in the experiment can be generalised.

External validity refers to whether the cause-and-effect relationships found in the experiment can be generalised. In other words, can the results be generalised beyond the experimental situation, and if so, to what populations, settings, times, independent variables and dependent variables can the results be projected?[11] Threats to external validity arise when the specific set of experimental conditions does not realistically take into account the interactions of other relevant variables in the real world.

It is desirable to have an experimental design that has both internal and external validity, but in applied marketing research we often have to trade one type of validity for another.[12] To control for extraneous variables, a researcher may conduct an experiment in an artificial environment. This enhances internal validity, but it may limit the generalisability of the results, thereby reducing external validity. For example, fast-food chains test customers' preferences for new formulations of menu items in test kitchens. Can the effects measured in this environment be generalised to fast-food outlets that may operate in multifarious environments? (Further discussion on the influence of artificiality on external validity may be found in the section of this chapter on laboratory versus field experimentation.) Regardless of the deterrents to external validity, if an experiment lacks internal validity, it may not be meaningful to generalise the results. Factors that threaten internal validity may also threaten external validity, the most serious of these being extraneous variables.

Extraneous variables

The need to control extraneous variables to establish internal and external validity has already been discussed. In this section, we classify extraneous variables in the following categories: history, maturation, testing effects, instrumentation, statistical regression, selection bias and mortality.

History

History
Specific events that are external to the experiment but that occur at the same time as the experiment.

Contrary to what the name implies, **history** (H) does not refer to the occurrence of events before the experiment. Rather, history refers to specific events that are external to the experiment but that occur at the same time as the experiment. These events may affect the dependent variable. Consider the following experiment:

$$O_1 \quad X_1 \quad O_2$$

where O_1 and O_2 are measures of personal loan applications in a specific region and X_1 represents a new promotional campaign. The difference $(O_2 - O_1)$ is the treatment effect. Suppose that the experiment revealed that there was no difference between O_2 and O_1. Can we then conclude that the promotional campaign was ineffective? Certainly not! The promotional campaign X_1 is not the only possible explanation of the difference between O_2 and O_1. The campaign might well have been effective. What if general economic conditions declined during the experiment and the local area was particularly hard hit by redundancies through several employers closing down their operations (history)? Conversely, even if there was some difference between O_2 and O_1, it may be incorrect to conclude that the campaign was effective if history was not controlled, because the experimental effects might have been confounded by history. The longer the time interval between observations, the greater the possibility that history will confound an experiment of this type.[13]

Maturation

Maturation
An extraneous variable attributable to changes in the test units themselves that occur with the passage of time.

Maturation (MA) is similar to history except that it refers to changes in the test units themselves. These changes are not caused by the impact of independent variables or treatments but occur with the passage of time. In an experiment involving people, maturation takes place as people become older, more experienced, tired, bored or uninterested. Tracking and market studies that span several months are vulnerable to maturation, since it is difficult to know how respondents are changing over time.

Maturation effects also extend to test units other than people. For example, consider the case in which the test units are banks. Banks change over time in terms of personnel, physical layout, decoration, and the range of products and services they have to offer.

Testing effects

Testing effects
Effects caused by the process of experimentation.

Testing effects are caused by the process of experimentation. Typically, these are the effects on the experiment of taking a measure on the dependent variable before and after the presentation of the treatment. There are two kinds of testing effects: (1) main testing effect (MT), and (2) interactive testing effect (IT).

Main testing effect
An effect of testing occurring when a prior observation affects a later observation.

The **main testing effect** (MT) occurs when a prior observation affects a later observation. Consider an experiment to measure the effect of advertising on attitudes towards a brand of beer. The respondents are given a pre-treatment questionnaire measuring background information and attitude towards the brand. They are then exposed to the test commercial embedded in a television programme. After viewing the commercial, the respondents again answer a questionnaire measuring, among other things, attitude towards the beer brand.

265

Suppose that there is no difference between the pre- and post-treatment attitudes. Can we conclude that the commercial was ineffective? An alternative explanation might be that the respondents tried to maintain consistency between their pre- and post-treatment attitudes. As a result of the main testing effect, post-treatment attitudes were influenced more by pre-treatment attitudes than by the treatment itself. The main testing effect may also be reactive, causing the respondents to change their attitudes simply because these attitudes have been measured. The main testing effect compromises the internal validity of the experiment.

Interactive testing effect
An effect in which a prior measurement affects the test unit's response to the independent variable.

In the **interactive testing effect** (IT), a prior measurement affects the test unit's response to the independent variable. Continuing with our beer advertising experiment, when people are asked to indicate their attitudes towards a brand, they become aware of that brand: they are sensitised to that brand and become more likely to pay attention to the test commercial than are people who were not included in the experiment. The measured effects are then not generalisable to the population; therefore, the interactive testing effects influence the experiment's external validity.

Instrumentation

Instrumentation
An extraneous variable involving changes in the measuring instrument, in the observers, or in the scores themselves.

Instrumentation (I) refers to changes in the measuring instrument, in the observers or in the scores themselves. Sometimes measuring instruments are modified during the course of an experiment. In the beer advertising experiment, using a newly designed questionnaire to measure the post-treatment attitudes could lead to variations in the responses obtained. Consider an experiment in which sales at a shoe shop are measured before and after exposure to a promotional offer of a discounted music festival ticket (treatment). A non-experimental price change between O_1 and O_2 results in a change in instrumentation, because sales will be measured using different unit prices. In this case, the treatment effect $(O_2 - O_1)$ could be attributed to a change in instrumentation.

As shown above, instrumentation effects are likely when interviewers make pre- and post-treatment measurements. The effectiveness of interviewers can be different at different times.

Statistical regression

Statistical regression
An extraneous variable that occurs when test units with extreme scores move closer to the average score during the course of the experiment.

Statistical regression (SR) effects occur when test units with extreme scores move closer to the average score during the course of the experiment. In the beer advertising experiment, suppose that in a pre-test measurement some respondents had either very favourable or very unfavourable attitudes towards the brand. On post-treatment measurement, their attitudes might have moved towards the average. Consumer attitudes change continuously for a wide variety of reasons. Consumers with extreme attitudes have more room for change, so variation may be more likely. This has a confounding effect on the experimental results, because the observed effect (change in attitude) may be attributable to statistical regression rather than to the treatment (test commercial).

Selection bias

Selection bias
An extraneous variable attributable to the improper assignment of test units to treatment conditions.

Selection bias (SB) refers to the improper assignment of test units to treatment conditions. This bias occurs when selection or assignment of test units results in treatment groups that differ on the dependent variable before the exposure to the treatment condition. If test units self-select their own groups or are assigned to groups on the basis of the researchers' judgement, selection bias is possible. For example, consider an experiment in which two different displays (old *static display* and new *audio-visual display*) are assigned to different bank branches. The banks in the two

groups may not be equivalent initially. They may vary with respect to an essential characteristic, such as branch size, which is likely to affect the sales of personal loans, regardless of which display was assigned to a bank.

Mortality

Mortality (MO) refers to the loss of test units while the experiment is in progress. This happens for many reasons, such as test units refusing to continue in the experiment. Mortality confounds results because it is difficult to determine whether the lost test units would respond in the same manner to the treatments as those that remain. Consider again the merchandising display experiment. Suppose that during the course of the experiment branch managers in three banks in the new *audio-visual display* drop out because they feel the noise is not conducive to negotiations with certain types of client. The researcher could not determine whether the average sales of the personal loans for the new display would have been higher or lower if these three banks had continued in the experiment.

The various categories of extraneous variables are not mutually exclusive; they can occur jointly and also interact with each other. To illustrate, testing–maturation–mortality refers to a situation in which, because of pre-treatment measurement, the respondents' beliefs and attitudes change over time and there is a differential loss of respondents from the various treatment groups.

Controlling extraneous variables

Extraneous variables represent alternative explanations of experimental results. They pose a serious threat to the internal and external validity of an experiment. Unless they are controlled, they affect the dependent variable and thus confound the results. For this reason, they are also called **confounding variables**. There are four ways of controlling extraneous variables: randomisation, matching, statistical control and design control.

Randomisation

Randomisation refers to the random assignment of test units to experimental groups by using random numbers. Treatment conditions are also randomly assigned to experimental groups. For example, respondents are randomly assigned to one of three experimental groups. One of the three versions of a test commercial, selected at random, is administered to each group. As a result of random assignment, extraneous factors can be represented equally in each treatment condition. Randomisation is the preferred procedure for ensuring the prior equality of experimental groups,[14] but it may not be effective when the sample size is small because it merely produces groups that are equal on average. It is possible, though, to check whether randomisation has been effective by measuring the possible extraneous variables and comparing them across the experimental groups.

Matching

Matching involves comparing test units on a set of key background variables before assigning them to the treatment conditions. In the display experiment, banks could be matched on the basis of turnover, size, proportion of retail to corporate clients, or location. Then one bank from each matched pair would be assigned to each experimental group.

Matching has two drawbacks. First, test units can be matched on only a few characteristics, so the test units may be similar on the variables selected but unequal on others. Second, if the matched characteristics are irrelevant to the dependent variable, then the matching effort has been futile.[15]

Statistical control

Statistical control
A method of controlling extraneous variables by measuring the extraneous variables and adjusting for their effects through statistical methods.

Statistical control involves measuring the extraneous variables and adjusting for their effects through statistical analysis. This was illustrated in Table 11.2, which examined the relationship (association) between purchase of shares and education, controlling for the effect of income. More advanced statistical procedures, such as analysis of covariance (ANCOVA), are also available. In ANCOVA, the effects of the extraneous variable on the dependent variable are removed by an adjustment of the dependent variable's mean value within each treatment condition. ANCOVA is discussed in more detail in Chapter 19.

Design control

Design control
A method of controlling extraneous variables that involves using specific experimental designs.

Design control involves the use of experiments designed to control specific extraneous variables. The types of controls possible by suitably designing the experiment are illustrated with the following example.

example

Experimenting with new products[16]

Controlled-distribution electronic test markets are used increasingly to conduct experimental research on new products. This method makes it possible to create a design that controls for several extraneous factors. The control can allow for the manipulation of variables that can affect the success of new products. In manipulating variables, it is possible to ensure that a new product:

- obtains the right level of supermarket acceptance and all commodity volume distribution,
- is positioned in the correct aisle in each supermarket,
- receives the right number of facings on the shelf,
- has the correct everyday price,
- never has out-of-stock problems, and
- obtains the planned level of trade promotion, display and price features on the desired time schedule.

Pre-experimental designs
Designs that do not control for extraneous factors by randomisation.

True experimental designs
Experimental designs distinguished by the fact that the researcher can randomly assign test units to experimental groups and also randomly assign treatments to experimental groups.

Quasi-experimental designs
Designs that apply part of the procedures of true experimentation yet lack full experimental control.

Statistical designs
Designs that allow for the statistical control and analysis of external variables.

By being able to control these variables, a high degree of internal validity can be obtained. ■

A classification of experimental designs

Experimental designs may be classified as **pre-experimental**, **true experimental**, **quasi-experimental**, and **statistical designs**: see Figure 11.1.

Pre-experimental designs do not employ randomisation procedures to control for extraneous factors. Examples of these designs include the one-shot case study, the one-group pre-test–post-test design, and the static group. In true experimental designs, the researcher can randomly assign test units to experimental groups and treatments to experimental groups. Included in this category are the pre-test–post-test control group design, the post-test-only control group design, and the Solomon four-group design. Quasi-experimental designs result when the researcher is unable to achieve full manipulation of scheduling or allocation of treatments to test units but can still apply part of the apparatus of the experimentation. Two such designs are time series and multiple time series designs. A statistical design is a series of basic experiments that allows for statistical control and analysis of external variables. Statistical designs are classified
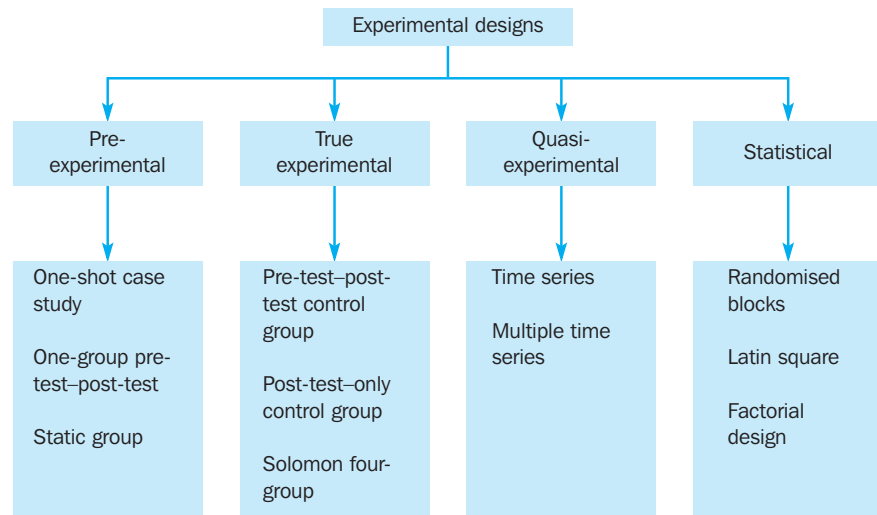
Figure 11.1
A classification of experimental designs

based on their characteristics and use. The important statistical designs include randomised block design, Latin square design and factorial designs.[17]

We begin our discussion with the first type of experimental design: pre-experimental.

Pre-experimental designs

These designs are characterised by an absence of randomisation. Three specific designs are described: the one-shot case study, the one-group pre-test–post-test design, and the static group design.

One-shot case study

One-shot case study
A pre-experimental design in which a single group of test units is exposed to a treatment X, and then a single measurement of the dependent variable is taken.

Also known as the after-only design, the **one-shot case study** may be symbolically represented as

$$X \quad O_1$$

A single group of test units is exposed to a treatment **X**, and then a single measurement on the dependent variable is taken (O_1). There is no random assignment of test units. Note that the symbol **R** is not used, because the test units are self-selected or selected arbitrarily by the researcher.

The danger of drawing valid conclusions from experiments of this type can be easily seen. They do not provide a basis of comparing the level of O_1 with what would happen when **X** was absent. In addition, the level of O_1 might be affected by many extraneous variables, including history, maturation, selection and mortality. Lack of control for these extraneous variables undermines the internal validity. For these reasons, the one-shot case study is more appropriate for exploratory than for conclusive research.

One-group pre-test–post-test design

One-group pre-test–post-test design
A pre-experimental design in which a group of test units is measured twice.

The **one-group pre-test–post-test design** may be symbolised as

$$O_1 \quad X \quad O_2$$

In this design, a group of test units is measured twice. There is no control group. First a pre-treatment measure is taken (O_1), then the group is exposed to the treatment (**X**). Finally, a post-treatment measure is taken (O_2). The treatment effect is computed

as $(O_2 - O_1)$ but the validity of this conclusion is questionable since extraneous variables are largely uncontrolled. History, maturation, testing (both main and interactive testing effects), instrumentation, selection, mortality and regression could possibly be present. The following example shows how this design is used.

Cinematic performance

It is possible to use the one-group pre-test–post-test design to measure the effectiveness of test commercials. Respondents are recruited to central cinema locations in different test cities. At the central location, respondents are first administered a personal interview to measure, among other things, attitudes towards the brand being portrayed in the commercial (O_1). Then they watch a TV programme containing the test commercial (X). After viewing the TV programme, the respondents are again administered a personal interview to measure attitudes towards the same brand (O_2). The effectiveness of the test commercial is measured as $(O_2 - O_1)$. ■

Static group design

Static group

A pre-experimental design in which there are two groups: the experimental group (EG), which is exposed to the treatment, and the control group (CG). Measurements on both groups are made only after the treatment, and test units are not assigned at random.

The **static group** is a two-group experimental design. One group, called the experimental group (EG), is exposed to the treatment, and the other, called the control group (CG), is not. Measurements on both groups are made only after the treatment, and test units are not assigned at random. This design may be symbolically described as

$$\begin{array}{lcc} \text{EG:} & \mathbf{X} & \mathbf{O}_1 \\ \text{CG:} & & \mathbf{O}_2 \end{array}$$

The treatment effect would be measured as $(O_1 - O_2)$. Notice that this difference could also be attributed to at least two extraneous variables (selection and mortality). Because test units are not randomly assigned, the two groups (EG and CG) may differ before the treatment, and selection bias may be present. There may also be mortality effects, as more test units may withdraw from the experimental group than from the control group. This would be particularly likely to happen if the treatment were unpleasant.

In practice, a control group is sometimes defined as the group that receives the current level of marketing activity, rather than a group that receives no treatment at all. The control group is defined in this way because it is difficult to reduce current marketing activities such as advertising and personal selling to zero. We illustrate the static group in the context of the GlobalCash Project.

Static group

A static group comparison to measure the effectiveness of a product placement (a particular character overtly using a branded product or service as part of the story) for a bank, within the context of a feature film, would be conducted as follows. Two groups of respondents would be recruited on the basis of convenience. Only the experimental group would be exposed to the feature film containing the product placement. Then, attitudes towards the bank of both the experimental and control group respondents would be measured. The effectiveness of the product placement would be measured as $(O_1 - O_2)$. ■

True experimental designs

The distinguishing feature of true experimental designs, compared with pre-experimental designs, is randomisation. In true experimental designs, the researcher randomly assigns test units to experimental groups and treatments to experimental groups. True experimental designs include the pre-test–post-test control group design, the post-test-only control group design, and the Solomon four-group design.

Pre-test–post-test control group design

In the **pre-test–post-test control group design**, test units are randomly assigned to either the experimental or the control group and a pre-treatment measure is taken on each group. Then the treatment is applied to the experimental group, and a post-treatment measure is taken from both groups. This design is symbolised as:

$$\text{EG:} \quad R \quad O_1 \quad X \quad O_2$$
$$\text{CG:} \quad R \quad O_3 \quad \quad O_4$$

The treatment effect (TE) is measured as

$$(O_2 - O_1) - (O_4 - O_3)$$

This design controls for most extraneous variables. Selection bias is eliminated by randomisation. The other extraneous effects are controlled as follows:

$$O_2 - O_1 = TE + H + MA + MT + IT + I + SR + MO$$
$$O_4 - O_3 = H + MA + MT + I + SR + MO$$
$$= EV \text{ (extraneous variables)}$$

where the symbols for the extraneous variables are as defined previously. The experimental result is obtained by

$$(O_2 - O_1) - (O_4 - O_3) = TE + IT$$

Interactive testing effect is not controlled, because of the effect of the pre-test measurement on the reaction of units in the experimental group to the treatment.

e x a m p l e

GlobalCash Project

Pre-test–post-test control group

In the context of measuring the effectiveness of a product placement in a feature film for a bank, a pre-test–post-test control group design would be implemented as follows. A sample of respondents would be selected at random. Half of these would be randomly assigned to the experimental group, and the other half would form the control group. Respondents in both groups would be administered a questionnaire to obtain a pre-test measurement on attitudes towards the bank. Only the respondents in the experimental group would be exposed to the feature film containing the product placement. Then, a questionnaire would be administered to respondents in both groups to obtain post-test measures on attitudes towards the bank. ■

As this example shows, the pre-test–post-test control group design involves two groups and two measurements on each group. A simpler design is the post-test-only control group design.

Post-test-only control group design

The **post-test-only control group design** does not involve any pre-measurement. It may be symbolised as

$$\text{EG:} \quad R \quad X \quad O_1$$
$$\text{CG:} \quad R \quad \quad O_2$$

The treatment effect is obtained by

$$O_1 - O_2 = TE$$

This design is fairly simple to implement. Because there is no pre-measurement, the testing effects are eliminated, but this design is sensitive to selection bias and mortality. It is assumed that the two groups are similar in terms of pre-treatment measures on the dependent variable because of the random assignment of test units to groups.

Because there is no pre-treatment measurement, this assumption cannot be checked. This design is also sensitive to mortality. It is difficult to determine whether those in the experimental group who discontinue the experiment are similar to their counterparts in the control group. Yet another limitation is that this design does not allow the researcher to examine changes in individual test units.

It is possible to control for selection bias and mortality through carefully designed experimental procedures. Examination of individual cases is often not of interest. On the other hand, this design possesses significant advantages in terms of time, cost and sample size requirements. It involves only two groups and only one measurement per group. Because of its simplicity, the post-test-only control group design is probably the most popular design in marketing research. Note that, except for pre-measurement, the implementation of this design is very similar to that of the pre-test–post-test control group design.

Solomon four-group design

Solomon four-group design
An experimental design that explicitly controls for interactive testing effects, in addition to controlling for all the other extraneous variables.

In this case, the researcher is not concerned with examining the changes in the attitudes of individual respondents. When this information is desired, the **Solomon four-group design** should be considered. The Solomon four-group design overcomes the limitations of the pre-test–post-test control group and post-test-only control group designs in that it explicitly controls for interactive testing effect, in addition to controlling for all the other extraneous variables. However, this design has practical limitations: it is expensive and time-consuming to implement. Hence, it is not considered further.[18]

In all true experimental designs, the researcher exercises a high degree of control. In particular, the researcher can control when the measurements are taken, on whom they are taken, and the scheduling of the treatments. Moreover, the researcher can randomly select the test units and randomly expose test units to the treatments. In some instances, the researcher cannot exercise this kind of control; then quasi-experimental designs should be considered.

Quasi-experimental designs

A quasi-experimental design results under the following conditions. First, the researcher can control when measurements are taken and on whom they are taken. Second, the researcher lacks control over the scheduling of the treatments and also is unable to expose test units to the treatments randomly.[19] Quasi-experimental designs are useful because they can be used in cases when true experimentation cannot be used, and because they are quicker and less expensive. Because full experimental control is lacking, the researcher must consider the specific variables that are not controlled. Popular forms of quasi-experimental designs are time series and multiple time series designs.

Time series design

Time series design
A quasi-experimental design that involves periodic measurements of the dependent variable for a group of test units. Then the treatment is administered by the researcher or occurs naturally. After the treatment, periodic measurements are continued to determine the treatment effect.

The **time series design** involves a series of periodic measurements on the dependent variable for a group of test units. The treatment is then administered by the researcher or occurs naturally. After the treatment, periodic measurements are continued to determine the treatment effect. A time-series experiment may be symbolised as:

$$O_1 \, O_2 \, O_3 \, O_4 \, O_5 \, O_6 \, O_7 \, O_8 \, O_9 \, O_{10}$$

This is a quasi-experiment, because there is no randomisation of test units to treatments, and the timing of treatment presentation, as well as which test units are exposed to the treatment, may not be within the researcher's control (hence there being no specific **X** symbolised above).

Taking a series of measurements before and after the treatment provides at least partial control for several extraneous variables. Maturation is at least partially controlled, because it would not affect O_5 and O_6 alone but would also influence other observations. By similar reasoning, main testing effect and statistical regression are controlled as well. If the test units are selected randomly or by matching, selection bias can be reduced. Mortality may pose a problem, but it can be largely controlled by paying a premium or offering other incentives to respondents.

The major weakness of the time series design is the failure to control history. Another limitation is that the experiment may be affected by the interactive testing effect because multiple measurements are being made on the test units. Nevertheless, time series designs are useful, as illustrated by this case. The effectiveness of a test commercial (**X**) may be examined by broadcasting the commercial a predetermined number of times and examining the data from a pre-existing test panel. Although the marketer can control the scheduling of the test commercial, it is uncertain when or whether the panel members are exposed to it. The panel members' purchases before, during and after the campaign are examined to determine whether the test commercial has a short-term effect, a long-term effect or no effect.

Multiple time series design

Multiple time series design
A time series design that includes another group of test units to serve as a control group.

The **multiple time series design** is similar to the time series design except that another group of test units is added to serve as a control group. Symbolically, this design may be described as

$$\text{EG:} \quad O_1 \; O_2 \; O_3 \; O_4 \; O_5 \quad X \quad O_6 \; O_7 \; O_8 \; O_9 \; O_{10}$$
$$\text{CG:} \quad O_{11} O_{12} O_{13} O_{14} O_{15} \qquad O_{16} O_{17} O_{18} O_{19} O_{20}$$

If the control group is carefully selected, this design can be an improvement over the simple time series experiment. The improvement lies in the ability to test the treatment effect twice: against the pre-treatment measurements in the experimental group and against the control group. To use the multiple time series design to assess the effectiveness of a commercial, the test panel example would be modified as follows. The test commercial would be shown in only a few of the test cities. Panel members in these cities would make up the experimental group. Panel members in cities where the commercial was not shown would constitute the control group.

Another application of multiple time series design is illustrated in the following example.

e x a m p l e

Splitting commercials shows their strength[20]

A multiple time series design was used to examine the build-up effect of increased advertising. The data were obtained from Burke Marketing Services from a split-cable TV advertising field experiment. In the split-cable system, one group of households was assigned to the experimental panel and an equivalent group was assigned to the control panel. The two groups were matched on demographic variables. Data were collected for 76 weeks. Both panels received the same level of advertising for the first 52 weeks for the brand in question. For the next 24 weeks, the experimental panel was exposed to twice as much advertising as the control panel. The results indicated that the build-up effect of advertising was immediate with a duration of the order of the purchase cycle. Information of this type can be useful in selecting advertising timing patterns (allocating a set of advertising exposures over a specified period to obtain maximum impact). ■

Table 11.3 Potential sources of invalidity of experimental designs

Design	Source of invalidity							
	Internal variables							External variables
	History	*Maturation*	*Testing*	*Instrumentation*	*Regression*	*Selection*	*Mortality*	*Interaction of testing and X*
Pre-experimental designs								
One-shot case study X O	–	–				–	–	
One-group pre-test–post-test design O X O	–	–	–	–	?			–
Static group comparison X O O	+	?	+	+	+	–	–	
True experimental designs								
Pre-test–post-test control group R O X O R O O	+	+	+	+	+	+	+	–
Post-test only control group design R X O R O	+	+	+	+	+	+	+	+
Quasi-experimental designs								
Time series O O O X O O O	–	+	+	?	+	+	+	–
Multiple time series O O O X O O O O O O O O O	+	+	+	+	+	+	–	

Note: A minus sign indicates a definite weakness, a plus sign indicates that the factor is controlled, a question mark denotes a possible source of concern, and a blank means that the factor is not relevant.

In concluding our discussion of pre-experimental, true experimental and quasi-experimental designs, we summarise in Table 11.3 the potential sources of invalidity that may affect each of these designs. In this table, a minus sign indicates a definite weakness, a plus sign indicates that the factor is controlled, a question mark denotes a possible source of concern, and a blank means that the factor is not relevant. It should be remembered that potential sources of invalidity are not the same as actual errors.

Statistical designs

Statistical designs consist of a series of basic experiments that allow for statistical control and analysis of external variables. In other words, several basic experiments are conducted simultaneously. Thus, statistical designs are influenced by the same

sources of invalidity that affect the basic designs being used. Statistical designs offer the following advantages:

1 The effects of more than one independent variable can be measured.
2 Specific extraneous variables can be statistically controlled.
3 Economical designs can be formulated when each test unit is measured more than once.

The most common statistical designs are the randomised block design, the Latin square design and the factorial design.

Randomised block design

Randomised block design
A statistical design in which the test units are blocked on the basis of an external variable to ensure that the various experimental and control groups are matched closely on that variable.

A **randomised block design** is useful when there is only one major external variable – such as sales, store size, or income of the respondent – that might influence the dependent variable. The test units are blocked or grouped on the basis of the external variable. The researcher must be able to identify and measure the blocking variable. By blocking, the researcher ensures that the various experimental and control groups are matched closely on the external variable.

Randomised block design

Let us extend the example examining the effectiveness of a product placement in a feature film for a bank. The purpose of this experiment would be to measure the impact of environmental concern in a film character related to the bank. Suppose that a pan-European bank like ABN-AMRO were to sponsor a film that included shots using their buildings, logos and examples of the way they run their business. They would naturally be concerned that the image portrayed in the film enhanced the corporate image that they wish to project.

To test this, three test film clips, A, B and C, show respectively a character in the film with no environmental concern, some environmental concern, and high environmental concern. Which of these would be the most effective? Management feels that the respondents' evaluation of the product placement will be influenced by the extent of their usage of a bank. So bank usage is identified as the blocking variable and the randomly selected respondents are

Character A from the test film clips, displaying 'no environmental concern' by sneaking up to and scaring pigeons.

classified into four blocks (heavy, medium, light, or non-users of the bank). Respondents from each block are randomly assigned to the treatment groups (test film clips A, B, and C). The results reveal that the some environmental concern commercial (B) was most effective overall (see Table 11.4). ■

Table 11.4 An example of a randomised block design

Block number	Bank usage	Treatment groups		
		Film A	Film B	Film C
1	*High*			
2	*Medium*			
3	*Light*			
4	*None*			

As this example illustrates, in most marketing research situations, external variables such as sales, bank size, bank type, bank location and characteristics of the respondent can influence the dependent variable. Therefore, randomised block designs are generally more useful than completely random designs. Their main limitation is that the researcher can control for only one external variable. When more than one variable must be controlled, the researcher must use Latin square or factorial designs.

Latin square design

Latin square design
A statistical design that allows for the statistical control of two non-interacting external variables in addition to the manipulation of the independent variable.

A **Latin square design** allows the researcher to control statistically two non-interacting external variables as well as to manipulate the independent variable. Each external or blocking variable is divided into an equal number of blocks or levels. The independent variable is also divided into the same number of levels. A Latin square is conceptualised as a table (see Table 11.5), with the rows and columns representing the blocks in the two external variables. The levels of the independent variable are then assigned to the cells in the table. The assignment rule is that each level of the independent variable should appear only once in each row and each column, as shown in Table 11.5.

Table 11.5 An example of a Latin square design

Bank usage	Interest in increasing electronic automation		
	High	*Medium*	*Low*
High	B	A	C
Medium	C	B	A
Light and None	A	C	B

Note: A, B and C denote the three test commercials, which have respectively no environmental concern, some concern and high concern.

Latin square design

To illustrate the Latin square design, suppose that in the previous example, in addition to controlling for bank usage, the researcher also wanted to control for interest in increasing the electronic automation of cash transactions (defined as high, medium or low). To implement a Latin square design, bank usage would also have to be blocked at three rather than four levels, (e.g. by combining the low and non-users into a single block). Assignments of the three test film clips could then be made as shown in Table 11.5. Note that each film clip – A, B or C – appears once, and only once, in each row and each column. ■

Although Latin square designs are popular in marketing research, they are not without limitations. They require an equal number of rows, columns and treatment levels, which is sometimes problematic. Note that, in the above example, the low users and non-patrons had to be combined to satisfy this requirement. In addition, only two external variables can be controlled simultaneously. Latin squares do not allow the researcher to examine interactions of the external variables with each other or with the independent variable. To examine interactions, factorial designs should be used.

Factorial design

Factorial design
A statistical experimental design used to measure the effects of two or more independent variables at various levels and to allow for interactions between variables.

A **factorial design** is used to measure the effects of two or more independent variables at various levels. Unlike the randomised block design and the Latin square, factorial designs allow for interactions between variables.[21] An interaction is said to take place when the simultaneous effect of two or more variables is different from the sum of their separate effects. For example, an individual's favourite drink might be coffee and her favourite temperature level might be cold, but this individual might not prefer cold coffee, leading to an interaction.

A factorial design may also be conceptualised as a table. In a two-factor design, each level of one variable represents a row and each level of another variable represents a column. Multidimensional tables can be used for three or more factors. Factorial designs involve a cell for every possible combination of treatment variables. Suppose that in the previous example, in addition to examining the effect of environmental concern, the researcher was also interested in simultaneously examining the effect of the amount of information about the bank that came over in the film clip. Further, the amount of bank information was also varied at three levels (high, medium and low). As shown in Table 11.6, this would require $3 \times 3 = 9$ cells. The respondents would be randomly selected and randomly assigned to the nine cells. Respondents in each cell would receive a specific treatment combination. For example, respondents in the upper left corner cell would view a film clip that had no environmental concern and low bank information. The results revealed a significant interaction between the two factors or variables. Respondents with a low amount of bank information preferred the high environmental concern film clip (C). Those with a high amount of bank information, however, preferred the no environmental concern film clip (A). Notice that, although Table 11.6 may appear somewhat similar to Table 11.4, the random assignment of respondents and data analysis are very different for the randomised block design and the factorial design.[22]

Table 11.6 **An example of a factorial design**

Amount of bank information	Amount of environmental concern		
	No concern	Some concern	High concern
Low			
Medium			
High			

Another example of a factorial design follows.

example

Price and information are for the dogs[23]

Burke Marketing Research conducted an experiment prior to the launch of a new dog food. They wished to determine the effect of price and competitive brand information on purchase intentions. A two-factor design was used. Price was manipulated to have four levels: one

Sally celebrates being chosen to take part in the experiment prior to the launch of a new dog food.

discount, two parity (or mid-range prices), and one premium. Competitive brand information was varied at two levels: whether or not information on competitive brands was provided. Approximately 240 respondents were randomly assigned to one of eight (4 × 2) treatment conditions. Respondents were asked to indicate their purchase intentions for the new product on a five-point scale. The results indicated that neither price nor competitive brand information had a significant effect on purchase intentions. ■

The main disadvantage of a factorial design is that the number of treatment combinations increases multiplicatively with an increase in the number of variables or levels. In the Burke Marketing Research example, if the price had been manipulated at six levels and competitive brand information at three levels (no information, partial information and full information), the number of cells would have jumped from 8 to 18. All the treatment combinations are required if all the main effects and interactions are to be measured. If the researcher is interested in only a few of the interactions or main effects, fractional factorial designs may be used. As their name implies, these designs consist of only a fraction or portion of the corresponding full factorial design.

Laboratory versus field experiments

Field environment
An experimental location set in actual market conditions.

Laboratory environment
An artificial setting for experimentation in which the researcher constructs the desired conditions.

Experiments may be conducted in a laboratory or field environment. A laboratory environment is an artificial one that the researcher constructs with the desired conditions specific to the experiment. The term **field environment** is synonymous with actual market conditions. Our experiment to measure the effectiveness of a film clip could be conducted in a **laboratory environment** by showing the film in a test cinema. The same experiment could also be conducted in a field environment by running the full test film (rather than clips) in conventional cinemas. The differences between the two environments are summarised in Table 11.7.

Table 11.7 **Laboratory versus field experiments**

Factor	Laboratory	Field
Environment	Artificial	Realistic
Control	High	Low
Reactive error	High	Low
Demand artefacts	High	Low
Internal validity	High	Low
External validity	Low	High
Time	Short	Long
Number of units	Small	Large
Ease of implementation	High	Low
Cost	Low	High

Laboratory experiments have the following advantages over field experiments:

■ The laboratory environment offers a high degree of control because it isolates the experiment in a carefully monitored environment. Therefore, the effects of history can be minimised.

■ A laboratory experiment also tends to produce the same results if repeated with similar subjects, leading to high internal validity.

■ Laboratory experiments tend to use a small number of test units, last for a shorter time, be more restricted geographically, and are easier to conduct than field experiments. Hence, they are generally less expensive as well.

Compared with field experiments, laboratory experiments suffer from some main disadvantages:

■ The artificiality of the environment may cause reactive error in that the respondents react to the situation itself rather than to the independent variable.[24]

Demand artefacts
Responses given because the respondents attempt to guess the purpose of the experiment and respond accordingly.

■ The environment may cause **demand artefacts**, a phenomenon in which the respondents attempt to guess the purpose of the experiment and respond accordingly. For example, while viewing the film clip, the respondents may recall pre-treatment questions about the brand and guess that the commercial is trying to change their attitudes towards the brand.[25]

■ Finally, laboratory experiments are likely to have lower external validity than field experiments. Because a laboratory experiment is conducted in an artificial environment, the ability to generalise the results to the real world may be diminished.

It has been argued that artificiality or lack of realism in a laboratory experiment need not lead to lower external validity. One must be aware of the aspects of the laboratory experiment that differ from the situation to which generalisations are to be made. External validity will be reduced only if these aspects interface with the independent variables explicitly manipulated in the experiment, as is often the case in applied marketing research. Another consideration, however, is that laboratory experiments allow for more complex designs than field experiments. Hence, the researcher can control for more factors or variables in the laboratory setting, which increases external validity.[26]

The researcher must consider all these factors when deciding whether to conduct laboratory or field experiments.[27] Field experiments are less common in marketing research than laboratory experiments, although laboratory and field experiments play complementary roles.[28]

Characteristics and limitations of experimental designs

In Chapter 3, we discussed three types of research designs: exploratory, descriptive and causal. Of these, it may be argued that causal designs are the most appropriate for inferring and measuring cause-and-effect relationships (though not the only way – as may be argued by the adherents to grounded research approaches, introduced in Chapter 6). Although descriptive survey data are often used to provide evidence of 'causal' relationships, these studies do not meet all the conditions required for causality. For example, it is difficult in descriptive studies to establish the prior equivalence of the respondent groups with respect to both the independent and dependent variables. On the other hand, an experiment can establish this equivalence by random assignment of test units to groups. In descriptive research, it is also difficult to establish time order of occurrence of variables. In an experiment, however, the researcher controls the timing of the measurements and the introduction of the treatment. Finally, descriptive research offers little control over other possible causal factors.

We do not wish to undermine the importance of descriptive research designs in marketing research. As mentioned in Chapter 3, descriptive research constitutes the most popular research design in marketing research, and we do not want to imply that it should never be used to examine causal relationships. Indeed, some authors have suggested procedures for drawing causal inferences from descriptive (non-experimental) data.[29] Rather, our intent is to alert the reader to the limitations of descriptive research for examining causal relationships. Likewise, we also want to make the reader aware of the limitations of experimentation.[30]

Experimentation is an important research design that gives the ability to infer causal relationships. However, it has limitations of time, cost and administration of an experiment.

Time

Experiments can be time-consuming, particularly if the researcher is interested in measuring the long-term effects of the treatment, such as the effectiveness of an advertising campaign. Experiments should last long enough so that the post-treatment measurements include most or all of the effects of the independent variables.

Cost

Experiments are often expensive. The requirements of experimental group, control group and multiple measurements significantly add to the cost of research.

Administration

Experiments can be difficult to administer. It may be impossible in measuring human activity to control for the effects of the extraneous variables, particularly in a field environment. Field experiments often interfere with a company's ongoing operations, and obtaining cooperation from the retailers, wholesalers and others involved may be difficult. Finally, competitors may deliberately contaminate the results of a field experiment. These limitations have given rise to the use of grounded theory approaches, especially in developing an understanding of consumer behaviour that is impossible to encapsulate through experiments.

Experimental design application: test marketing

Test marketing
An application of a controlled experiment done in limited, but carefully selected, test markets. It involves a replication of the planned national marketing programme for a product in test markets.

Test markets
A carefully selected part of the marketplace particularly suitable for test marketing.

Standard test market
A test market in which the product is sold through regular distribution channels. For example, no special considerations are given to products simply because they are being test-marketed.

Test marketing, also called market testing, is an application of a controlled experiment conducted in limited but carefully selected parts of the marketplace called **test markets**. It involves a replication of a planned national marketing programme in test markets. Often, the marketing mix variables (independent variables) are varied in test marketing and the sales (dependent variable) are monitored so that an appropriate national marketing strategy can be identified. The two major objectives of test marketing are (1) to determine market acceptance of the product, and (2) to test alternative levels of marketing mix variables. Test-marketing procedures may be classified as standard test markets, controlled and mini-market tests, and simulated test marketing.

Standard test market

In a **standard test market**, test markets are selected and the product is sold through regular distribution channels. Typically, the company's own sales force is responsible for distributing the product. Sales personnel stock the shelves, restock, and take inventory at regular intervals. One or more combinations of marketing mix variables (product, price, distribution and promotional levels) are employed.

Designing a standard test market involves deciding what criteria are to be used for selecting test markets, how many test markets to use, and the duration of the test. Test markets must be carefully selected. In general, the more test markets that can be used, the better. If resources are limited, at least two test markets should be used for each programme variation to be tested. Where external validity is important, however, at least four test markets should be used. The criteria for selection of test markets may be summarised as:[31]

1 Large enough to produce meaningful projections. They should contain at least 2% of the potential target population.
2 Representative demographically.
3 Representative with respect to product consumption behaviour.
4 Representative with respect to media usage.
5 Representative with respect to competition.
6 Relatively isolated in terms of media and physical distribution.
7 Having normal historical development in the product class.
8 Having marketing research and auditing services available.
9 Not overtested.

The duration of the test depends on the repurchase cycle for the product, the probability of competitive response, cost considerations, the initial consumer response, and company philosophy. The test should last long enough for repurchase activity to be observed. This indicates the long-term impact of the product. If competitive reaction to the test is anticipated, the duration should be short. The cost of the test is also an important factor. The longer a test, the more it costs, and at some point the value of additional information is outweighed by its costs. Recent evidence suggests that tests of new brands should run for at least 10 months. An empirical analysis found that the final test market share was reached in 10 months 85% of the time and in 12 months 95% of the time. Test marketing can be very beneficial to a product's successful introduction, as the following example demonstrates.

Bass joins exclusive Czech beer club[32]

Bass has acquired 34% of Staropramen, a Prague brewer. It launched the Czech beer in six-month test markets in Manchester and Liverpool in bottles, and on draft and in bottles in London. The introduction was backed with a comprehensive promotional package designed to encourage consumer trial and future purchase. This included sampling nights, point-of-sale material and glassware. ■

A standard test market, such as the Bass example, constitutes a one-shot case study. In addition to the problems associated with this design, test marketing faces two unique problems. First, competitors often take actions such as increasing their promotional efforts to contaminate the test marketing programme. When Procter & Gamble test-marketed a hand-and-body lotion, the market leader, Cheeseborough Ponds, started a competitive buy-one-get-one-free promotion for its flagship brand, Vaseline Intensive Care lotion. This encouraged consumers to stock up on Vaseline Intensive Care lotion and, as a result, the Procter & Gamble product did poorly in the test market. Procter & Gamble still launched the line nationally. Ponds again countered with the same promotional strategy. Vaseline Intensive Care settled with a market share of 22% while Procter & Gamble achieved but 4%.[33] Another problem is that, while a firm's test marketing is in progress, competitors have an opportunity to beat it to the national market[34].

Sometimes it is not feasible to implement a standard test market using the company's personnel. Instead, the company must seek help from an outside supplier, in which case the controlled test market may be an attractive option.

Controlled test market

Controlled test market
A test-marketing programme conducted by an outside research company in field experimentation. The research company guarantees distribution of the product in retail outlets that represent a predetermined percentage of the market.

In a **controlled test market**, the entire test-marketing programme is conducted by an outside research company. The research company guarantees distribution of the product in retail outlets that represent a predetermined percentage of the market. It handles warehousing and field sales operations, such as stocking shelves, selling and stock control. The controlled test market includes both mini-market (or forced distribution) tests and the smaller controlled store panels. This service is provided by a number of research firms, including A.C. Nielsen.

Simulated test market

Simulated test market
A quasi-test market in which respondents are preselected; they are then interviewed and observed on their purchases and attitudes towards the product.

Also called a laboratory test or test market simulation, a **simulated test market** yields mathematical estimates of market share based on initial reaction of consumers to a new product. The procedure works as follows. Typically, respondents are intercepted in busy locations, such as shopping centres, and pre-screened for product usage. The selected individuals are exposed to the proposed new product concept and given an opportunity to buy the new product in a real-life or laboratory environment. Those who purchase the new product are interviewed about their evaluation of the product and repeat purchase intentions. The trial and repeat-purchase estimates so generated are combined with data on proposed promotion and distribution levels to project a share of the market.[35]

Simulated test markets can be conducted in 16 weeks or fewer. The information they generate is confidential and the competition cannot obtain it. They are also relatively inexpensive. Simulated test markets can cost around 10% of a standard test market.

Determining a test-marketing strategy

The first decision to be made is whether or not to test market the proposed new product, or whatever element of the marketing programme that is under consideration. As shown in Figure 11.2, this decision must take into account the competitive environ-
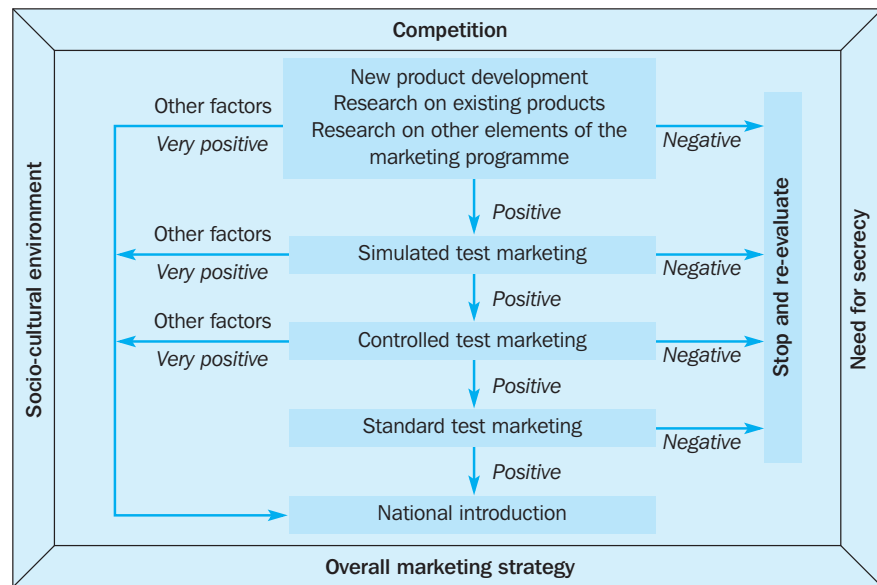
Figure 11.2
Selecting a test-marketing strategy

ment; the socio-cultural environment, particularly consumer preferences and past behaviours; the need to keep the firm's marketing efforts secret; and the overall marketing strategy of the firm. If the marketing research already undertaken to develop the new product provides compelling positive evidence, or if factors such as pre-empting competitive moves dominate, the new product may well be introduced nationally without test marketing. If the decision is to conduct test marketing, however, simulated test marketing may be conducted first, followed by controlled test marketing, then standard test marketing, and, if the results are positive, national introduction. Of course, very positive results at any stage may directly lead to national introduction, circumventing subsequent testing.

International marketing research

If field experiments are difficult to conduct in developed Western economies, the challenge they pose is greatly increased in the international arena. In many countries, the marketing, economic, structural, information and technological environment is not developed to the extent that it is in Europe and the United States. For example, in many countries, TV stations are owned and operated by a government that may place severe restrictions on television advertising. This makes field experiments that manipulate advertising levels extremely difficult. Consider, for example, M & M/Mars, which has set up massive manufacturing facilities in Russia and advertises its sweets on television. Yet, the sales potential has not been realised. Is Mars advertising too much, too little or just enough? Although the answer could be determined by conducting a field experiment that manipulated the level of advertising, such causal research is not feasible given the tight control by the Russian government on television stations. Likewise, the lack of major supermarkets in the Baltic states makes it difficult for Procter & Gamble to conduct field experiments to determine the effect of in-store promotions on the sales of its detergents.

In some countries in Asia, Africa and South America, a majority of the population live in small towns and villages. Yet basic infrastructure such as roads, transportation and warehouse facilities are lacking, making it difficult to achieve

▶

desired levels of distribution. Even when experiments are designed in such countries, it is difficult to control for the time order of occurrence of variables and the absence of other possible causal factors, two of the necessary conditions for causality. Because the researcher has little control over the environment, control of extraneous variables is particularly problematic. Furthermore, it may not be possible to address this problem by adopting the most appropriate experimental design, as environmental constraints may make that design infeasible.

Thus, the internal and external validity of field experiments conducted overseas is generally lower than in Europe and the United States. Although pointing to the difficulties of conducting field experiments in other countries, we do not wish to imply that such causal research cannot or should not be conducted. On the contrary, as the following example indicates, creatively designed field experiments can result in rich findings.

| *example* | ### What you hear is what you get[36] |

PepsiCo's strategy to fight arch-rival Coca-Cola in France was through increased spending on advertisements. Part of Coca-Cola's campaign was to use singers and celebrities such as Rod Stewart, Tina Turner, Gloria Estefan and M.C. Hammer in their commercials as well as publicity tours. Marketing research, however, revealed that overplaying American celebrities was detrimental in the French market. Pepsi thought that this was probably a weakness Coke had because Europeans considered Coca-Cola's marketing effort as 'too American' in Europe. Pepsi, therefore, decided to use taste as a competitive tool. The key was to highlight the product superiority, although not directly, as comparative advertising was prohibited in France. They came up with music as a means of communicating. How did this work?

Research showed that attitude towards the brand was influenced by attitude towards the ad, especially in low-involvement products such as soft drinks. Sweet and melodious music played for Pepsi could transfer good feelings from the music to the Pepsi brand. Similarly, repugnant and undesirable music played for Coke could also transfer from the music to the Coke brand. This mechanism is called classical conditioning. To test these hypotheses, a two-factor experiment was designed. The two factors would be the type of music and the brand preferred, each varied at two levels as shown. A test commercial for each experimental condition was run for a month, with each commercial being played in a different city. At the end of the campaign, central location interviews were used to examine the effect on the brand.

		Brand type	
		Pepsi	*Coke*
Music type	*Good*		
	Bad		

The results of a similar experiment indicated positive effects for 'good' music and negative effects for 'bad' music. Pepsi designed its advertising based on these findings. Subsequently, retail sales in France increased, although it remained in second position, after Coke. ■

Ethics in marketing research

As was explained in Chapter 10, it is often believed that, if respondents are aware of the purpose of a research project, they may give biased responses. In these situations, a deliberate attempt is made by the researcher to disguise the purpose of the research. This is often necessary with experimentation, where disguise is needed to produce valid results. Take, for example, a project conducted to determine the effectiveness of television commercials for a breakfast cereal. The respondents are recruited and brought to a central facility. They are told that they will be watching a television programme on nutrition and then will be asked some questions. Interspersed in the programme is a test commercial for the breakfast cereal as well as commercials for some other products (filler commercials). After viewing the programme and the commercials, the respondents are given a questionnaire to complete. The questionnaire obtains evaluations on the programme content, the test commercial and some of the filler commercials. Note that the evaluations of the programme content and the filler commercials are not of interest but are obtained to reinforce the nature of the disguise. If the respondents knew that the true purpose was to determine the effectiveness of the test commercial, their responses might be biased. Disguising the purpose of the research, however, should not lead to deception.

Although this seems like a paradox, one solution is to disclose the possible existence of deception before the start of the experiment and allow the participants the right to redress at the conclusion of the experiment. The following four items should be conveyed: (1) inform respondents that in an experiment of this nature a disguise of the purpose is often required for valid results; (2) inform them of the general nature of the experiment and what they will be asked to do; (3) make sure they know that they can leave the experiment at any time; and (4) inform them that the study will be fully explained after the data have been gathered and at that time they may request that their information be withdrawn.

Debriefing
After a disguised experiment, informing test subjects what the experiment was about and how the experimental manipulations were performed.

The procedure outlined in item (4) is called **debriefing**. It could be argued that disclosure in this way would also bias results. There is evidence, however, indicating that data collected from subjects informed of the possibility of deception and those not informed are similar.[37] Debriefing can alleviate the stress caused by the experiment and make the experiment a learning experience for the respondents. However, if not handled carefully, debriefing itself can be unsettling to subjects. In the breakfast cereal example above, respondents may find it disheartening that they spent their time evaluating a cereal commercial. The researcher should anticipate and address this issue in the debriefing session.

One further ethical concern in experimentation involves using the appropriate experimental design to control errors caused by extraneous variables. It is the responsibility of the researcher to use the most applicable experimental design for the problem. As the following example illustrates, determining the most appropriate experimental design for the problem requires not only an initial evaluation but also continuous monitoring.

example

Correcting errors early: a stitch in time saves nine[38]

A marketing research firm specialising in advertising research examined the effectiveness of a television commercial for Nike athletic shoes. A one-group pre-test–post-test design was used. Attitudes held by the respondents towards Nike athletic shoes were obtained prior to being exposed to a sports programme on TV and several commercials, including the one for

Nike. Attitudes were again measured after viewing the programme and the commercials. Initial evaluation based on a small sample found the one-group pre-test–post-test design adopted in this study to be susceptible to demand artefacts: respondents attempt to guess the purpose of the experiment and respond accordingly. Because time and financial constraints make redesigning the study difficult at best, the research continued with correction. Continuing a research project after knowing errors were made in the early stages is not ethical behaviour. Experimental design problems should be disclosed immediately to the client. Decisions whether to redesign or accept the flaws should be made jointly. ■

Internet and computer applications

The Internet can be a useful vehicle for conducting causal research. Different experimental treatments can be displayed on different Websites. Respondents can then be recruited to visit these sites and respond to a questionnaire that obtains information on the dependent and extraneous variables. Thus, the Internet can provide a mechanism for controlled experimentation, although in a laboratory type of environment.

An example of testing the effectiveness of advertisements can be used to illustrate the use of the Internet in causal research. Different advertisements can be posted on different Websites. Matched or randomly selected respondents can be recruited to visit these sites, with one group visiting only one site. If any pre-treatment measures have to be obtained, respondents can answer a questionnaire posted on the site. Then they are exposed to a particular advertisement on that site. After viewing the advertisement, the respondents answer additional questions providing post-treatment measures. Control groups can also be implemented in a similar way. Thus, all types of experimental designs that we have considered can be implemented in this manner.

To complement the Internet, microcomputers and mainframe software can be used in the design and analysis of experiments. For example, the statistical analysis package Minitab can be used to design experiments – see **www.minitab.com** for details and to download a demo. Minitab is similar in use to SPSS (**www.spss.com**) or SAS (**www.sas.com**). Minitab includes functions and documentation specifically for industrial-quality control work in which factorial designs are encountered. For example, researchers investigating restaurant atmosphere might want to examine some of the interactions of independent variables. The dependent variable in this experiment could be the respondent's rating of the restaurant as a setting for a romantic meal. Three factors would be included in this $2 \times 2 \times 2$ study. Assuming two lighting levels (i.e. low or medium), two sound types (piped music or live music), and two olfactory stimuli (i.e. spicy smells or sweet/confectionery smells), the best combinations of restaurant atmospherics can be examined.

The marketing research company A. C. Nielsen are major proponents of test-marketing designs. Visit their Website on **www.acnielsen.com**, click on 'test marketing' and review their approaches under the headings of 'Live/In-store New Product Testing, 'Live/In-Store Existing Product Testing', and 'Simulated Product Testing – BASES'.

Summary

The scientific notion of causality implies that we can never prove that *X* causes *Y*. At best, we can only infer that *X* is one of the causes of *Y* in that it makes the occurrence of *Y* probable. Three conditions must be satisfied before causal inferences can be made: (1) concomitant variation, which implies that *X* and *Y* must vary together in a hypothesised way; (2) time order of occurrence of variables, which implies that *X* must precede *Y*; and (3) elimination of other possible causal factors, which implies that competing explanations must be ruled out. Experiments provide the most convincing evidence of all three conditions. An experiment is formed when one or more independent variables are manipulated or controlled by the researcher and their effect on one or more dependent variables is measured.

In designing an experiment, it is important to consider internal and external validity. Internal validity refers to whether the manipulation of the independent variables actually caused the effects on the dependent variables. External validity refers to the generalisability of experimental results. For the experiment to be valid, the researcher must control the threats imposed by extraneous variables, such as history, maturation, testing (main and interactive testing effects), instrumentation, statistical regression, selection bias and mortality. There are four ways of controlling extraneous variables: randomisation, matching, statistical control and design control.

Experimental designs may be classified as pre-experimental, true experimental, quasi-experimental and statistical designs. An experiment may be conducted in a laboratory environment or under actual market conditions in a real-life setting. Only causal designs encompassing experimentation are appropriate for inferring cause-and-effect relationships.

Although experiments have limitations in terms of time, cost and administration, they are becoming increasingly popular in marketing. Test marketing is an important application of experimental design.

The internal and external validity of field experiments conducted in developing nations is generally lower than in the developed Western economies. The level of development in many countries is lower and the researcher lacks control over many of the marketing variables. The ethical issues involved in conducting causal research include disguising the purpose of the experiment. Debriefing can be used to address some of these issues.

Questions

1 What are the requirements for inferring a causal relationship between two variables?

2 Differentiate between internal and external validity.

3 List any five extraneous variables and give an example to show how each can reduce internal validity.

4 Describe the various methods for controlling extraneous sources of variation.

5 What is the key characteristic that distinguishes true experimental designs from pre-experimental designs?

6 List the steps involved in implementing the post-test-only control group design. Describe the design symbolically.

7 What is a time series experiment? When is it used?

8 How is a multiple time series design different from a basic time series design?

9 What advantages do statistical designs have over basic designs?

10 What are the limitations of the Latin square design?

11 Compare the characteristics of laboratory and field experimentation.

12 Should descriptive research be used for investigating causal relationships? Why or why not?

13 What is test marketing? What are the three types of test marketing?

14 What is the main difference between a standard test market and a controlled test market?

15 Describe how simulated test marketing works.

Notes

1 Dwek, R., 'Prediction of success', *Marketing* (POP and Field Marketing Supplement) (17 April 1997), XII–XIII; 'POP radio test airs the ads in-store', *Marketing News* (24 October 1986), 16.

2 Bausall, R.B., *Conducting Meaningful Experiments* (Thousand Oaks, CA: Sage, 1994).

3 For several references on the use of experiments in marketing, see Gardner, D.M. and Belk, R.W., *A Basic Bibliography on Experimental Design in Marketing* (Chicago, IL: American Marketing Association, 1980).

4 Boruch, R.F., *Randomized Experiments for Planning and Evaluation* (Thousand Oaks, CA: Sage, 1994).

5 Witthaus, M., 'POP stars', *Marketing Week* 20(16) (17 July 1997), 37–41.

6 Wyner, G.A., 'Experimental design', *Marketing Research: A Magazine of Management and Applications* 9(3) (Fall 1997), 39–41; Brown, S.R. and Melamed, L.E., *Experimental Design and Analysis* (Newbury Park, CA: Sage, 1990).

7 To see a study employing experimental designs: Spence, M.T. and Brucks, M., 'The moderating effect of problem characteristics on experts' and novices' judgments', *Journal of Marketing Research* 34 (May 1997), 233–47.

8 Love, A.A., 'Companies want to cut coupons, but consumers demand bargains', *Marketing News* 31(10) (12 May 1997), 15; Shoemaker, R.W. and Tibrewala, V., 'Relating coupon redemption rates to past purchasing of the brand', *Journal of Advertising Research* 25 (October–November 1985), 40–7.

9 Banks, S., *Experimentation in Marketing* (New York: McGraw-Hill, 1965), 168–79.

10 In addition to internal and external validity, there also exist construct and statistical conclusion validity. Construct validity addresses the question of what construct, or characteristic, is in fact being measured and is discussed in Chapter 12 on measurement and scaling. Statistical conclusion validity addresses the extent and statistical significance of the covariation that exists in the data and is discussed in the chapters on data analysis. See also: Cook, T.D. and Campbell, D.T., *Quasi-Experimentation* (Chicago, IL: Rand McNally, 1979), 52–3; and Campbell, D.T. and Stanley, J.C., *Experimental and Quasi Experimental Designs for Research* (Chicago, IL: Rand McNally, 1966).

11 Bordia, P., 'Face-to-face computer-mediated communication: a synthesis of the experimental literature', *Journal of Business Communication* 34(1) (January 1997), 99–120; Bowen, D.M., 'Work group research: past strategies and future opportunities', *IEEE Transactions on Engineering Management* 42(1) (February 1995), 30–8; Lynch, J.G., Jr, 'On the external validity of experiments in consumer research', *Journal of Consumer Research* 9 (December 1982), 225–44.

12 Argyris, C., 'Actionable knowledge: design causality in the service of consequential theory', *Journal of Applied Behavioural Science* 32(4) (December 1966), 390-406; Lynch, J.G., Jr, 'The role of external validity in theoretical research', Calder, B.J., Phillips, L.W. and Tybout, A., 'Beyond external validity', and McGrath, J.E and Brinberg, D., 'External validity and the research process', *Journal of Consumer Research* (June 1983), 109–11, 112–14 and 115–24.

13 Hirsch, A.R., 'Effects of ambient odors on slot machine usage in a Las Vegas casino', *Psychology and Marketing* 12(7) (October 1995), 585–94.

14 Durier, C., Monod, H. and Bruetschy, A., 'Design and analysis of factorial sensory experiments with carry-over effects', *Food Quality and Preference* 8(2) (March 1997), 141–9; Nelson, L.S., 'Notes on the use of randomization in experimentation', *Journal of Quality Technology* 28(1) (January 1996), 123–6.

15 Selart, M., 'Structure compatibility and restructuring in judgement and choice', *Organisation Behaviour and Human Decision Processes* 65(2) (February 1996), 106–16; Barker Bausell, R., *Conducting Meaningful Experiments* (Thousand Oaks, CA: Sage, 1994).

16 Spethmann, B., 'Choosing a test market', *Brandweek* 36(19) (8 May 1995), 42–3; Tarshis, A.M., 'Natural sell-in avoids pitfalls of controlled tests', *Marketing News* (24 October 1986), 14.

17 Other experimental designs are also available. See Gunter, B., 'Fundamental issues in experimental design', *Quality Progress* 29(6) (June 1996), 105–13; Winer, R.S., 'Analysis of advertising experiments', *Journal of Advertising Research* (June 1980), 25–31.

18 For an application of the Solomon four-group design, see Mizerski, R.W., Allison, N.K. and Calvert, S., 'A controlled field study of corrective advertising using multiple exposures and a commercial medium', *Journal of Marketing Research* 17 (August 1980), 341–8.

19 Moorman, C., 'A quasi-experiment to assess the consumer and informational determinants of nutrition information-processing activities – the case of the Nutrition Labeling and Education Act', *Journal of Public Policy and Marketing* 15(1) (Spring 1996), 28–44.

20 Lodish, L.M., Abraham, M.M., Livelsberger, J. and Lubetkin, B., 'A summary of fifty-five in-market experimental estimates of the long-term effects of TV advertising', *Marketing Science* (Summer 1995), G133–40; Krishnamurthi, L., Narayan, J. and Raj, S.P., 'Intervention analysis of a field experiment to assess the build-up effect of advertising', *Journal of Marketing Research* 23 (November 1986), 337–45.

21 For applications of factorial designs, see Bryce, W.J., Day, R. and Olney, T.J., 'Commitment approach to motivating community recycling: New Zealand curbside trial', *Journal of Consumer Affairs*, 31(1) (Summer 1997), 27–52; Leclerc, F., Schmitt, B.H. and Dube, L., 'Foreign branding and its effects on product perceptions and attitudes', *Journal of Marketing Research* 31 (May 1994), 263–70.

22 Jones, G.E. and Kavanagh, M.J., 'An experimental examination of the effects of individual and situational factors on unethical behavioural intentions in the workplace', *Journal of Business Ethics* 15(5) (May 1996), 511–23; Bemmaor, A.C. and Mouchoux, D., 'Measuring the short-term effect of in-store promotion and retail advertising on brand sales: a factorial experiment', *Journal of Marketing Research* 28 (May 1991), 202–14.

23 Dickson, J.R. and Wilby, C.P., 'Concept testing with and without product trial', *Journal of Product Innovation Management* 14(2) (March 1997), 117–25; Miller, J.B., Bruvold, N.T. and Kernan, J.B., 'Does competitive-set information affect the results of concept tests?', *Journal of Advertising Research* (April–May 1987), 16–23.

24 Bone, P.F., 'Word-of-mouth effects on short-term and long-term product judgments', *Journal of Business Research* 32(3) (March 1995), 213–23; Barnes, J.H. Jr and Seymour, D.T., 'Experimenter bias: task, tools, and time', *Journal of the Academy of Marketing Science* (Winter 1980), 1–11.

25 Perrien, J., 'Repositioning demand artifacts in consumer research', *Advances in Consumer Research* 24, (1997) 267–71; Shrimp, T.A., Hyatt, E.M. and Snyder, D.J., 'A critical appraisal of demand artifacts in consumer research', *Journal of Consumer Research* 18(3) (December 1991), 272–83.

26 Monden, Y., Akter, M. and Kubo, N., 'Target costing performance based on alternative participation and evaluation method: a laboratory experiment', *Managerial and Decision Economics* 18(2) (March 1997), 113–29.

27 Farris, P.W. and Reibstein, D.J., 'Overcontrol in advertising experiments', *Journal of Advertising Research* (June–July 1984), 37–42.

28 Alston, R.M. and Nowell, C., 'Implementing the voluntary contribution game: a field experiment', *Journal of Economic Behaviour and Organisation* 31(3) (December 1996), 357–68; and Houston, M.J. and Rothschild, M.L., 'Policy-related experiments on information provision: a normative model and explication', *Journal of Marketing Research* (November 1980), 432–49.

29 Spanos, A., 'On theory testing in econometrics: modeling with nonexperimental data', *Journal of Econometrics* 67(1) (May 1995), 189–226; Blalock, H.M. Jr, *Causal Inferences in Non-experimental Research* (Chapel Hill, NC: University of North Carolina Press, 1964).

30 In some situations, surveys and experiments can complement each other and may both be used. For example, the results obtained in laboratory experiments may be further examined in a field survey. See Johnston, W.J. and Kim, K., 'Performance, attribution, and expectancy linkages in personal selling', *Journal of Marketing* 58 (October 1994), 68–81.

31 Romeo, P., 'Testing, testing', *Restaurant Business* 97(2) (15 January 1998), 12. Reprinted with permission from *Marketing News* published by the American Marketing Association, Chicago, 1 March 1985, p. 15.

32 'Bass joins exclusive Czech beer club', *Grocer* 216 (7174) (22 October 1994), 34.

33 Mehegan, S., 'Vaseline ups ante via anti-bacterial', *Brandweek* 38(21) (26 May 1997), 1, 6.

34 'Why new products are bypassing the market test', *Management Today* (October 1995), 12.

35 Hayes, D.J., Shogren, J.F., Fox, J.A. and Kliebenstein, B., 'Test marketing new food products using a multitrial nonhypothetical experimental auction', *Psychology and Marketing* 13(4) (July 1996), 365–79.

36 Crumley, B., 'French cola wars', *Advertising Age* (17 December 1990), 22.

37 Schmitt, B.H., 'Contextual priming of visual information in advertisements', *Psychology and Marketing* 11(1) (January/February 1994), 1–14.

38 Peterson, B., 'Ethics revisited', *Marketing Research: A Magazine of Management and Applications* 8(4) (Winter 1996), 47–8.

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection
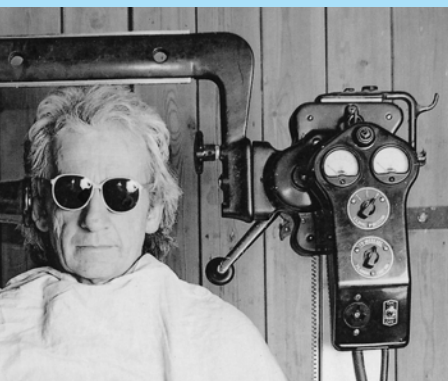
Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Measurement and scaling: fundamentals, comparative and non-comparative scaling

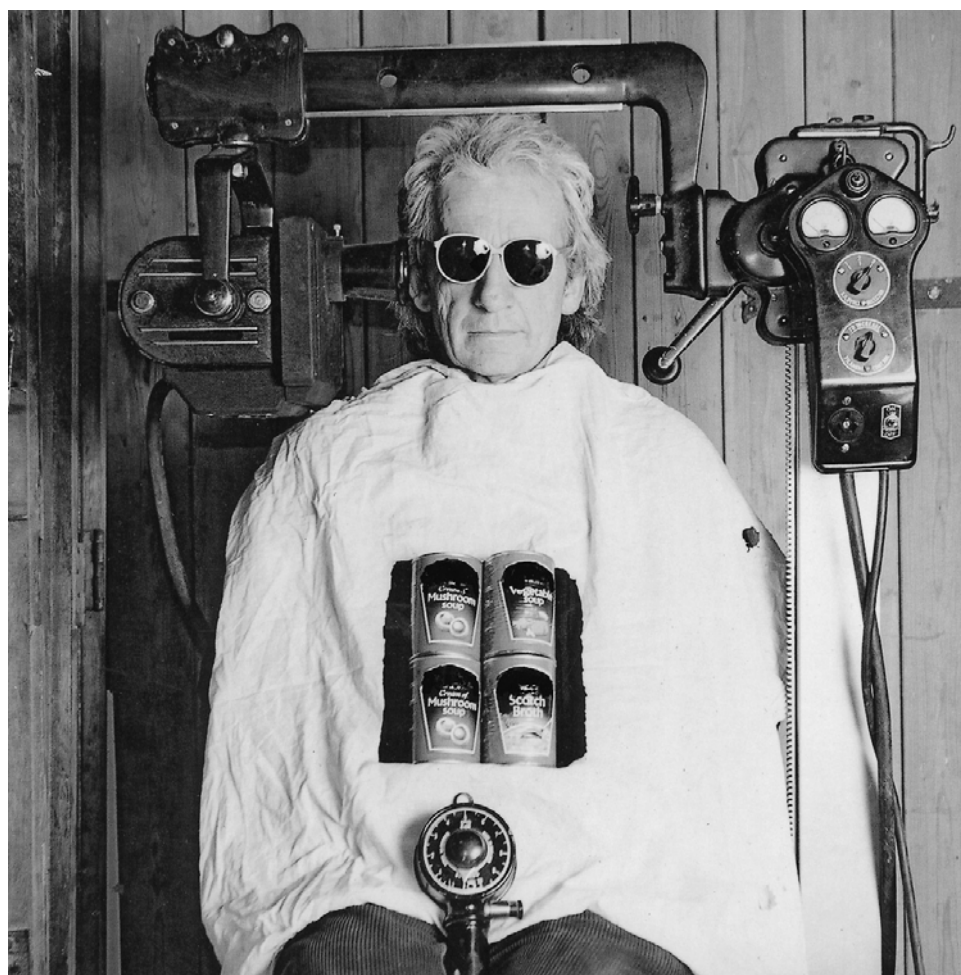## Objectives

After reading this chapter, you should be able to:

1 introduce the concepts of measurement and scaling and show how scaling may be considered an extension of measurement;

2 discuss the primary scales of measurement and differentiate nominal, ordinal, interval and ratio scales;

3 classify and discuss scaling techniques as comparative and non-comparative and describe the comparative techniques of paired comparison, rank order, constant sum and Q-sort scaling;

4 explain the concept of verbal protocols and discuss how they could be employed to measure consumer response to advertising;

5 describe the non-comparative scaling techniques, distinguish between continuous and itemised rating scales, and explain Likert, semantic differential and Stapel scales;

6 discuss the decisions involved in constructing itemised rating scales;

7 discuss the criteria used for scale evaluation and explain how to assess reliability, validity and generalisability;

8 discuss the considerations involved in implementing the primary scales of measurement in an international setting;

9 understand the ethical issues involved in selecting scales of measurement.

'When you can measure what you are speaking about and express it in numbers, you know something about it'. – *Lord Kelvin*

Overview

Once the marketing researcher has a clear understanding of what they wish to understand in their target respondents, they should consider the concepts of scaling and measurement. These concepts are vital in developing questionnaires or 'instruments of measurement' that will fulfil their research objectives in the most accurate manner. This chapter describes the concepts of scaling and measurement and discusses four primary scales of measurement: nominal, ordinal, interval and ratio. We describe and illustrate both comparative and non-comparative scaling techniques in detail. The comparative techniques, consisting of paired comparison, rank order, constant sum and Q-sort scaling, are discussed and illustrated with examples. The non-comparative techniques are composed of continuous and itemised rating scales. We discuss and illustrate the popular itemised rating scales – the Likert, semantic differential and Stapel scales – as well as the construction of multi-item rating scales. We show how scaling techniques should be evaluated in terms of reliability and validity and consider how the researcher selects a particular scaling technique. Mathematically derived scales are also presented. The considerations involved in implementing scaling techniques when researching international markets are discussed. The chapter concludes with a discussion of several ethical issues that arise in scale construction. We begin with an example of how the use of different types of scale can give quite different powers of analysis and interpretation.

Will your instrument of measurement do any better in revealing consumer preferences for soups?

Numbers, rankings and ratings: France is on top

According to the international football federation (FIFA) 2001 end-of-year rankings, world champions France maintained their supremacy at the top with 812 points and Argentina the second spot with 802 points. The top ten countries were as follows:

Number	Country	2001 Ranking	Points
1	Argentina	2	802
2	Brazil	3	793
3	Colombia	5	739
4	England	10	712
5	France	1	812
6	Holland	8	722
7	Italy	6	734
8	Mexico	9	714
9	Portugal	4	741
10	Spain	7	730

Note that the countries have been placed in alphabetical order and that at first glance this gives the impression that South American countries have performed better than European countries. An alphabetical order is used to illustrate the first column 'number'. The 'number' assigned to denote countries is not in any way related to their football playing capabilities but simply serves the purpose of identification, e.g. drawing numbered balls to decide which teams may play each other in a competition. This identification number constitutes a nominal scale, which says nothing about the respective performances of the countries. So whilst Holland is numbered 6 and Portugal is numbered 9, this does not reflect the superior performance of Portugal.

A much clearer way to present the list would be to place the countries in the order of their ranking, placing France at the top and England at the bottom of the table. The ranking would represent an ordinal scale, where it would be clear to see that the lower the number, the better the performance. But what is still missing from the ranking is the magnitude of differences between the countries.

The only way to really understand how much one country is better than another is to examine the points awarded to each country. The points awarded out of 1000 represent an interval scale. Based on the points awarded, note that only two points separate the closely ranked Portugal (741) and Colombia (739), or Mexico (714) and England (712), but that the difference between Brazil (793) ranked at No 3 and Portugal (741) ranked at No 4 is 52 points. ■

Measurement and scaling

Measurement
The assignment of numbers or other symbols to characteristics of objects according to certain pre-specified rules.

Measurement means assigning numbers or other symbols to characteristics of objects according to certain pre-specified rules.[1] We measure not the object but some characteristic of it. Thus, we do not measure consumers, only their perceptions, attitudes, preferences or other relevant characteristics. In marketing research, numbers are usually assigned for one of two reasons. First, numbers permit statistical analysis of the resulting data. Second, numbers facilitate universal communication of measurement rules and results.

The most important aspect of measurement is the specification of rules for assigning numbers to the characteristics. The assignment process must be isomorphic i.e., there must be one-to-one correspondence between the numbers and the characteristics being measured. For example, the same euro (€) figures can be assigned to households with identical annual incomes. Only then can the numbers be associated with specific characteristics of the measured object, and vice versa. In addition, the rules for assigning numbers should be standardised and applied uniformly. They must not change over objects or time.

Scaling may be considered an extension of measurement. Scaling involves creating a continuum upon which measured objects are located. To illustrate, consider a scale for locating consumers according to the characteristic 'attitude towards banks'. Each respondent is assigned a number indicating an unfavourable attitude (measured as 1), a neutral attitude (measured as 2) or a favourable attitude (measured as 3). Measurement is the actual assignment of 1, 2 or 3 to each respondent. Scaling is the process of placing the respondents on a continuum with respect to their attitude towards banks. In our example, scaling is the process by which respondents would be classified as having an unfavourable, neutral or positive attitude.

Primary scales of measurement

There are four primary scales of measurement: nominal, ordinal, interval and ratio.[2] These scales are illustrated in Figure 12.1, and their properties are summarised in Table 12.1 and discussed in the following sections.
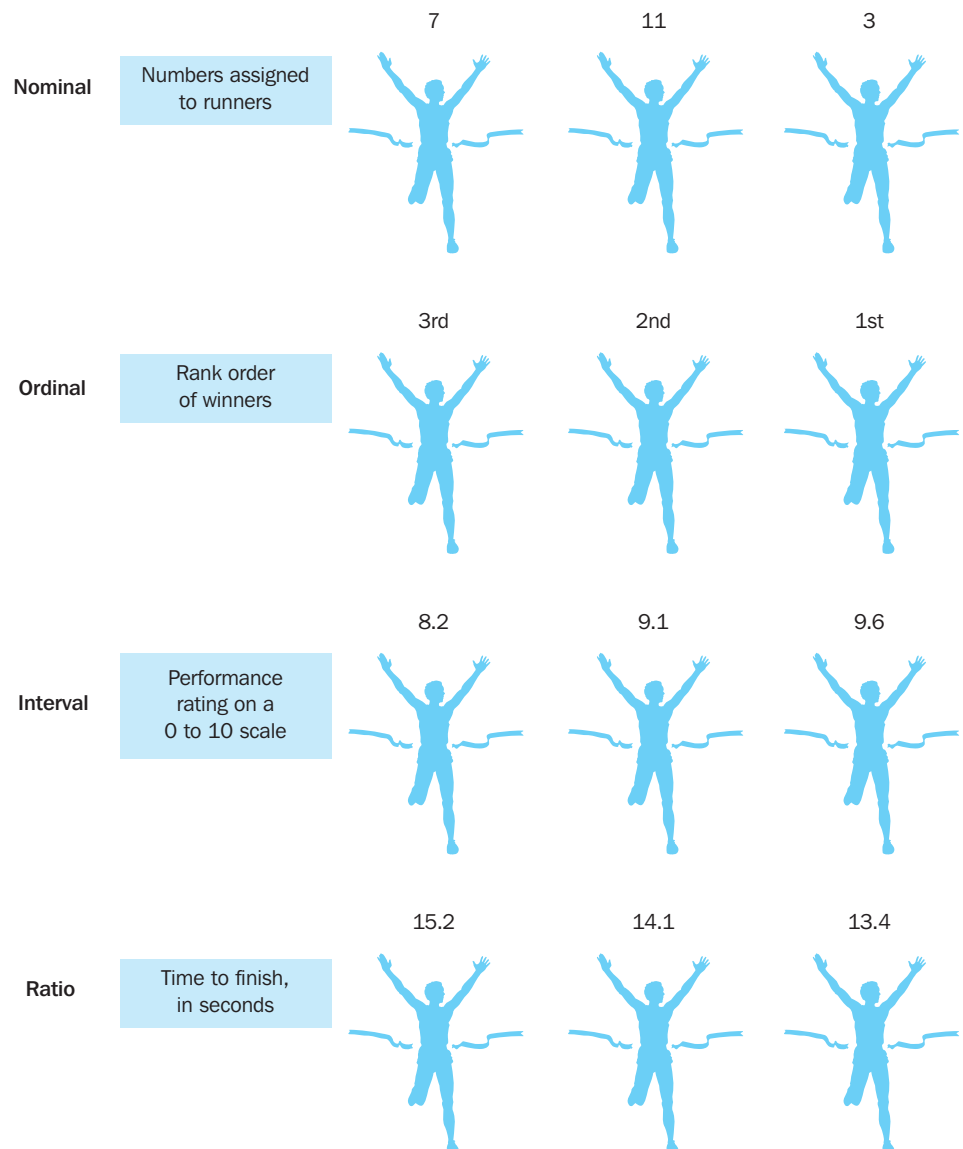
Figure 12.1
An illustration of primary scales of measurement

Table 12.1 Primary scales of measurement

Scale	Basic characteristics	Common examples	Marketing example	Permissible statistics	
				Descriptive	Inferential
Nominal	Numbers identify and classify objects	Student registration numbers, numbers on football players' shirts	Gender classification, bank types	Percentages, mode	Chi-square, binomial test
Ordinal	Numbers indicate the relative positions of the objects but not the magnitude of differences between them	Rankings of the top 4 teams in the football World Cup	Ranking of service quality delivered by a number of banks. Rank order of favourite television programmes	Percentile, median	Rank-order correlation, Friedman ANOVA
Interval	Differences between objects can be compared; zero point is arbitrary	Temperature (Fahrenheit, Celsius)	Attitudes, opinions, index numbers	Range, mean, standard deviation	Product-moment correlations, t-tests, ANOVA, regression, factor analysis
Ratio	Zero point is fixed; ratios of scale values can be computed	Length, weight	Age, income, costs, sales, market shares	Geometric mean, harmonic mean	Coefficient of variation

Nominal scale

Nominal scale
A scale whose numbers serve only as labels or tags for identifying and classifying objects with a strict one-to-one correspondence between the numbers and the objects.

A **nominal scale** is a figurative labelling scheme in which the numbers serve only as labels or tags for identifying and classifying objects. For example, the numbers assigned to the respondents in a study constitute a nominal scale, thus a female respondent may be assigned a number 1 and a male respondent 2. When a nominal scale is used for the purpose of identification, there is a strict one-to-one correspondence between the numbers and the objects. Each number is assigned to only one object, and each object has only one number assigned to it.

Common examples include student registration numbers at their college or university and numbers assigned to football players or jockeys in a horse race. In marketing research, nominal scales are used for identifying respondents, brands, attributes, banks and other objects.

When used for classification purposes, the nominally scaled numbers serve as labels for classes or categories. For example, you might classify the control group as group 1 and the experimental group as group 2. The classes are mutually exclusive and collectively exhaustive. The objects in each class are viewed as equivalent with respect to the characteristic represented by the nominal number. All objects in the same class have the same number, and no two classes have the same number.

The numbers in a nominal scale do not reflect the amount of the characteristic possessed by the objects. For example, a high number on a football player's shirt does not imply that the footballer is a better player than one with a low number or vice versa. The same applies to numbers assigned to classes. The only permissible operation on the numbers in a nominal scale is counting. Only a limited number of statistics, all of which are based on frequency counts, are permissible. These include percentages, mode, chi-square and binomial tests (see Chapter 18). It is not meaningful to compute an average student registration number, the average gender of respondents in a survey, or the number assigned to an average bank, as in the following example.

Nominal scale

In the GlobalCash Project, the numbers 1 through to 250 were assigned to named banks (see extracts from the list in Table 12.2). Thus, bank 48 referred to Credit Lyonnais in France. It did not imply that Credit Lyonnais was in any way superior or inferior to Den Danske Bank, which was assigned the number 54. Any reassignment of the numbers, such as transposing the numbers assigned to Credit Lyonnais and Den Danske Bank, would have no effect on the numbering system, because the numerals did not reflect any characteristics of the banks. It is meaningful to make statements such as '25 per cent of French respondents named Credit Lyonnais as their main bank'. Although the average of the assigned numbers is 50.5, it is not meaningful to state that the number of the average bank is 50.5. ■

Table 12.2 Illustration of primary scales of measurement

No.	Nominal scale	Ordinal scale	Interval scale		Ratio scale
	Bank	Preference rankings	Preference ratings		
			1–7	11–17	
1	ABN AMRO	1	10	7	17	60%
11	Banco Bilbao Vizcaya			4	14	0%
23	Bank Brussels Lambert			5	15	0%
27	Bank of Ireland			7	17	0%
37	Budapest Bank			5	15	0%
44	Citibank	3	50	5	15	30%
48	Credit Lyonnais			6	16	0%
54	Den Danske Bank			6	16	0%
56	Deutsche Bank	2	25	7	17	10%
80	Okobank Finland			2	12	0%

Ordinal scale

Ordinal scale
A ranking scale in which numbers are assigned to objects to indicate the relative extent to which some characteristic is possessed. Thus, it is possible to determine whether an object has more or less of a characteristic than some other object.

An **ordinal scale** is a ranking scale in which numbers are assigned to objects to indicate the relative extent to which the objects possess some characteristic. An ordinal scale allows you to determine whether an object has more or less of a characteristic than some other object, but not how much more or less. Thus, an ordinal scale indicates relative position, not the magnitude of the differences between the objects. The object ranked first has more of the characteristic as compared with the object ranked second, but whether the object ranked second is a close second or a poor second is not known. Common examples of ordinal scales include quality rankings, rankings of teams in a tournament and occupational status. In marketing research, ordinal scales are used to measure relative attitudes, opinions, perceptions and preferences. Measurements of this type include 'greater than' or 'less than' judgements from the respondents.

In an ordinal scale, as in a nominal scale, equivalent objects receive the same rank. Any series of numbers can be assigned that preserves the ordered relationships between the objects.[3] In other words, any monotonic positive (order preserving) transformation of the scale is permissible, since the differences in numbers are void of any meaning other than order (see the following example). For these reasons, in addition to the counting operation allowable for nominal scale data, ordinal scales permit the use of statistics based on centiles. It is meaningful to calculate percentile, quartile, median (Chapter 18), rank-order correlation (Chapter 20) or other summary statistics from ordinal data.

Ordinal scale

Table 12.2 gives a particular respondent's preference rankings. Respondents ranked three banks in order of who they preferred to do business with, showing their 'lead bank', 'second bank' and 'third bank', by assigning a rank 1 to the lead, rank 2 to the second bank, and so on. Note that ABN AMRO (ranked 1) is preferred to Deutsche Bank (ranked 2), but how much it is preferred we do not know. Also, it is not necessary that we assign numbers from 1 to 3 to obtain a preference ranking. The second ordinal scale, which assigns a number 10 to ABN AMRO, 25 to Deutsche and 50 to Citibank, is an equivalent scale, as it was obtained by a monotonic positive transformation of the first scale. The two scales result in the same ordering of the banks according to preference. ■

Interval scale

Interval scale
A scale in which the numbers are used to rank objects such that numerically equal distances on the scale represent equal distances in the characteristic being measured.

In an **interval scale**, numerically equal distances on the scale represent equal values in the characteristic being measured. An interval scale contains all the information of an ordinal scale, but it also allows you to compare the differences between objects. The difference between any two scale values is identical to the difference between any other two adjacent values of an interval scale. There is a constant or equal interval between scale values. The difference between 1 and 2 is the same as the difference between 2 and 3, which is the same as the difference between 5 and 6. A common example in everyday life is a temperature scale. In marketing research, attitudinal data obtained from rating scales are often treated as interval data.[4]

In an interval scale, the location of the zero point is not fixed. Both the zero point and the units of measurement are arbitrary. Hence, any positive linear transformation of the form $y = a + bx$ will preserve the properties of the scale. Here, x is the original scale value, y is the transformed scale value, b is a positive constant, and a is any constant. Therefore, two interval scales that rate objects A, B, C and D as 1, 2, 3 and 4 or as 22, 24, 26 and 28 are equivalent. Note that the latter scale can be derived from the former by using $a = 20$ and $b = 2$ in the transforming equation.

Because the zero point is not fixed, it is not meaningful to take ratios of scale values. As can be seen, the ratio of D to B values changes from 2:1 to 7:6 when the scale is transformed. Yet, ratios of differences between scale values are permissible. In this process, the constants a and b in the transforming equation drop out in the computations. The ratio of the difference between D and B values to the difference between C and B values is 2:1 in both the scales.

Statistical techniques that may be used on interval scale data include all those that can be applied to nominal and ordinal data in addition to the arithmetic mean, standard deviation (Chapter 18), product moment correlations (Chapter 20), and other statistics commonly used in marketing research. Certain specialised statistics such as geometric mean, harmonic mean and coefficient of variation, however, are not meaningful on interval scale data. The GlobalCash example gives a further illustration of an interval scale.

Interval scale

In Table 12.2, a respondent's preferences for conducting any transactions with the 10 named banks were expressed on a seven-point rating scale, where a higher number represents a greater preference for a bank. What is being measured differs from the ordinal scale example. In the ordinal scale example, preference was expressed for the bank a respondent liked to work with. In this example, preference is expressed for banks based upon their ability to handle transactions, a specific element of their working relationship. The most preferred bank might be one that the respondent does not have an account with; it could be one that handles payments from its most profitable customer. We can see that, although Den Danske received a preference rating of 6 and Okobank a rating of 2, this does not mean that Den

Danske is preferred three times as much as Okobank. When the ratings are transformed to an equivalent 11 to 17 scale (next column), the ratings for these banks become 16 and 12, and the ratio is no longer 3:1. In contrast, the ratios of preference differences are identical on the two scales. The ratio of the preference difference between ABN AMRO and Okobank to the preference difference between Budapest Bank and Okobank is 5:3 on both scales. ■

Ratio scale

Ratio scale
The highest scale. This scale allows the researcher to identify or classify objects, rank order the objects, and compare intervals or differences. It is also meaningful to compute ratios of scale values.

A **ratio scale** possesses all the properties of the nominal, ordinal and interval scales, and, in addition, an absolute zero point. Thus, in ratio scales we can identify or classify objects, rank the objects, and compare intervals or differences. It is also meaningful to compute ratios of scale values. Not only is the difference between 2 and 5 the same as the difference between 14 and 17, but also 14 is seven times as large as 2 in an absolute sense. Common examples of ratio scales include height, weight, age and money. In marketing, sales, costs, market share and number of customers are variables measured on a ratio scale.

Ratio scales allow only proportionate transformations of the form $y = bx$, where b is a positive constant. One cannot add an arbitrary constant, as in the case of an interval scale. An example of this transformation is provided by the conversion of yards to feet ($b = 3$). The comparisons between the objects are identical whether made in yards or feet.

All statistical techniques can be applied to ratio data. These include specialised statistics such as geometric mean, harmonic mean and coefficient of variation. The ratio scale is further illustrated in the context of the GlobalCash example.

example

GlobalCash Project

Ratio scale

In the ratio scale illustrated in Table 12.2, respondents were asked to indicate the percentage of business transactions that they conduct with each of the banks they named as lead, second and third banks. Note that since this respondent conducted 60 per cent of their business transactions with ABN AMRO and 10 per cent with Deutsche Bank, this person conducted six times as much business in ABN AMRO compared with Deutsche Bank. Also, the zero point is fixed, since 0 means that the respondent did not do any business with that bank (though their clients may have, hence their exposure to the bank). Note that the rank order of the amount of business conducted with the three banks does not match the rank order of preference of doing business with a particular bank – named as 'lead', 'second' and 'third'. Companies working in a particular region where a bank has expertise may be 'forced' (perhaps by being a subsidiary of a much larger organisation that has made the decision) to conduct business with that bank. They may prefer to do business with another bank, but the percentage of business they conduct with it is low because that bank may be lacking expertise in a particular industry or region. ■

The four primary scales discussed above do not exhaust the measurement level categories. It is possible to construct a nominal scale that provides partial information on order (the partially ordered scale). Likewise, an ordinal scale can convey partial information on distance, as in the case of an ordered **metric scale**. A discussion of these scales is beyond the scope of this text.[5]

Metric scale
A scale that is either interval or ratio in nature.

A comparison of scaling techniques

Comparative scales
One of two types of scaling techniques in which there is direct comparison of stimulus objects with one another.

The scaling techniques commonly employed in marketing research can be classified into comparative and non-comparative scales (see Figure 12.2).

Comparative scales involve the direct comparison of stimulus objects. For example, respondents may be asked whether they prefer Coke or Pepsi. Comparative scale
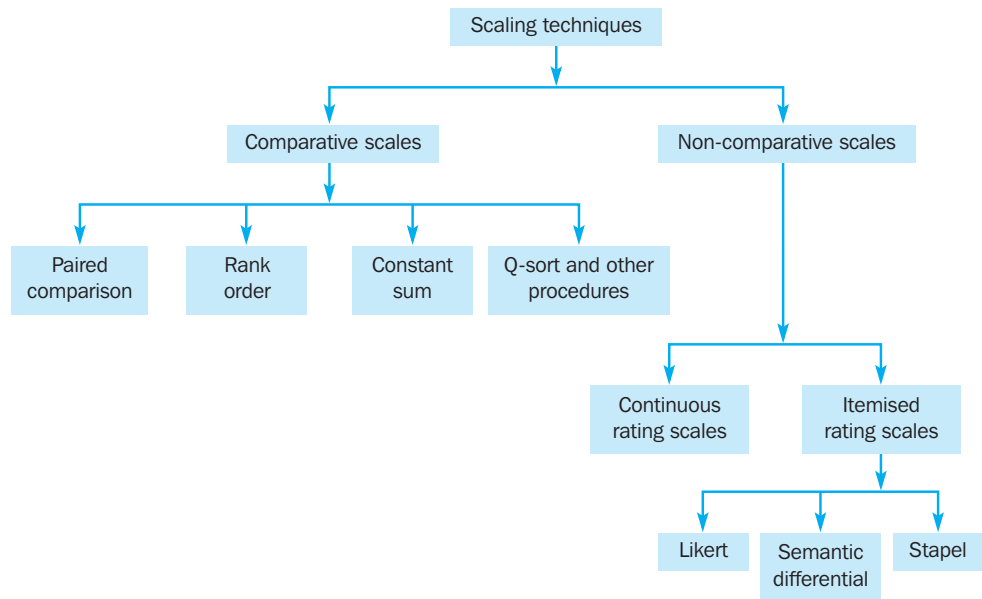
Figure 12.2
A classification of scaling techniques

Non-metric scale
A scale that is either nominal or ordinal in nature.

data must be interpreted in relative terms and have only ordinal or rank order properties. For this reason, comparative scaling is also referred to as **non-metric scaling**. As shown in Figure 12.2, comparative scales include paired comparisons, rank order, constant sum scales, Q-sort and other procedures.

The major benefit of comparative scaling is that small differences between stimulus objects can be detected. As they compare the stimulus objects, respondents are forced to choose between them. In addition, respondents approach the rating task from the same known reference points. Consequently, comparative scales are easily understood and can be applied easily. Other advantages of these scales are that they involve fewer theoretical assumptions, and they also tend to reduce halo or **carryover effects** from

Carryover effects
Where the evaluation of a particular scaled item significantly affects the respondent's judgement of subsequent scaled items.

one judgement to another.[6] The major disadvantages of comparative scales include the ordinal nature of the data and the inability to generalise beyond the stimulus objects scaled. For instance, to compare Virgin Cola with Coke and Pepsi the researcher would have to do a new study. These disadvantages are substantially overcome by the non-comparative scaling techniques.

Non-comparative scales
One of two types of scaling techniques in which each stimulus object is scaled independently of the other objects in the stimulus set. Also called monadic scale.

In **non-comparative scales**, also referred to as monadic or metric scales, each object is scaled independently of the others in the stimulus set. The resulting data are generally assumed to be interval or ratio scaled.[7] For example, respondents may be asked to evaluate Coke on a 1 to 6 preference scale (1 = not at all preferred, 6 = greatly preferred). Similar evaluations would be obtained for Pepsi and Virgin Cola. As can be seen in Figure 12.2, non-comparative scales can be continuous rating or itemised rating scales. The itemised rating scales can be further classified as Likert, semantic differential or Stapel scales. Non-comparative scaling is the most widely used scaling technique in marketing research.

Comparative scaling techniques

Paired comparison scaling

Paired comparison scaling
A comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object in the pair according to some criterion. The data obtained are ordinal in nature.

As its name implies, in **paired comparison scaling** a respondent is presented with two objects and asked to select one according to some criterion.[8] The data obtained are ordinal in nature. A respondent may state that he or she prefers Belgian chocolate to Swiss, likes Kellogg's cereals better than supermarket home brands, or likes Adidas

more than Nike. Paired comparison scales are frequently used when the stimulus objects are physical products. Coca-Cola is reported to have conducted more than 190,000 paired comparisons before introducing New Coke.[9] Paired comparison scaling is the most widely used comparative scaling technique.

Figure 12.3 shows paired comparison data obtained to assess a respondent's bottled beer preferences. As can be seen, this respondent made 10 comparisons to evaluate five brands. In general, with n brands, $[n(n-1)/2]$ paired comparisons include all possible pairings of objects.[10]

Paired comparison data can be analysed in several ways.[11] The researcher can calculate the percentage of respondents who prefer one stimulus over another by summing the matrices of Figure 12.3 for all the respondents, dividing the sum by the number of respondents, and multiplying by 100. Simultaneous evaluation of all the stimulus objects is also possible. Under the assumption of transitivity, it is possible to convert paired comparison data to a rank order.

Transitivity of preference implies that if brand A is preferred to B, and brand B is preferred to C, then brand A is preferred to C. To arrive at a rank order, the researcher determines the number of times each brand is preferred by summing the column entries in Figure 12.3. Therefore, this respondent's order of preference, from most to least preferred, is Carlsberg, Holsten, Stella Artois, Budvar and Grolsch. It is also possible to derive an interval scale from paired comparison data using the Thurstone case V procedure. Refer to the appropriate literature for a discussion of this procedure.[12]

Several modifications of the paired comparison technique have been suggested. One involves the inclusion of a neutral/no difference/no opinion response. Another extension is graded paired comparisons. In this method, respondents are asked which brand in the pair is preferred and how much it is preferred. The degree of preference may be expressed by how much more the respondent is willing to pay for the preferred brand. The resulting scale is a euro metric scale. Another modification of paired comparison scaling is widely used in obtaining similarity judgements in multidimensional scaling (see Chapter 24).

Paired comparison scaling is useful when the number of brands is limited, since it requires direct comparison and overt choice. With a large number of brands, however, the number of comparisons becomes unwieldy. Other disadvantages are that viola-

Transitivity of preference
An assumption made to convert paired comparison data to rank order data. It implies that if brand A is preferred to brand B, and brand B is preferred to brand C, then brand A is preferred to brand C.

Instructions
We are going to present you with ten pairs of bottled beer brands. For each pair, please indicate which of the two brands of beer in the pair you prefer.

Recording form

	Holsten	Stella Artois	Grolsch	Carlsberg	Budvar
Holsten		0	0	1	0
Stella Artois	1[a]		0	1	0
Grolsch	1	1		1	1
Carlsberg	0	0	0		0
Budvar	1	1	0	1	
Number of times preferred[b]	3	2	0	4	1

[a] 1 in a particular box means that the brand in that column was preferred over the brand in the corresponding row. A 0 means that the row brand was preferred over the column brand.

[b] The number of times a brand was preferred is obtained by summing the 1s in each column.

Figure 12.3
Obtaining bottled beer preferences using paired comparisons

tions of the assumption of transitivity may occur, and the order in which the objects are presented may bias the results.[13] Paired comparisons bear little resemblance to the marketplace situation, which involves selection from multiple alternatives. Also respondents may prefer one object over certain others, but they may not like it in an absolute sense.

Rank order scaling

Rank order scaling
A comparative scaling technique in which respondents are presented with several objects simultaneously and asked to order or rank them according to some criterion.

After paired comparisons, the most popular comparative scaling technique is **rank order scaling**. In rank order scaling respondents are presented with several objects simultaneously and asked to order or rank them according to some criterion. For example, respondents may be asked to rank brands of cars according to overall preference. As shown in Figure 12.4, these rankings are typically obtained by asking the respondents to assign a rank of 1 to the most preferred brand, 2 to the second most preferred, and so on, until a rank of n is assigned to the least preferred brand. Like paired comparison, this approach is also comparative in nature, and it is possible that the respondent may dislike the brand ranked 1 in an absolute sense. Furthermore, rank order scaling also results in ordinal data. See Table 12.2, which uses rank order scaling to derive an ordinal scale.

Rank order scaling is commonly used to measure attributes of products and services as well as preferences for brands. Rank order data are frequently obtained from respondents in conjoint analysis (see Chapter 24), since rank order scaling forces the respondent to discriminate among the stimulus objects. Moreover, compared with paired comparisons, this type of scaling process more closely resembles the shopping environment. It also takes less time and eliminates intransitive responses. If there are n stimulus objects, only $(n-1)$ scaling decisions need be made in rank order scaling. However, in paired comparison scaling, $[n(n-1)/2]$ decisions would be required. Another advantage is that most respondents easily understand the instructions for ranking. The major disadvantage is that this technique produces only ordinal data.

Finally, under the assumption of transitivity, rank order data can be converted to equivalent paired comparison data, and vice versa. This point was illustrated by

Instructions

Rank the various brands of car in order of preference. Begin by picking out the one brand that you like most and assign it a number 1. Then find the second most preferred brand and assign it a number 2. Continue this procedure until you have ranked all the brands of car in order of preference. The least preferred brand should be assigned a rank of 10.

No two brands should receive the same rank number.

The criterion of preference is entirely up to you. There is no right or wrong answer. Just try to be consistent.

	Brand	Rank order
1	Porsche	
2	Jaguar	
3	BMW	
4	Bristol	
5	Aston Martin	
6	Mercedes	
7	McLaren	
8	Ferrari	
9	Lamborghini	
10	Bentley	

Figure 12.4
Preference for car brands using rank order scaling

examining the 'Number of times preferred' in Figure 12.3. Hence, it is possible to derive an interval scale from rankings using the Thurstone case V procedure. Other approaches for deriving interval scales from rankings have also been suggested.[14]

Constant sum scaling

Constant sum scaling
A comparative scaling technique in which respondents are required to allocate a constant sum of units such as points, euros, chits, stickers or chips among a set of stimulus objects with respect to some criterion.

In **constant sum scaling**, respondents allocate a constant sum of units, such as points or euros, among a set of stimulus objects with respect to some criterion. As shown in Figure 12.5, respondents may be asked to allocate 100 points to attributes of bottled beers in a way that reflects the importance they attach to each attribute. If an attribute is unimportant, the respondent assigns it zero points. If an attribute is twice as important as some other attribute, it receives twice as many points. The sum of all the points is 100. Hence the name of the scale.

The attributes are scaled by counting the points assigned to each one by all the respondents and dividing by the number of respondents. These results are presented for three groups, or segments, of respondents in Figure 12.5. Segment I attaches overwhelming importance to price. Segment II considers a high alcoholic level to be of prime importance. Segment III values bitterness, hop flavours, fragrance and the aftertaste. Such information cannot be obtained from rank order data unless they are transformed into interval data. Note that the constant sum also has an absolute zero; 10 points are twice as many as 5 points, and the difference between 5 and 2 points is the same as the difference between 57 and 54 points. For this reason, constant sum scale data are sometimes treated as metric. Although this may be appropriate in the limited context of the stimuli scaled, these results are not generalisable to other stimuli not included in the study. Hence, strictly speaking, the constant sum should be considered an ordinal scale because of its comparative nature and the resulting lack of generalisability. It can be seen that the allocation of points in Figure 12.5 is influenced by the specific attributes included in the evaluation task.

The main advantage of the constant sum scale is that it allows for fine discrimination among stimulus objects without requiring too much time. It has two primary

---

Instructions

Below are eight attributes of bottled beers. Please allocate 100 points among the attributes so that your allocation reflects the relative importance you attach to each attribute. The more points an attribute receives, the more important an attribute is. If an attribute is not at all important, assign it no points. If an attribute is twice as important as some other attribute, it should receive twice as many points.

Form

AVERAGE RESPONSES OF THREE SEGMENTS			
	Attribute	Segment I	Segment II	Segment III
1	Bitterness	8	2	17
2	Hop flavours	2	4	20
3	Fragrance	3	9	19
4	Country where brewed	9	17	4
5	Price	53	5	7
6	High alcohol level	7	60	9
7	Aftertaste	5	0	15
8	Package design	13	3	9
	Sum	100	100	100

Figure 12.5
Importance of bottled beer attributes using a constant sum scale

disadvantages, however. Respondents may allocate more or fewer units than those specified. For example, a respondent may allocate 108 or 94 points. The researcher must modify such data in some way or eliminate this respondent from analysis. Another potential problem is rounding error if too few units are used. On the other hand, the use of a large number of units may be too taxing on the respondent and cause confusion and fatigue.

Q-sort and other procedures

Q-sort scaling was developed to discriminate among a relatively large number of objects quickly. This technique uses a rank order procedure in which objects are sorted into piles based on similarity with respect to some criterion. For example, respondents are given 100 attitude statements on individual cards and asked to place them into 11 piles, ranging from 'most highly agreed with' to 'least highly agreed with'. The number of objects to be sorted should not be less than 60 nor more than 140; a reasonable range is 60 to 90 objects.[15] The number of objects to be placed in each pile is pre-specified, often to result in a roughly normal distribution of objects over the whole set.

Another comparative scaling technique is magnitude estimation.[16] In this technique, numbers are assigned to objects such that ratios between the assigned numbers reflect ratios on the specified criterion. For example, respondents may be asked to indicate whether they agree or disagree with each of a series of statements measuring attitude towards banks. Then they assign a number between 0 to 100 to each statement to indicate the intensity of their agreement or disagreement. Providing this type of number imposes a cognitive burden on the respondents.

Another particularly useful procedure (that could be viewed as a very structured combination of observation and depth interviewing) for measuring cognitive responses or thought processes consists of **verbal protocols**. Respondents are asked to 'think out loud' and verbalise anything going through their heads while making a decision or performing a task.[17] The researcher says 'If you think anything, say it aloud, no matter how trivial the thought may be.' Even with such an explicit instruction, the respondent may be silent. At these times, the researcher will say 'Remember to say aloud everything you are thinking.' Everything that the respondent says is tape recorded. This record of the respondent's verbalised thought processes is referred to as a protocol.[18]

Protocols have been used to measure consumers' cognitive responses in actual shopping trips as well as in simulated shopping environments. An interviewer accompanies the respondent and holds a microphone into which the respondent talks. Protocols, thus collected, have been used to determine the attributes and cues used in making purchase decisions, product usage behaviour, and the impact of the shopping environment on consumer decisions. Protocol analysis has also been employed to measure consumer response to advertising. Immediately after seeing an ad, the respondent is asked to list all the thoughts that came to mind while watching the ad. The respondent is given a limited amount of time to list the thoughts so as to minimise the probability of collecting thoughts generated after, rather than during, the message. After the protocol has been collected, the individual's thoughts or cognitive responses can be coded into three categories as illustrated in Table 12.3.[19]

Table 12.3 **Coded verbal protocols**

Category	Definition	Example
Support argument	Support the claim made by the message	'Diet Coke tastes great'
Counter-argument	Refute the claim made by the message	'Diet Coke has an aftertaste'
Source derogation	Negative opinion about the source of the message	'Coca-Cola is not an honest company'

Protocols are, typically, incomplete. The respondent has many thoughts that she or he cannot or will not verbalise. The researcher must take the incomplete record and infer from it a measure of the underlying cognitive response.

Non-comparative scaling techniques

Respondents using a non-comparative scale employ whatever rating standard seems appropriate to them. They do not compare the object being rated either with another object or to some specified standard, such as 'your ideal brand'. They evaluate only one object at a time; thus, non-comparative scales are often referred to as monadic scales. Non-comparative techniques consist of continuous and itemised rating scales, which are described in Table 12.4 and discussed in the following sections.

Table 12.4 **Basic non-comparative scales**

Scale	Basic characteristics	Examples	Advantages	Disadvantages
Continuous rating scale	Place a mark on a continuous line	Reaction to TV commercials	Easy to construct	Scoring can be cumbersome unless computerised
Itemised rating scales				
Likert scale	Degree of agreement on a 1 (strongly disagree) to 5 (strongly agree) scale	Measurement of attitudes	Easy to construct, administer and understand	More time-consuming
Semantic differential scale	Seven-point scale with bipolar labels	Brand product and company images	Versatile	Controversy as to whether the data are interval
Stapel scale	Unipolar 10-point scale, −5 to +5, without a neutral point (zero)	Measurement of attitudes and images	Easy to construct, administered over phone	Confusing and difficult to apply

Continuous rating scale

In a **continuous rating scale**, also referred to as a graphic rating scale, respondents rate the objects by placing a mark at the appropriate position on a line that runs from one extreme of the criterion variable to the other. Thus, the respondents are not restricted to selecting from marks previously set by the researcher. The form of the continuous scale may vary considerably. For example, the line may be vertical or horizontal; scale points, in the form of numbers or brief descriptions, may be provided; and if provided, the scale points may be few or many. Three versions of a continuous rating scale are illustrated in Figure 12.6.

Once the respondent has provided the ratings, the researcher divides the line into as many categories as desired and assigns scores based on the categories into which the ratings fall. In Figure 12.6, the respondent exhibits a favourable attitude towards Dresdner. These scores are typically treated as interval data. The advantage of continuous scales is that they are easy to construct; however, scoring is cumbersome and unreliable.[20] Moreover, continuous scales provide little new information. Hence, their use in marketing research has been limited. Recently, however, with the increased

How would you rate Dresdner Bank in handling pan-European cash management transactions?

Version 1
Probably the worst ~~~~~~~~~~~~~~~~~~~~~~~~✓~~~~~~~~~Probably the best

Version 2
Probably the worst ~~~~~~~~~~~~~~~~~~~~~~~~✓~~~~~~~~~Probably the best
               0   10   20   30   40   50   60   70   80   90   100

Version 3
                    Very bad   *Neither good nor bad*   *Very good*
Probably the worst ~~~~~~~~~~~~~~~~~~~~~~~~✓~~~~~~~~~Probably the best
               0   10   20   30   40   50   60   70   80   90   100
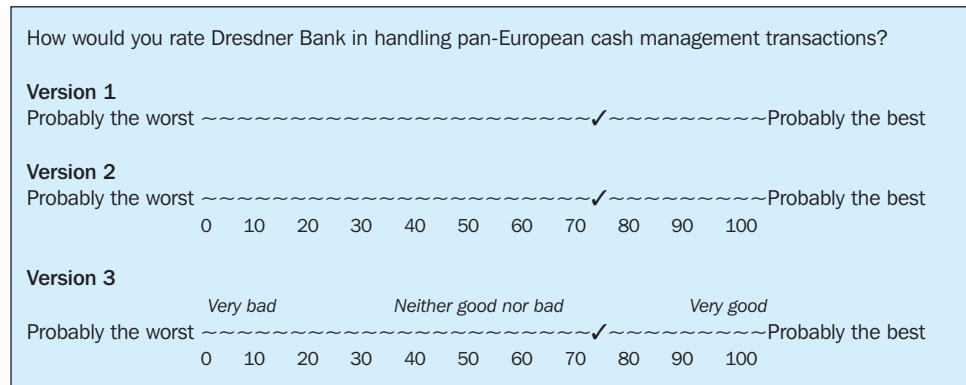
Figure 12.6
Continuous rating scale

Note: In the forthcoming examples using Dresdner Bank, the scores presented do not represent actual measurements taken in GlobalCash.

popularity of computer-assisted personal interviewing and other technologies, their use has become more frequent.[21]

Itemised rating scales

Itemised rating scale
A measurement scale having numbers or brief descriptions associated with each category. The categories are ordered in terms of scale position.

In an **itemised rating scale**, respondents are provided with a scale that has a number or brief description associated with each category. The categories are ordered in terms of scale position; and the respondents are required to select the specified category that best describes the object being rated. Itemised rating scales are widely used in marketing research and form the basic components of more complex scales, such as multi-item rating scales. We first describe the commonly used itemised rating scales – the Likert, semantic differential and Stapel scales – and then examine the major issues surrounding the use of itemised rating scales.

Likert scale

Likert scale
A measurement scale with five response categories ranging from 'strongly disagree' to 'strongly agree' that requires respondents to indicate a degree of agreement or disagreement with each of a series of statements related to the stimulus objects.

Named after its developer, Rensis Likert, the **Likert scale** is a widely used rating scale that requires the respondents to indicate a degree of agreement or disagreement with each of a series of statements about the stimulus objects.[22] Typically, each scale item has five response categories, ranging from 'strongly disagree' to 'strongly agree'. We illustrate with a Likert scale for evaluating attitudes toward Dresdner Bank.

To conduct the analysis, each statement is assigned a numerical score, ranging either from –2 to +2 or from 1 to 5. The analysis can be conducted on an item-by-item basis (profile analysis), or a total (summated) score can be calculated for each respondent by summing across items. Suppose that the Likert scale in Figure 12.7 was used to measure attitudes towards the Bank of Ireland as well as Dresdner. Profile analysis would involve comparing the two banks in terms of the average respondent ratings for each item, such as level of electronic banking support, level of transaction detail, and service levels. The summated approach is most frequently used, and as a result, the Likert scale is also referred to as a summated scale.[23] When using this approach to determine the total score for each respondent on each bank, it is important to use a consistent scoring procedure so that a high (or low) score consistently reflects a favourable response. This requires that the categories assigned to the negative statements by the respondents be scored by reversing the scale. Note that for a negative statement (e.g. Dresdner has poor customer operational support), an agreement reflects an unfavourable response, whereas for a positive statement (e.g. Dresdner delivers high quality banking services), agreement represents a favourable

response. Accordingly, a 'strongly agree' response to a favourable statement and a 'strongly disagree' response to an unfavourable statement would both receive scores of 5.[24] In the example in Figure 12.7, if a higher score is to denote a more favourable attitude, the scoring of items 2, 4, 5 and 7 will be reversed. The respondent to this set of statements has an attitude score of 26. Each respondent's total score for each bank is calculated. A respondent will have the most favourable attitude towards the bank with the highest score. The procedure for developing summated Likert scales is described later in the section on the development and evaluation of scales.

The Likert scale has several advantages. It is easy to construct and administer, and respondents readily understand how to use the scale, making it suitable for Internet surveys, mail, telephone or personal interviews. The major disadvantage of the Likert scale is that it takes longer to complete than other itemised rating scales because respondents have to read and fully reflect upon each statement.

Semantic differential scale

Semantic differential
A seven-point rating scale with end points associated with bipolar labels.

The **semantic differential** is a seven-point rating scale with end points associated with bipolar labels that have semantic meaning. In a typical application, respondents rate objects on a number of itemised, seven-point rating scales bounded at each end by one of two bipolar adjectives, such as 'cold' and 'warm'.[25] We illustrate this scale in Figure 12.8 by presenting a respondent's evaluation of Dresdner on five attributes.

The respondents mark the blank that best indicates how they would describe the object being rated.[26] Thus, in our example, Dresdner Bank is evaluated as somewhat

Instructions

Listed below are different opinions about Dresdner Bank. Please indicate how strongly you agree or disagree with each by putting a tick next to your choice on the following scale:

1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1 Dresdner delivers high quality banking services	1	2✓	3	4	5
2 Dresdner has poor customer operational support	1	2✓	3	4	5
3 I prefer to conduct transactions with Dresdner	1	2	3✓	4	5
4 Dresdner does not have a good European branch network	1	2	3	4✓	5
5 The electronic banking security at Dresdner is terrible	1	2	3	4✓	5
6 Account managers at Dresdner display great knowledge of cash management	1✓	2	3	4	5
7 I do not like Dresdner advertisements	1	2	3	4✓	5
8 Dresdner has good published quality standards	1	2	3	4✓	5
9 Dresdner provides credit on excellent terms	1	2✓	3	4	5

Figure 12.7
The Likert scale

Instructions

What does Dresdner Bank mean to you? The following descriptive scales, bounded at each end by bipolar adjectives, summarises characteristics of the bank. Please mark ✗ the blank that best indicates what the Dresdner Bank means to you.

Form

Dresdner Bank is:

Powerful :__:__:__:__:__:✗:__: Weak
Unreliable :__:__:__:__:__:✗:__: Reliable
Modern   :__:__:__:__:__:__:✗: Old-fashioned
Hi-tech   :__:__:__:__:__:✗:__: Low-tech
Careful   :__:✗:__:__:__:__:__: Careless

Figure 12.8
Semantic differential scale

weak, reliable, very old fashioned, low-tech and careful. The negative adjective or phrase sometimes appears at the left side of the scale and sometimes at the right. This controls the tendency of some respondents, particularly those with very positive or very negative attitudes, to mark the right- or left-hand sides without reading the labels.

Individual items on a semantic differential scale may be scored either on a –3 to +3 or on a 1 to 7 scale. The resulting data are commonly analysed through profile analysis. In profile analysis, means or median values on each rating scale are calculated and compared by plotting or statistical analysis. This helps determine the overall differences and similarities among the objects. To assess differences across segments of respondents, the researcher can compare mean responses of different segments. Although the mean is most often used as a summary statistic, there is some controversy as to whether the data obtained should be treated as an interval scale.[27] On the other hand, in cases when the researcher requires an overall comparison of objects, such as to determine bank preference, the individual item scores are summed to arrive at a total score.

Its versatility makes the semantic differential a popular rating scale in marketing research. It has been widely used in comparing brand, product and company images. It has also been used to develop advertising and promotion strategies and in new product development studies.[28] Several modifications of the basic scale have been proposed.[29]

Stapel scale

Stapel scale
A scale for measuring attitudes that consists of a single adjective in the middle of an even-numbered range of values.

The **Stapel scale**, named after its developer, Jan Stapel, is a unipolar rating scale with 10 categories numbered from –5 to +5, without a neutral point (zero).[30] This scale is usually presented vertically. Respondents are asked to indicate by selecting an appropriate numerical response category how accurately or inaccurately each term describes the object. The higher the number, the more accurately the term describes the object, as shown in Figure 12.9. In this example, Dresdner is perceived as not having high-quality products and having somewhat poor service.

The data obtained by using a Stapel scale can be analysed in the same way as semantic differential data. The Stapel scale produces results similar to the semantic differential.[31] The Stapel scale's advantages are that it does not require a pre-test of the adjectives or phrases to ensure true bipolarity and that it can be administered over the telephone. Some researchers, however, believe the Stapel scale is confusing and difficult to apply. Of the three itemised rating scales considered, the Stapel scale is used least.[32] Nonetheless, this scale merits more attention than it has received.

Instructions

Please evaluate how accurately each word or phrase describes Dresdner Bank. Select a positive number for the phrases you think describe the bank accurately. The more accurately you think the phrase describes the bank, the larger the plus number you should choose. You should select a minus number for the phrases you think do not describe the bank accurately. The less accurately you think the phrase describes the bank, the larger the negative number you should choose. You can select any number from +5 for phrases you think are very accurate, to –5 for phrases you think are very inaccurate.

Form

Dresdner Bank

+5	+5
+4	+4
+3	+3
+2	+2✗
+1	+1

High quality   *Poor service*

–1	–1
–2	–2
–3	–3
–4 ✗	–4
–5	–5

Figure 12.9
The Stapel scale

Itemised rating scale decisions

As is evident from the discussion so far, non-comparative itemised rating scales can take many different forms. The researcher must make six major decisions when constructing any of these scales:

1 The number of scale categories to use
2 Balanced versus unbalanced scale
3 Odd or even number of categories
4 Forced versus non-forced choice
5 The nature and degree of the verbal description
6 The physical form of the scale.

Number of scale categories

Two conflicting considerations are involved in deciding the number of scale categories or response options. The greater the number of scale categories, the finer the discrimination among stimulus objects that is possible. On the other hand, most respondents cannot handle more than a few categories. Traditional guidelines suggest that the appropriate number of categories should be between five and nine.[33] Yet there is no single optimal number of categories. Several factors should be taken into account in deciding on the number of categories.

If the respondents are interested in the scaling task and are knowledgeable about the objects, many categories may be employed. On the other hand, if the respondents are not very knowledgeable or involved with the task, fewer categories should be used. Likewise, the nature of the objects is also relevant. Some objects do not lend themselves to fine discrimination, so a small number of categories is sufficient. Another important factor is the mode of data collection. If telephone interviews are involved, many categories may confuse the respondents. Likewise, space limitations may restrict the number of categories in mail questionnaires.

How the data are to be analysed and used should also influence the number of categories. In situations where several scale items are added together to produce a single score for each respondent, five categories are sufficient. The same is true if the researcher wishes to make broad generalisations or group comparisons. If, however, individual responses are of interest or if the data will be analysed by sophisticated statistical techniques, seven or more categories may be required. The size of the correlation coefficient, a common measure of relationship between variables (Chapter 20), is influenced by the number of scale categories. The correlation coefficient decreases with a reduction in the number of categories. This, in turn, has an impact on all statistical analysis based on the correlation coefficient.[34]

Balanced versus unbalanced scale

Balanced scale
A scale with an equal number of favourable and unfavourable categories.

In a **balanced scale**, the number of favourable and unfavourable categories is equal; in an unbalanced scale, the categories are unequal.[35] Examples of balanced and unbalanced scales are given in Figure 12.10.

In general, in order to obtain objective data, the scale should be balanced. If the distribution of responses is likely to be skewed, however, either positively or negatively, an unbalanced scale with more categories in the direction of skewness may be appropriate. If an unbalanced scale is used, the nature and degree of imbalance in the scale should be taken into account in data analysis.

Odd or even number of categories

With an odd number of categories, the middle scale position is generally designated as neutral or impartial. The presence, position and labelling of a neutral category can have a significant influence on the response. The Likert scale is a balanced rating scale with an odd number of categories and a neutral point. [36]

The decision to use an odd or even number of categories depends on whether some of the respondents may be neutral on the response being measured. If a neutral or indifferent response is possible from at least some of the respondents, an odd number of categories should be used. If, on the other hand, the researcher wants to force a response or believes that no neutral or indifferent response exists, a rating scale with an even number of categories should be used. A related issue is whether the choice should be forced or non-forced.

Forced versus non-forced choice

Forced rating scale
A rating scale that forces respondents to express an opinion because a 'no opinion' or 'no knowledge' option is not provided.

On **forced rating scales** the respondents are forced to express an opinion because a 'no opinion' option is not provided. In such a case, respondents without an opinion may mark the middle scale position. If a sufficient proportion of the respondents do not have opinions on the topic, marking the middle position will distort measures of central tendency and variance. In situations where the respondents are expected to have no opinion, as opposed to simply being reluctant to disclose it, the accuracy of data may be improved by a non-forced scale that includes a 'no opinion' category.[37]
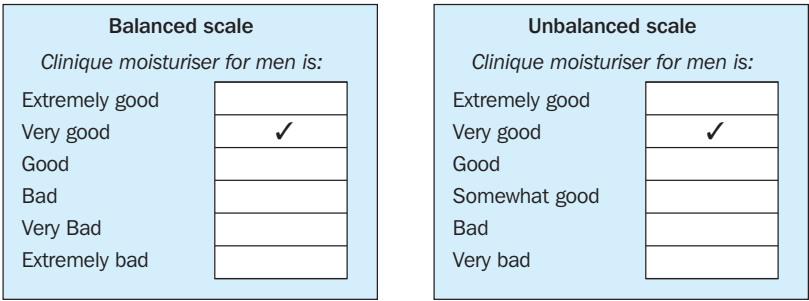
Figure 12.10
Balanced and unbalanced scales

Balanced scale		Unbalanced scale	
Clinique moisturiser for men is:		*Clinique moisturiser for men is:*	
Extremely good		Extremely good	
Very good	✓	Very good	✓
Good		Good	
Bad		Somewhat good	
Very Bad		Bad	
Extremely bad		Very bad	

Nature and degree of verbal description

The nature and degree of verbal description associated with scale categories varies considerably and can have an effect on the responses. Scale categories may have verbal, numerical or even pictorial descriptions. Furthermore, the researcher must decide whether to label every scale category, label only some scale categories, or label only extreme scale categories. Surprisingly, providing a verbal description for each category may not improve the accuracy or reliability of the data. Yet, an argument can be made for labelling all or many scale categories to reduce scale ambiguity. The category descriptions should be located as close to the response categories as possible.

The strength of the adjectives used to anchor the scale may influence the distribution of the responses. With strong anchors (1 = completely disagree, 7 = completely agree), respondents are less likely to use the extreme scale categories. This results in less variable and more peaked response distributions. Weak anchors (1 = generally disagree, 7 = generally agree), in contrast, produce uniform or flat distributions. Procedures have been developed to assign values to category descriptors to result in balanced or equal interval scales.[38]

Physical form of the scale

A number of options are available with respect to scale form or configuration. Scales can be presented vertically or horizontally. Categories can be expressed by boxes, discrete lines or units on a continuum and may or may not have numbers assigned to them. If numerical values are used, they may be positive, negative or both. Several possible configurations are presented in Figure 12.11.

Two unique rating scale configurations used in marketing research are the thermometer scale and the smiling face scale. For the thermometer scale, the higher the temperature the more favourable the evaluation. Likewise, happier faces indicate eval-
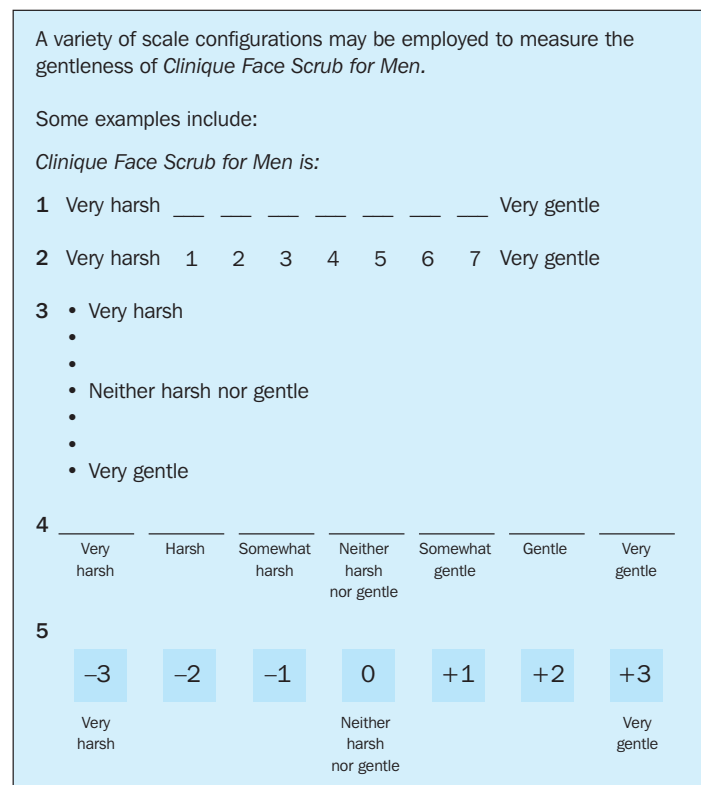
Figure 12.11
Rating scale configurations

uations that are more favourable. These scales are especially useful for children.[39] Examples of these scales are shown in Figure 12.12. Table 12.5 summarises the six decisions in designing rating scales.

Thermometer scale

Instructions

Please indicate how much you like McDonald's 'Big Macs' by colouring in the thermometer with your blue pen. Start at the bottom and colour up to the temperature that shows how much you prefer McDonalds 'Big Macs'.

Form

Like very much

100°
75°
50°
25°
0°

Dislike very much

Smiling face scale

Instructions

Please tell me how much you like Barbie Doll by pointing to the face that best shows how much you like it. If you did not like the Barbie Doll at all, you would point to Face 1. If you liked it very much, you would point to Face 5. Now tell me, how much did you like the Barbie Doll?

Form

1   2   3   4   5

Figure 12.12
Some unique rating scale configurations

Table 12.5 Summary of itemised rating scale decisions

1. Number of categories	Although there is no single, optimal number, traditional guidelines suggest that there should be between five and nine categories.
2. Balanced versus unbalanced	In general, the scale should be balanced to obtain objective data.
3. Odd or even number of categories	If a neutral or indifferent scale response is possible from at least some of the respondents, an odd number of categories should be used.
4. Forced versus unforced	In situations where the respondents are expected to have no opinion, the accuracy of the data may be improved by a non-forced scale.
5. Verbal description	An argument can be made for labelling all or many scale categories. The category descriptions should be located as close to the response categories as possible.
6. Physical form	A number of options should be tried and the best one selected.

The development and evaluation of scales

The development of multi-item rating scales requires considerable technical expertise.[40] Figure 12.13 presents a sequence of operations needed to construct multi-item scales.

The characteristic to be measured is frequently called a construct. Scale development begins with an underlying theory of the construct being measured. Theory is necessary not only for constructing the scale but also for interpreting the resulting scores. The next step is to generate an initial pool of scale items. Typically, this is based on theory, analysis of secondary data and qualitative research. From this pool, a reduced set of potential scale items is generated by the judgement of the researcher and other knowledgeable individuals. Some qualitative criterion is adopted to aid their judgement. The reduced set of items may still be too large to constitute a scale. Thus, further reduction is achieved in a quantitative manner.

Data are collected on the reduced set of potential scale items from a large pre-test sample of respondents. The data are analysed using techniques such as correlations, factor analysis, cluster analysis, discriminant analysis and statistical tests discussed later in this book. As a result of these statistical analyses, several more items are eliminated, resulting in a purified scale. The purified scale is evaluated for reliability and validity by collecting more data from a different sample (these concepts will be explained on page 313). On the basis of these assessments, a final set of scale items is selected. As can be seen from Figure 12.13, the scale development process is an iterative one with several feedback loops.[41]

A multi-item scale should be evaluated for accuracy and applicability.[42] As shown in Figure 12.14, this involves an assessment of reliability, validity and generalisability of the scale. Approaches to assessing reliability include test–re-test reliability, alternative-forms
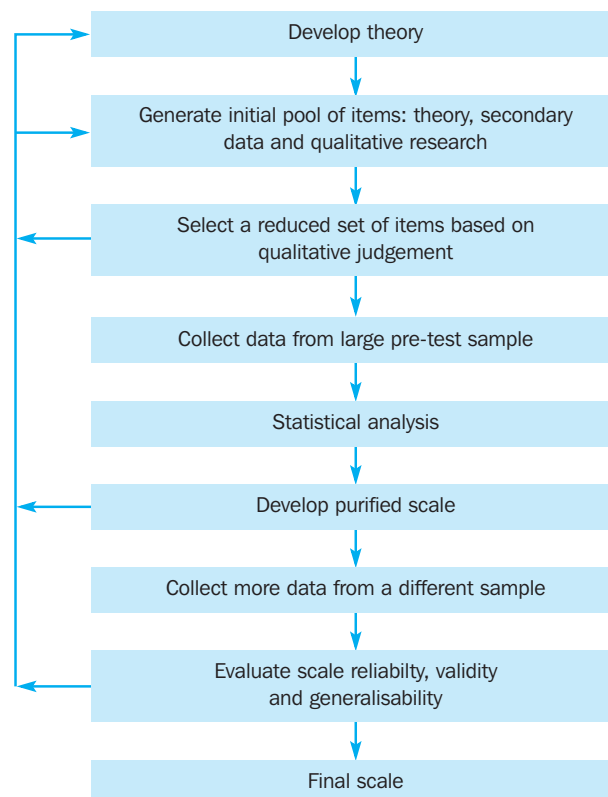
Figure 12.13
Development of a multi-item scale
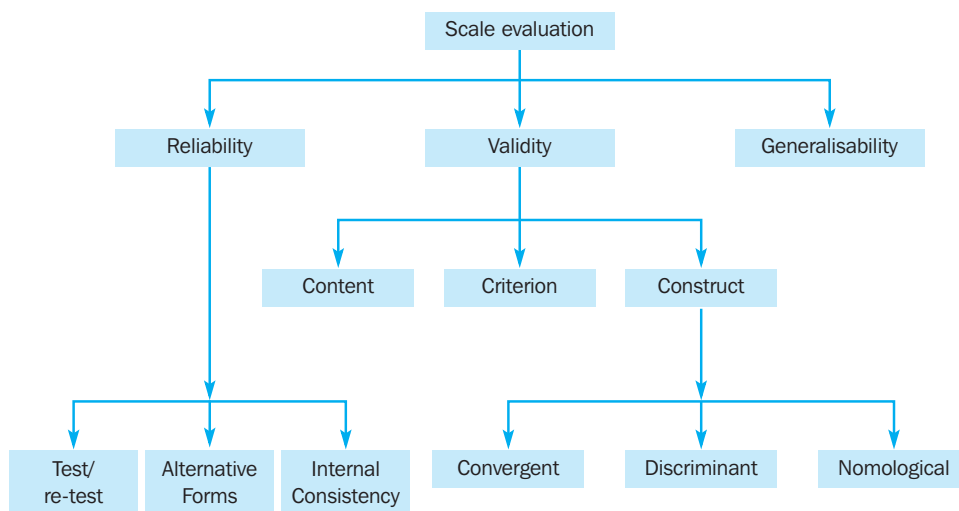
Figure 12.14
Scale evaluation

reliability and internal consistency reliability. Validity can be assessed by examining content validity, criterion validity and construct validity.

Before we can examine reliability and validity we need an understanding of measurement accuracy; it is fundamental to scale evaluation.

Measurement accuracy

Measurement error
The variation in the information sought by the researcher and the information generated by the measurement process employed.

A measurement is a number that reflects some characteristic of an object. A measurement is not the true value of the characteristic of interest but rather an observation of it. A variety of factors can cause **measurement error**, which results in the measurement or observed score being different from the true score of the characteristic being measured (see Table 12.6).

Table 12.6 Potential sources of error in measurement

1	Other relatively stable characteristics of the individual that influence the test score, such as intelligence, social desirability and education
2	Short-term or transient personal factors, such as health, emotions, fatigue
3	Situational factors, such as the presence of other people, noise and distractions
4	Sampling of items included in the scale: addition, deletion or changes in the scale items
5	Lack of clarity of the scale, including the instructions or the items themselves
6	Mechanical factors, such as poor printing, overcrowding items in the questionnaire, and poor design
7	Administration of the scale, such as differences among interviewers
8	Analysis factors, such as differences in scoring and statistical analysis

True score model
A mathematical model that provides a framework for understanding the accuracy of measurement.

The **true score model** provides a framework for understanding the accuracy of measurement.[43] According to this model,

$$X_O = X_T + X_S + X_R$$

where X_O = the observed score or measurement
X_T = the true score of the characteristic
X_S = systematic error
X_R = random error

Note that the total measurement error includes the systematic error, X_S, and the random error, X_R. **Systematic error** affects the measurement in a constant way. It represents stable factors that affect the observed score in the same way each time the measurement is made, such as mechanical factors (see Table 12.6). **Random error**, on the other hand, is not constant. It represents transient factors that affect the observed score in different ways each time the measurement is made, such as short-term transient personal factors or situational factors (see Table 12.6). The distinction between systematic and random error is crucial to our understanding of reliability and validity.

Reliability

Reliability refers to the extent to which a scale produces consistent results if repeated measurements are made.[44] Systematic sources of error do not have an adverse impact on reliability, because they affect the measurement in a constant way and do not lead to inconsistency. In contrast, random error produces inconsistency, leading to lower reliability. Reliability can be defined as the extent to which measures are free from random error, X_R. If $X_R = 0$, the measure is perfectly reliable.

Reliability is assessed by determining the proportion of systematic variation in a scale. This is done by determining the association between scores obtained from different administrations of the scale. If the association is high, the scale yields consistent results and is therefore reliable. Approaches for assessing reliability include the test–re-test, alternative forms, and internal consistency methods.

In **test–re-test reliability**, respondents are administered identical sets of scale items at two different times, under as nearly equivalent conditions as possible. The time interval between tests or administrations is typically two to four weeks. The degree of similarity between the two measurements is determined by computing a correlation coefficient (see Chapter 20). The higher the correlation coefficient, the greater the reliability.

Several problems are associated with the test–re-test approach to determining reliability. First, it is sensitive to the time interval between testing. Other things being equal, the longer the time interval, the lower the reliability. Second, the initial measurement may alter the characteristic being measured. For example, measuring respondents' attitude towards low-alcohol beer may cause them to become more health conscious and to develop a more positive attitude towards low-alcohol beer. Third, it may be impossible to make repeated measurements (for example, the research topic may be the respondent's initial reaction to a new product). Fourth, the first measurement may have a carryover effect to the second or subsequent measurements. Respondents may attempt to remember answers they gave the first time. Fifth, the characteristic being measured may change between measurements. For example, favourable information about an object between measurements may make a respondent's attitude more positive. Finally, the test–re-test reliability coefficient can be inflated by the correlation of each item with itself. These correlations tend to be higher than correlations between different scale items across administrations. Hence, it is possible to have high test–re-test correlations because of the high correlations between the same scale items measured at different times even though the correlations between different scale items are quite low. Because of these problems, a test–re-test approach is best applied in conjunction with other approaches, such as alternative-forms reliability.[45]

In **alternative-forms reliability**, two equivalent forms of the scale are constructed. The same respondents are measured at two different times, usually two to four weeks apart (e.g. by initially using Likert scaled items and then using Stapel scaled items). The scores from the administrations of the alternative scale forms are correlated to assess reliability.[46]

The two forms should be equivalent with respect to content, i.e. each scale item should attempt to measure the same items. The main problems with this approach are that it is difficult, time-consuming and expensive to construct an equivalent form of the scale. In a strict sense, it is required that the alternative sets of scale items should have the same means, variances and intercorrelations. Even if these conditions are satisfied, the two forms may not be equivalent in content. Thus, a low correlation may reflect either an unreliable scale or non-equivalent forms.

Internal consistency reliability is used to assess the reliability of a summated scale where several items are summed to form a total score. In a scale of this type, each item measures some aspect of the construct measured by the entire scale, and the items should be consistent in what they indicate about the construct. This measure of reliability focuses on the internal consistency of the set of items forming the scale.

The simplest measure of internal consistency is **split-half reliability**. The items on the scale are divided into two halves and the resulting half scores are correlated. High correlations between the halves indicate high internal consistency. The scale items can be split into halves based on odd- and even-numbered items or randomly. The problem is that the results will depend on how the scale items are split. A popular approach to overcoming this problem is to use the coefficient alpha.

The **coefficient alpha**, or Cronbach's alpha, is the average of all possible split-half coefficients resulting from different ways of splitting the scale items.[47] This coefficient varies from 0 to 1, and a value of 0.6 or less generally indicates unsatisfactory internal consistency reliability. An important property of coefficient alpha is that its value tends to increase with an increase in the number of scale items. Therefore, coefficient alpha may be artificially, and inappropriately, inflated by including several redundant scale items.[48] Another coefficient that can be employed in conjunction with coefficient alpha is coefficient beta. Coefficient beta assists in determining whether the averaging process used in calculating coefficient alpha is masking any inconsistent items.

Some multi-item scales include several sets of items designed to measure different aspects of a multidimensional construct. For example, bank image is a multidimensional construct that includes country of origin, range of products, quality of products, service of bank personnel, credit terms, investment rates, convenience of location, and physical layout of branches. Hence, a scale designed to measure bank image would contain items measuring each of these dimensions. Because these dimensions are somewhat independent, a measure of internal consistency computed across dimensions would be inappropriate. If several items are used to measure each dimension, however, internal consistency reliability can be computed for each dimension.

Validity

The **validity** of a scale may be considered as the extent to which differences in observed scale scores reflect true differences among objects on the characteristic being measured, rather than systematic or random error. Perfect validity requires that there be no measurement error ($X_O = X_T$, $X_R = 0$, $X_S = 0$). Researchers may assess content validity, criterion validity or construct validity.[49]

Content validity, sometimes called face validity, is a subjective but systematic evaluation of how well the content of a scale represents the measurement task at hand. The researcher or someone else examines whether the scale items adequately cover the entire domain of the construct being measured. Thus, a scale designed to measure bank image would be considered inadequate if it omitted any of the major dimensions (range of products, quality of products, service of bank personnel, etc.). Given its subjective nature, content validity alone is not a sufficient measure of the validity of a scale, yet it aids in a common-sense interpretation of the scale scores. A more formal evaluation can be obtained by examining criterion validity.

Internal consistency reliability
An approach for assessing the internal consistency of the set of items, where several items are summated in order to form a total score for the scale.

Split-half reliability
A form of internal consistency reliability in which the items constituting the scale are divided into two halves and the resulting half scores are correlated.

Coefficient alpha
A measure of internal consistency reliability that is the average of all possible split-half coefficients resulting from different splittings of the scale items.

Validity
The extent to which a measurement represents characteristics that exist in the phenomenon under investigation.

Content validity
A type of validity, sometimes called face validity, that consists of a subjective but systematic evaluation of the representativeness of the content of a scale for the measuring task at hand.

Criterion validity
A type of validity that examines whether the measurement scale performs as expected in relation to other selected variables as meaningful criteria.

Concurrent validity
A type of validity that is assessed when the data on the scale being evaluated and on the criterion variables are collected at the same time.

Predictive validity
A type of validity that is concerned with how well a scale can forecast a future criterion.

Construct validity
A type of validity that addresses the question of what construct or characteristic the scale is measuring. An attempt is made to answer theoretical questions of why a scale works and what deductions can be made concerning the theory underlying the scale.

Convergent validity
A measure of construct validity that measures the extent to which the scale correlates positively with other measures of the same construct.

Discriminant validity
A type of construct validity that assesses the extent to which a measure does not correlate with other constructs from which it is supposed to differ.

Nomological validity
A type of validity that assesses the relationship between theoretical constructs. It seeks to confirm significant correlations between the constructs as predicted by a theory.

Criterion validity reflects whether a scale performs as expected in relation to other selected variables (criterion variables) as meaningful criteria. If, for example, a scale is designed to measure loyalty in customers, criterion validity might be determined by comparing the results generated by this scale with results generated by observing the extent of repeat purchasing. Based on the time period involved, criterion validity can take two forms, concurrent validity and predictive validity.

Concurrent validity is assessed when the data on the scale being evaluated (e.g. loyalty scale) and the criterion variables (e.g. repeat purchasing) are collected at the same time. The scale being developed and the alternative means of encapsulating the criterion variables would be administered simultaneously and the results compared.

Predictive validity is concerned with how well a scale can forecast a future criterion. To assess predictive validity, the researcher collects data on the scale at one point in time and data on the criterion variables at a future time. For example, attitudes towards how loyal customers feel to a particular brand could be used to predict future repeat purchases of that brand. The predicted and actual purchases are compared to assess the predictive validity of the attitudinal scale.

Construct validity addresses the question of what construct or characteristic the scale is, in fact, measuring. When assessing construct validity, the researcher attempts to answer theoretical questions about why the scale works and what deductions can be made concerning the underlying theory. Thus, construct validity requires a sound theory of the nature of the construct being measured and how it relates to other constructs. Construct validity is the most sophisticated and difficult type of validity to establish. As Figure 12.14 shows, construct validity includes convergent, discriminant and nomological validity.

Convergent validity is the extent to which the scale correlates positively with other measurements of the same construct. It is not necessary that all these measurements be obtained by using conventional scaling techniques. **Discriminant validity** is the extent to which a measure does not correlate with other constructs from which it is supposed to differ. It involves demonstrating a lack of correlation among differing constructs. **Nomological validity** is the extent to which the scale correlates in theoretically predicted ways with measures of different but related constructs. A theoretical model is formulated that leads to further deductions, tests and inferences.

An example of construct validity can be evaluated in the following example. A researcher seeks to provide evidence of construct validity in a multi-item scale, designed to measure the concept of 'self-image'. These findings would be sought:[50]

- High correlations with other scales designed to measure self-concepts and with reported classifications by friends (convergent validity)
- Low correlations with unrelated constructs of brand loyalty and variety-seeking (discriminant validity)
- Brands that are congruent with the individual's self-concept are more preferred, as postulated by the theory (nomological validity)
- A high level of reliability.

Note that a high level of reliability was included as evidence of construct validity in this example. This illustrates the relationship between reliability and validity.

Relationship between reliability and validity

The relationship between reliability and validity can be understood in terms of the true score model. If a measure is perfectly valid, it is also perfectly reliable. In this case, $X_O = X_T$, $X_R = 0$, and $X_S = 0$. Thus, perfect validity implies perfect reliability. If a

measure is unreliable, it cannot be perfectly valid, since at a minimum $X_O = X_T + X_R$. Furthermore, systematic error may also be present, that is, $X_S \neq 0$. Thus, unreliability implies invalidity. If a measure is perfectly reliable, it may or may not be perfectly valid, because systematic error may still be present ($X_O = X_T + X_S$). In other words, a reliable scale can be constructed to measure 'customer loyalty' but it may not necessarily be a valid measurement of 'customer loyalty'. Conversely, a valid measurement of 'customer loyalty' has to be reliable. Reliability is a necessary, but not sufficient, condition for validity.

Generalisability

The degree to which a study based on a sample applies to the population as a whole.

Generalisability

Generalisability refers to the extent to which one can generalise from the observations at hand to a universe of generalisations. The set of all conditions of measurement over which the investigator wishes to generalise is the universe of generalisation. These conditions may include items, interviewers, and situations of observation. A researcher may wish to generalise a scale developed for use in personal interviews to other modes of data collection, such as mail and telephone interviews. Likewise, one may wish to generalise from a sample of items to the universe of items, from a sample of times of measurement to the universe of times of measurement, from a sample of observers to a universe of observers, and so on.[51]

In generalisability studies, measurement procedures are designed to investigate each universe of interest by sampling conditions of measurement from each of them. For each universe of interest, an aspect of measurement called a facet is included in the study. Traditional reliability methods can be viewed as single-facet generalisability studies. A test–re-test correlation is concerned with whether scores obtained from a measurement scale are generalisable to the universe scores across all times of possible measurement. Even if the test–re-test correlation is high, nothing can be said about the generalisability of the scale to other universes. To generalise to other universes, generalisability theory procedures must be employed.

Choosing a scaling technique

In addition to theoretical considerations and evaluation of reliability and validity, certain practical factors should be considered in selecting scaling techniques for a particular marketing research problem.[52] These include the level of information (nominal, ordinal, interval or ratio) desired, the capabilities of the respondents, the characteristics of the stimulus objects, the method of administration, the context, and cost.

As a general rule, using the scaling technique that will yield the highest level of information feasible in a given situation will permit using the greatest variety of statistical analyses. Also, regardless of the type of scale used, whenever feasible, several scale items should measure the characteristic of interest. This provides more accurate measurement than a single-item scale. In many situations, it is desirable to use more than one scaling technique or to obtain additional measures using mathematically derived scales.

Mathematically derived scales

All the scaling techniques discussed in this chapter require the respondents to directly evaluate the constructs that the researcher believes to comprise the object of study, e.g. the cognitive state of customer satisfaction. In contrast, mathematical scaling

techniques allow researchers to infer respondents' evaluations of the constructs of the object of study. These evaluations are inferred from the respondents' overall judgements. Two popular mathematically derived scaling techniques are multidimensional scaling and conjoint analysis, which are discussed in detail in Chapter 24.

International marketing research

In designing the scale or response format, respondents' educational or literacy levels should be taken into account.[53] One approach is to develop scales that are pan-cultural, or free of cultural biases. Of the scaling techniques we have considered, the semantic differential scale may be said to be pan-cultural. It has been tested in a number of countries and has consistently produced similar results. The consistency of results occurred in the following example where Xerox successfully used a Russian translation of an equivalent English semantic differential scale.

example

Copying the name Xerox[54]

Xerox was a name well received in the former Soviet Union since the late 1960s. In fact, the act of copying documents was called Xeroxing, a term coined after the name of the company. It was a brand name people equated with quality. With the disintegration of the Soviet Union into the Commonwealth of Independent States, however, Xerox's sales started to fall. The management initially considered this problem to be the intense competition with strong competitors such as Canon, Ricoh, Mitsubishi and Minolta. First attempts to make the product more competitive did not help. Subsequently, marketing research was undertaken to measure the image of Xerox and its competitors in Russia. Semantic differential scales were used, as examples of this type of scale translated well in other countries and were thus considered pan-cultural. The bipolar labels used were carefully tested to ensure that they had the intended semantic meaning in the Russian context. ■

Although the semantic differential worked well in the Russian context, an alternative approach is to develop scales that use a self-defined cultural norm as a base referent. For example, respondents may be required to indicate their own anchor point and position relative to a culture-specific stimulus set. This approach is useful for measuring attitudes that are defined relative to cultural norms (e.g. attitude towards marital roles). In developing response formats, verbal rating scales appear to be the most suitable. Even less educated respondents can readily understand and respond to verbal scales. Special attention should be devoted to determining equivalent verbal descriptors in different languages and cultures. The end points of the scale are particularly prone to different interpretations. In some cultures, 1 may be interpreted as best, whereas in others it may be interpreted as worst, regardless of how it is scaled. It is important that the scale end points and the verbal descriptors be employed in a manner consistent with the culture.

Finally, in international marketing research, it is critical to establish the equivalence of scales and measures used to obtain data from different countries. This topic is complex and is discussed in some detail in Chapter 26.

Ethics in marketing research

Ethical issues can arise in the construction of non-comparative scales. Consider, for example, the use of scale descriptors. The descriptors used to frame a scale can be manipulated to bias results in any direction. They can be manipulated to generate a positive view of the client's brand or a negative view of a competitor's brand. A researcher who wants to project the client's brand favourably can ask respondents to indicate their opinion of the brand on several attributes using seven-point scales framed by the descriptors 'extremely poor' to 'good'. Using a strongly negative descriptor with only a mildly positive one has an interesting effect. As long as the product is not the worst, respondents will be reluctant to rate the product extremely poorly. In fact, respondents who believe the product to be only mediocre will end up responding favourably. Try this yourself. How would you rate BMW cars on the following attributes?

Reliability	Horrible	1	2	3	4	5	6	7	Good
Performance	Very poor	1	2	3	4	5	6	7	Good
Quality	One of the worst	1	2	3	4	5	6	7	Good
Prestige	Very low	1	2	3	4	5	6	7	Good

Did you find yourself rating BMW cars positively? Using this same technique, a researcher can negatively bias evaluations of competitors' products by providing mildly negative descriptors against strong positive descriptors.

Thus we see how important it is to use balanced scales with comparable positive and negative descriptors. When this guide is not practised, responses are biased and should be interpreted accordingly. This concern also underscores the need to adequately establish the reliability, validity and generalisability of scales before using them in a research project. Scales that are invalid, unreliable or not generalisable to the target market provide the client with flawed results and misleading findings, thus raising serious ethical issues. The researcher has a responsibility to both the client and respondents to ensure the applicability and usefulness of the scale.

Internet and computer applications

All the primary scales of measurement that we have considered can be implemented on the Internet. The same is true for the commonly used comparative scales. Paired comparisons involving verbal, visual or auditory comparisons can be implemented with ease. However, taste, smell and touch comparisons are difficult to implement. It may also be difficult to implement specialised scales such as the Q-sort. The process of implementing comparative scales may be facilitated by searching the Internet for similar scales that have been implemented by other researchers.

Continuous rating scales may be easily implemented on the Internet. The cursor can be moved on the screen in a continuous fashion to select the exact position on the scale that best describes the respondent's evaluation. Moreover, the scale values can be automatically scored by the computer, thus increasing the speed and accuracy of processing the data.

Similarly, it is also easy to implement all of the three itemised rating scales on the Internet. Again, you can use the Internet to search for and locate cases and examples where scales have been used by other researchers. It is also possible that other researchers have reported reliability and validity assessments for multi-item scales. Before generating new scales, a researcher should first examine similar scales used by other researchers and consider using them if they meet their measurement objectives.

Summary

Measurement is the assignment of numbers or other symbols to characteristics of objects according to set rules. Scaling involves the generation of a continuum upon which measured objects are located. The four primary scales of measurement are nominal, ordinal, interval and ratio. Of these, the nominal scale is the most basic in that the numbers are used only for identifying or classifying objects. In the ordinal scale, the numbers indicate the relative position of the objects but not the magnitude of difference between them. The interval scale permits a comparison of the differences between the objects. Because it has an arbitrary zero point, however, it is not meaningful to calculate ratios of scale values on an interval scale. The highest level of measurement is represented by the ratio scale in which the zero point is fixed. The researcher can compute ratios of scale values using this scale. The ratio scale incorporates all the properties of the lower-level scales.

Scaling techniques can be classified as comparative or non-comparative. Comparative scaling involves a direct comparison of stimulus objects. Comparative scales include paired comparisons, rank order, constant sum and the Q-sort. The data obtained by these procedures have only ordinal properties. Verbal protocols, where the respondent is instructed to think out loud, can be used for measuring cognitive responses.

In non-comparative scaling, each object is scaled independently of the other objects in the stimulus set. The resulting data are generally assumed to be interval or ratio scaled. Non-comparative rating scales can be either continuous or itemised. The itemised rating scales are further classified as Likert, semantic differential, or Stapel scales. When using non-comparative itemised rating scales, the researcher must decide on the number of scale categories, balanced versus unbalanced scales, an odd or even number of categories, forced versus non-forced choices, the nature and degree of verbal description, and the physical form or configuration.

Multi-item scales consist of a number of rating scale items. These scales should be evaluated in terms of reliability and validity. Reliability refers to the extent to which a scale produces consistent results if repeated measurements are made. Approaches to assessing reliability include test–re-test, alternative forms and internal consistency. The validity of a measurement may be assessed by evaluating content validity, criterion validity and construct validity.

The choice of particular scaling techniques in a given situation should be based on theoretical and practical considerations. Generally, the scaling technique used should be the one that will yield the highest level of information feasible. Also, multiple measures should be obtained.

In international marketing research, special attention should be devoted to determining equivalent verbal descriptors in different languages and cultures. The misuse of scale descriptors also raises serious ethical concerns. The researcher has a responsibility to both the client and respondents to ensure the applicability and usefulness of scales.

Questions

1 What is measurement?

2 Highlight any marketing phenomena that you feel may be problematic in terms of assigning numbers to characteristics of those phenomena.

3 Describe and illustrate, with examples, the differences between a nominal and an ordinal scale.

4 What are the advantages of a ratio scale over an interval scale? Are these advantages significant?

5 What is a comparative rating scale ?

6 What is a paired comparison? What are the advantages and disadvantages of paired comparison scaling?

7 Describe the constant sum scale. How is it different from the other comparative rating scales?

8 Identify the type of scale (nominal, ordinal, interval or ratio) used in each of the following. Give reasons for your choice.

  (a) I like to listen to the radio when I am revising for exams

      Disagree   *Agree*
       1   2   3   4   5

  (b) How old are you? _____

  (c) Rank the following activities in terms of your preference by assigning a rank from 1 to 5 (1 = most preferred, 2 = second most preferred, etc.).
    (i)   Reading magazines
    (ii) Watching television
    (iii) Going to the cinema
    (iv) Shopping for clothes
    (v) Eating out

  (d) What is your university/college registration number? _____

  (e) In an average weekday, how much time do you spend doing class assignments?
    (i)   Less than 15 minutes
    (ii) 15 to 30 minutes
    (iii) 31 to 60 minutes
    (iv) 61 to 120 minutes
    (v) More than 120 minutes

  (f) How much money did you spend last week in the Student Union Bar? _____

9 Describe the semantic differential scale and the Likert scale. For what purposes are these scales used?

10 What are the major decisions involved in constructing an itemised rating scale? How many scale categories should be used in an itemised rating scale? Why?

11 Should an odd or even number of categories be used in an itemised rating scale?

12 How does the nature and degree of verbal description affect the response to itemised rating scales?

13 What is reliability? What are the differences between test–re-test and alternative-forms reliability?

14 What is validity? What is criterion validity? How is it assessed?

15 How would you select a particular scaling technique?

Notes

1 Gofton, K., 'If it moves measure it', *Marketing* (Marketing Technique Supplement) (4 September 1997), 17; Doucette, W.R. and Wiederholt, J.B., 'Measuring product meaning for prescribed medication using a means-end model', *Journal of Health Care Marketing* 12 (March 1992), 48–54; Nunnally, J.C., *Psychometric Theory*, 2nd edn (New York: McGraw-Hill, 1978), 3.

2 Srinivasan, V. and Park, C.S., 'Surprising robustness of the self-explicated approach to customer preference structure measurement', *Journal of Marketing Research* 34 (May 1997), 286–91; Stevens, S., 'Mathematics, measurement and psychophysics', in Stevens, S. (ed.), *Handbook of Experimental Psychology* (New York: Wiley, 1951).

3 Cook, W.D., Kress, M. and Seiford, L.M., 'On the use of ordinal data in data envelopment analysis', *Journal of the Operational Research Society* 44(2) (February 1993), 133–40; Barnard, N.R. and Ehrenberg, A.S.C., 'Robust measures of consumer brand beliefs', *Journal of Marketing Research* 27 (November 1990), 477–84; Perreault Jr, W.D. and Young, F.W., 'Alternating least squares optimal scaling: analysis of non-metric data in marketing research', *Journal of Marketing Research* 17 (February 1980), 1–13.

4 Lynn, M. and Harriss, J., 'The desire for unique consumer products: a new individual difference scale', *Psychology and Marketing* 14(6) (September 1997), 601–16.

5 For a discussion of these scales, refer to Genest, C. and Zhang, S.S., 'A graphical analysis of ratio-scaled paired comparison data', *Management Science* 42(3) (March 1996), 335–49; Coombs, C.H., 'Theory and methods of social measurement', in Festinger, L. and Katz, D. (eds) *Research Methods in the Behavioral Sciences* (New York: Holt, Rinehart & Winston, 1953).

6 Bastell, R.R. and Wind, Y., 'Product development: current methods and needed developments', *Journal of the Market Research Society* 8 (1980), 122–6.

7 There is, however, some controversy regarding this issue. See Kang, M. and Stam, A., 'PAHAP: a pairwise aggregated hierarchical analysis of ratio-scale preferences', *Decision Sciences* 25(4) (July/August 1994), 607–24.

8 Kellogg, D.L. and Chase, R.B., 'Constructing an empirically derived measure for customer contact', *Management Science* 41(11) (November 1995), 1734–49; Corfman, K.P., 'Comparability and comparison levels used in choices among consumer products', *Journal of Marketing Research* 28 (August 1991), 368–74.

9 Rickard, L., 'Remembering New Coke', *Advertising Age* 66(16) (17 April 1995), 6; 'Coke's flip-flop underscores risks of consumer taste tests', *Wall Street Journal* (18 July 1985), 25.

10 It is not necessary to evaluate all possible pairs of objects, however. Procedures such as cyclic designs can significantly reduce the number of pairs evaluated. A treatment of such procedures may be found in Malhotra, N.K., Jain, A.K. and Pinson, C., 'The robustness of MDS configurations in the case of incomplete data', *Journal of Marketing Research* 25 (February 1988), 95–102.

11 For an advanced application involving paired comparison data, see Genest, C. and Zhang, S.S., 'A graphical analysis of ratio-scaled paired comparison data', *Management Science* 42(3) (March 1996), 335–49.

12 Likert, R., Roslow, S. and Murphy, G., 'A simple and reliable method of scoring the Thurstone Attitude Scales', *Personnel Psychology* 46(3) (Autumn 1993), 689–90; Thurstone, L.L., *The Measurement of Values* (Chicago, IL: University of Chicago Press, 1959). For an application of the case V procedure, see Malhotra, N.K., 'Marketing linen services to hospitals: a conceptual framework and an empirical investigation using Thurstone's case V analysis', *Journal of Health Care Marketing* 6 (March 1986), 43–50.

13 Daniles, E. and Lawford, J., 'The effect of order in the presentation of samples in paired comparison tests', *Journal of the Market Research Society* 16 (April 1974), 127–33.

14 Herman, M.W. and Koczkodaj, W.W., 'A Monte Carlo study of pairwise comparison', *Information Processing Letters* 57(1) (15 January 1996), 25–9.

15 Kerlinger, F., *Foundations of Behavioral Research*, 3rd edn (New York: Holt, Rinehart & Winston, 1973), 583–92.

16 Noel, N.M. and Hanna, N., 'Benchmarking consumer perceptions of product quality with price: an exploration', *Psychology and Marketing* 13(6) (September 1996), 591–604; Jan-Benedict, E., Steenkamp, M. and Wittink, D.R., 'The metric quality of full-profile judgements and the number of attribute levels effect in conjoint analysis', *International Journal of Research in Marketing* 11(3) (June 1994), 275–86.

17 Hayes, J.R., 'Issues in protocol analysis', in Ungson, G.R. and Braunste, D.N. (eds), *Decision Making: An Interdisciplinary Inquiry* (Boston, MA: Kent Publishing, 1982), 61–77.

18 For an application of verbal protocols, see Harrison, D.A., McLaughlin, M.E. and Coalter, T.M., 'Context, cognition and common method variance: psychometric properties and verbal protocol evidence', *Organizational Behavior and Human Decision Processes* 68(3) (December 1996), 246–61; Gardial, S.F., Clemons, D.S., Woodruff, R.B., Schumann, D.W. and Bums, M.J., 'Comparing consumers' recall of prepurchase and postpurchase product evaluation experiences', *Journal of Consumer Research* 20 (March 1994), 548–60.

19 Mick, D.G., 'Levels of subjective comprehension in advertising processing and their relations to ad perceptions, attitudes, and memory', *Journal of Consumer Research* 18 (March 1992), 411–24; Wright, P.L., 'Cognitive processes mediating acceptance of advertising', *Journal of Marketing Research* 10 (February 1973), 53–62; Wright, P.L., 'Cognitive responses to mass media advocacy and cognitive choice processes', in Petty, R., Ostrum, T. and Brock, T. (eds), *Cognitive Responses to Persuasion* (New York: McGraw-Hill, 1978).

20 Murphy, I.P., 'RAMS helps Best Western tout worldwide positioning', *Marketing News* 31(1) (6 January 1996), 25.

21 For arguments in favour of graphic rating scales, see Lampert, S.I., 'The Attitude Pollimeter: a new attitude scaling device', *Journal of Marketing Research* (November 1979), 578–82; Tillinghast, D.S., 'Direct magnitude scales in public opinion surveys', *Public Opinion Quarterly* (Fall 1980), 377–84.

22 Albaum, G., 'The Likert scale revisited – an alternative version', *Journal of the Market Research Society* 39(2) (April 1997), 331–48; Brody, C.J. and Dietz, J., 'On the dimensionality of 2-question format Likert attitude scales', *Social Science Research* 26(2) (June 1997), 197–204; Likert, R., 'A technique for the measurement of attitudes', *Archives of Psychology* 140 (1932).

23 However, when the scale is multidimensional, each dimension should be summed separately. See Aaker, J.L., 'Dimensions of brand personality', *Journal of Marketing Research* 34 (August 1997), 347–56.

24 Herche, J. and Engelland, B., 'Reversed-polarity items and scale unidimensionality', *Journal of the Academy of Marketing Science* 24(4) (Fall 1996), 366–74.

321

25 Chandler, T.A. and Spies, C.J., 'Semantic differential comparisons of attributions and dimensions among respondents from 7 nations', *Psychological Reports* 79 (3 pt 1) (December 1996), 747–58.

26 Millar, R. and Brotherton, C., 'Measuring the effects of career interviews on young people – a preliminary study', *Psychological Reports* 79 (3 pt 2) (December 1996), 1207–15.

27 There is little difference in the results based on whether the data are ordinal or interval; however, see Garciapena, M.D., *et al.*, 'Development and validation of an inventory for measuring job satisfaction among family physicians', *Psychological Reports* 79(1) (August 1996), 291–301; Gaiton, J., 'Measurement scales and statistics: resurgence of an old misconception', *Psychological Bulletin* 87 (1980), 564–7.

28 Reisenwitz, T.H. and Wimbush Jr, G.J., 'Over-the-counter pharmaceuticals: exploratory research of consumer preferences toward solid oral dosage forms', *Health Marketing Quarterly* 13(4) (1996), 47–61; Malhotra, S., Van Auken, S. and Lonial, S.C., 'Adjective profiles in television copy testing', *Journal of Advertising Research* (August 1981), 21–5.

29 Evans, R.H., 'The upgraded semantic differential: a further test', *Journal of the Market Research Society* 22 (1980), 143–7; Swan, S.E. and Futrell, C.M., 'Increasing the efficiency of the retailer's image study', *Journal of the Academy of Marketing Science* (Winter 1980), 51–7.

30 Stapel, J., 'About 35 years of market research in the Netherlands', *Markonderzock Kwartaalschrift* 2 (1969), 3–7.

31 Hawkins, D.I., Albaum, G. and Best, R., 'Stapel scale or semantic differential in marketing research?', *Journal of Marketing Research* 11 (August 1974), 318–22; Menezes, D. and Elbert, N.E., 'Alternative semantic scaling formats for measuring store image: an evaluation', *Journal of Marketing Research* 16 (February 1979), 80–7.

32 Devellis, R.F., *Scale Development: Theories and Applications* (Thousand Oaks, CA: Sage, 1991); Etzel, M.J., Williams, T.G., Rogers, J.C. and Lincoln, D.J., 'The comparability of three Stapel scale forms in a marketing setting', in Bush, R.F. and Hunt, S.D. (eds), *Marketing Theory: Philosophy of Science Perspectives* (Chicago, IL: American Marketing Association, 1982), 303–6.

33 Coleman, A.M., Norris, C.E. and Peterson, C.C., 'Comparing rating scales of different lengths – equivalence of scores from 5-point and 7-point scales', *Psychological Reports* 80(2) (April 1997), 355–62; Viswanathan, M., Bergen, M. and Childers, T., 'Does a single response category in a scale completely capture a response?', *Psychology and Marketing* 13(5) (August 1996), 457–79; Cox III, E.P., 'The optimal number of response alternatives for a scale: a review', *Journal of Marketing Research* 17 (November 1980), 407–22; Reynolds, F.D. and Neter, J., 'How many categories for respondent classification', *Journal of the Market Research* Society 24 (October 1982), 345–6; and Lawrence, R.J., 'Reply', *Journal of the Market Research Society* 24 (October 1982), 346–8.

34 Alwin, D.F., 'Feeling thermometers versus 7-point scales – which are better', *Sociological Methods and Research* 25(3) (February 1997), 318–40; Givon, M.M. and Shapira, Z., 'Response to rating scales: a theoretical model and its application to the number of categories problem', *Journal of Marketing Research* (November 1984), 410–19; Stem Jr, D.E. and Noazin, S., 'The effects of number of objects and scale positions on graphic position scale reliability', in Lusch, R.E. *et al.*, *1985 AMA Educators' Proceedings* (Chicago, IL: American Marketing Association, 1985), 370–2.

35 Watson, D., 'Correcting for acquiescent response bias in the absense of a balanced scale – an application to class-consciousness', *Sociological Methods and Research* 21(1) (August 1992), 52–88; Schuman, H. and Presser, S., *Questions and Answers in Attitude Surveys* (New York: Academic Press, 1981), 179–201.

36 Spagna, G.J., 'Questionnaires: which approach do you use?', *Journal of Advertising Research* (February–March 1984), 67–70.

37 Hasnich, K.A., 'The job descriptive index revisited: questions about the question mark', *Journal of Applied Psychology* 77(3) (June 1992), 377–82; Schneider, K.C., 'Uninformed response rate in survey research', *Journal of Business Research* (April 1985), 153–62.

38 Gannon, K.M. and Ostrom, T.M., 'How meaning is given to rating scales – the effects of response language on category activation', *Journal of Experimental Social Psychology* 32(4) (July 1996), 337–60; Friedman, H.H. and Leefer, J.R., 'Label versus position in rating scales', *Journal of the Academy of Marketing Science* (Spring 1981), 88–92.

39 Alwin, D.F., 'Feeling thermometers versus 7-point scales – which are better', *Sociological Methods and Research* 25(3) (February 1997), 318–40.

40 For an example of a multi-item scale, see Aaker, J.L., 'Dimensions of brand personality', *Journal of Marketing Research* 34 (August 1997), 347–56; Kohli, A.K., Jaworski, B.J. and Kumar, A., 'MARKOR: a measure of market orientation', *Journal of Marketing Research* 31 (November 1993), 467–77. An application may be found in Siguaw, J.A., Brown, G. and Widing II, R.E., 'The influence of the market orientation of the firm on sales force behavior and attitude', *Journal of Marketing Research* 31 (February 1994), 106–16.

41 For example, see Singhapakdi, A., Vitell, S.J., Rallapalli, K.C. and Kraft, K.L., 'The perceived role of ethics and social responsibility: a scale development', *Journal of Business Ethics* 15(11) (November 1996), 1131–40.

42 Kim, K. and Frazier, G.L., 'Measurement of distributor commitment in industrial channels of distribution', *Journal of Business Research* 40(2) (October 1997), 139–54; Greenleaf, E.A., 'Improving rating scale measures by detecting and correcting bias components in some response styles', *Journal of Marketing Research* 29 (May 1992), 176–88.

43 The true score model is not the only theory of measurement. See Lord, E.M. and Novick, M.A., *Statistical Theories of Mental Test-Scores* (Reading, MA: Addison-Wesley, 1968).

44 Wilson, E.J., 'Research design effects on the reliability of rating scales in marketing – an update on Churchill and Peter', *Advances in Consumer Research* 22 (1995), 360–5; Perreault Jr, W.D. and Leigh, L.E., 'Reliability of nominal data based on qualitative judgements', *Journal of Marketing Research* 25 (May 1989), 135–48; Peter, J.P., 'Reliability: a review of psychometric basics and recent marketing practices', *Journal of Marketing Research* 16 (February 1979), 6–17.

45 Lam, S.S.K. and Woo, K.S., 'Measuring service quality: a test–re-test reliability investigation of SERVQUAL', *Journal of the Market Research Society* 39(2) (April 1997), 381–96.

46 Armstrong, D., Gosling, A., Weinman, J. and Marteau, T., 'The place of inter-rater reliability in qualitative research: an empirical study', *Sociology: The Journal of the British Sociological Association* 31(3) (August 1997), 597–606; Segal, M.N., 'Alternate form conjoint reliability', *Journal of Advertising Research* 4 (1984), 31–8.

47 Cronbach, L.J., 'Coefficient alpha and the internal structure of tests', *Psychometrika* 16 (1951), 297–334.

48 Peterson, R.A., 'A meta-analysis of Cronbach's coefficient alpha', *Journal of Consumer Research* 21 (September 1994), 381–91.

49 McTavish, D.G., 'Scale validity – a computer content analysis approach', *Social Science Computer Review* 15(4) (Winter 1997), 379–93; Singh, J. and Rhoads, G.K., 'Boundary role ambiguity in marketing-oriented positions: a multidimensional, multifaceted operationalization', *Journal of Marketing Research* 28 (August 1991), 328–38; Peter, J.P., 'Construct validity: a review of basic issues and marketing practices', *Journal of Marketing Research* 18 (May 1981), 133–45.

50 For further details on validity, see Sirgy, M.J., Grewal, D., Mangleburg, T.F., Park, J. *et al.*, 'Assessing the predictive ability of two methods of measuring self-image congruence', *Journal of the Academy of Marketing Science* 25(3) (Summer 1997), 229–41; Spiro, R.L. and Weitz, B.A., 'Adaptive selling: conceptualization, measurement, and nomological validity', J*ournal of Marketing Research* 27 (February 1990), 61–9.

51 For a discussion of the generalisability theory and its applications in marketing research, see Abe, S., Bagozzi, R.P. and Sadarangani, P., 'An investigation of construct validity and generalizability in the self concept: self consciousness in Japan and the United States', *Journal of International Consumer Marketing* 8(3,4) (1996), 97–123; Rentz, J.O., 'Generalisability theory: a comprehensive method for assessing and improving the dependability of marketing measures', *Journal of Marketing Research* 24 (February 1987), 19–28.

52 Hinkin, T.R., 'A review of scale development practices in the study of organisations', *Journal of Management* 21(5) (1995), 967–88.

53 Mullen, M.R., Milne, G.R. and Didow, N.M., 'Determining cross-cultural metric equivalence in survey research: a new statistical test', *Advances in International Marketing* 8 (1996), 145–57; Gencturk, E., Childers, T.L. and Ruekert, R.W., 'International marketing involvement – the construct, dimensionality, and measurement', *Journal of International Marketing* 3(4) (1995), 11–37.

54 Mckay, B., 'Xerox fights trademark battle', *Advertising Age International* (27 April 1992), 1–39.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Questionnaire design

Objectives

After reading this chapter, you should be able to:

1 explain the purpose of a questionnaire and its objectives of asking questions that respondents can and will answer, encouraging respondents, and minimising response error;

2 understand the array of trade-offs that have to be made in the total process of questionnaire design;

3 describe the process of designing a questionnaire, the steps involved, and guidelines that must be followed at each step;

4 discuss the considerations involved in designing questionnaires for international marketing research;

5 understand the ethical issues involved in questionnaire design.

The questionnaire must motivate the respondent to cooperate, become involved, and provide complete, honest and accurate answers.

Overview

This chapter discusses the importance of questionnaires and how the marketing researcher must put themselves 'in the shoes' of target respondents in order to design an effective questionnaire. There are no scientific principles that can guarantee an optimal questionnaire; it is more of a craft that is honed through experience. Through experience, the questionnaire designer will see that, given the information needs of the decision-makers they support and characteristics of target respondents, they must make a series of trade-offs. These trade-offs are described and illustrated. Next, we describe the objectives of a questionnaire and the steps involved in designing questionnaires. We provide several guidelines for developing sound questionnaires. The considerations involved in designing questionnaires when conducting international marketing research are discussed. The chapter concludes with a discussion of several ethical issues that arise in questionnaire design.

We begin with an example that establishes the main principles of good questionnaire design, i.e. to engage respondents and stimulate their interest. These are vital to providing complete and accurate answers. The second example illustrates how the characteristics of target respondents affect the design and implementation of a questionnaire.

---

e x a m p l e

Sugging
The use of marketing research to deliberately disguise a sales effort.

The field-good factor[1]

The lifeblood of the marketing research industry is busy people who are bombarded with database marketing, **sugging** and catalogue distributors. They find it difficult to distinguish them from genuine research interviewers. David Jenkins, Chief Executive of Kantor: 'The problem is partly self induced in terms of sending out long, boring questionnaires and not respecting the value of people's time. There is a big onus on clients to resist the temptation to squeeze every piece of information they can from one interview. The clients signing off questionnaires do not then subject themselves to the fieldwork reality. We've been good at training our interviewers to get bad questionnaires completed. Once it comes back as data, cleaned up and sanitised, the reality of what has happened in the field and therefore affects the data quality, never surfaces. The vast bulk of what we try and do in consumer research is ambush people to try and get their opinions for free. This ought to be a transaction. It may not be money but consumers need to get something back for their time. A charitable donation, information, increasingly as we do Internet research you could trade something electronic of value – access to data and the like'. ■

---

e x a m p l e

Finding the elusive young through cooperative parents[2]

ACCESS to Youth is a specialist survey that comprises around 1100 in-home interviews targeted at 7- to 19-year-olds. In surveying this age group it was noted that children become fatigued more quickly than adults, and this could affect data quality. In this project strict controls were placed on the interview length – it did not exceed 10 minutes. Wording was kept simple; complicated questions must be avoided. The research showed that the ease with which children were able to answer questions correlated with age. The older they were, the easier they found the question; and the more times they were asked the question the easier they found it. It was also found that children found it more difficult to understand the traditional four- or five-point agree/disagree scale and so a simplified three-point scale was used. Interviewers should be correctly briefed on how to conduct interviews with children. They needed to be patient and must treat the child as an equal – impatience or a condescending manner affects the child's ease and hence influences the answers they give. ■

Questionnaire definition

As discussed in Chapter 10, survey and observation are the two sets of techniques for obtaining quantitative primary data in descriptive research. Both methods require some procedure for standardising the data collection process so that the data obtained are internally consistent and can be analysed in a uniform and coherent manner. If 40 different interviewers conduct personal interviews or make observations in different parts of the country, the data they collect will not be comparable unless they follow specific guidelines and ask questions and record answers in a standard way. A standardised questionnaire or form will ensure comparability of the data, increase speed and accuracy of recording, and facilitate data processing.

Questionnaire
A structured technique for data collection consisting of a series of questions, written or verbal, that a respondent answers.

A **questionnaire**, whether it is called a schedule, interview form or measuring instrument, is a formalised set of questions for obtaining information from respondents. Typically, a questionnaire is only one element of a data collection package that might also include (1) fieldwork procedures, such as instructions for selecting, approaching and questioning respondents (see Chapter 16); (2) some reward, gift or payment offered to respondents; and (3) communication aids, such as maps, pictures, advertisements and products (as in personal interviews) and return envelopes (in mail surveys).

Any questionnaire has three specific objectives. First, it must translate the information needed into a set of specific questions that the respondents can and will answer. Developing questions that respondents can and will answer and that will yield the desired information is difficult. Two apparently similar ways of posing a question may yield different information. Hence, this objective is a challenge.

Second, a questionnaire must uplift, motivate and encourage the respondent to become involved in the interview, to cooperate, and to complete the interview. Figure 13.1 uses a basic marketing model of exchange of values between two parties to illustrate this point. Before designing any questionnaire or indeed any research technique, the researcher must evaluate 'what is the respondent going to get out of this'. In other words, the marketing researcher must have an empathy with the respondent and appreciate what they go through when approached and questioned. Such an appreciation of what respondents go through affects the design of how they are approached, the stated purpose of the research, the rewards for taking part and the whole process of capturing the data.

Not all respondents are the same in what they seek from a questionnaire or interview process. In Figure 13.1, some respondents may want to see some personal benefit, perhaps a tangible reward, while others may be happy to see the social benefits. Taking care in appreciating what the respondent expects from the questioning process nurtures responses that have been well thought through, are honest and accurate.

Third, a questionnaire should minimise response error. The potential sources of error in research designs were discussed in Chapter 3, where response error was defined as the error that arises when respondents give inaccurate answers or when their answers are mis-recorded or mis-analysed. A questionnaire can be a major source of response error. Minimising this error is an important objective of questionnaire design.

Questionnaire design process

The great weakness of questionnaire design is lack of theory. Because there are no scientific principles that guarantee an optimal or ideal questionnaire, questionnaire design is a skill acquired through experience. Similarly, the correct grammatical use of language does

What the respondent may want from the researcher:

- Tangible reward
- Confidentiality
- Interesting subject and experience
- Personal benefits from seeing the research completed
- Social benefits from seeing the research completed
- Being 'chosen' as a respondent with expertise on the subject
- Research organisation known for excellence in research
- Rapport and trust

Marketing researcher — Exchange of values — Respondents
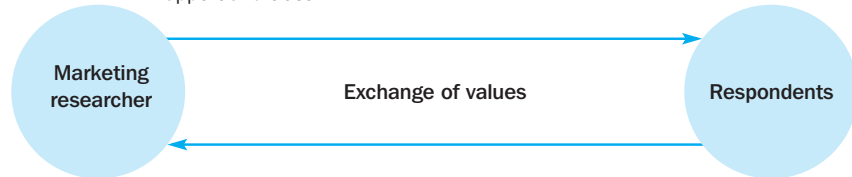
What the researcher wants from respondents:

- Honesty
- Takes in the reason for the study
- Follows the instructions in completing the study
- Thinks through the issues before forming an answer
- Says good things about the rationale for marketing research
- Says good things about the research process

Figure 13.1
Exchange of values between marketing researchers and respondents

not guarantee the optimal questionnaire. There may be certain respondents who do not communicate in a 'correct' grammatical manner; such questionnaires may be confusing and meaningless. Therefore, this section presents guidelines and rules to help develop the craft of questionnaire design. Although these guidelines and rules can help you avoid major mistakes, the fine-tuning of a questionnaire comes from the creativity of a skilled researcher.[3] Developing the craft of questionnaire design requires the creative 'trade-off' of many factors. Figure 13.2 helps to illustrate some of the trade-offs that the questionnaire designer faces, that collectively make the design process problematic.

The design process is founded upon generating information that will effectively support marketing decision-makers. Establishing the nature of marketing problems

Designing a questionnaire to match the language and logic of target respondents is a craft that requires many tools.
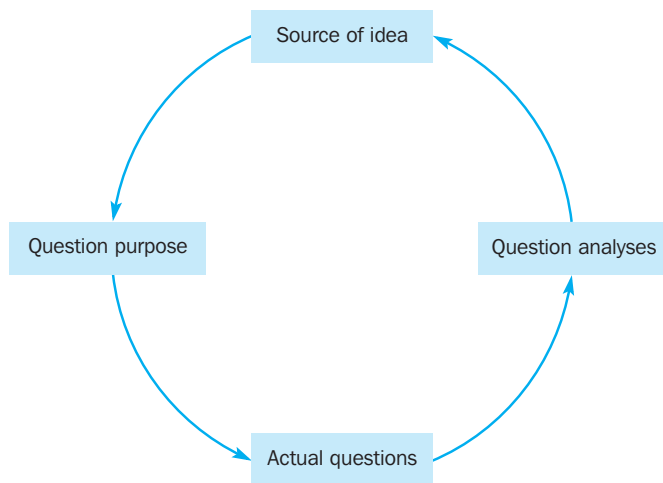
Figure 13.2
Trade-offs faced by the questionnaire designer

and corresponding marketing research problems, i.e. defining the nature of effective support, was discussed in Chapter 2. Different techniques and sources of information were outlined to help in the diagnosis process, which feed directly into the stages set out below.

1 The 'source of idea' represents the culmination of marketing decision-maker and marketing researchers' diagnoses and the data they have available at the time of commissioning a marketing research project.

2 From the diagnoses, and the statement of marketing and research problems, emerge specific research questions. Based upon the diagnoses, the purpose of each potential question should be established, i.e. 'question purposes'. Some research problems may be tackled through actual measurements in questionnaires. Other research problems may not be tackled by questionnaires. For example, in the GlobalCash Project, banks wished to know which criteria companies use to choose a bank. Tackling this issue is a straightforward task of establishing criteria and applying scales of importance to these criteria. Banks further wished to know what forces were shaping these criteria and how these forces interacted. The latter issues are qualitative in nature and were captured in follow-up in-depth interviews. In establishing question purposes, priorities have to be set of what can be measured using a questionnaire, and out of the array of issues that could be measured, what are seen as the most, down to the least, important.

3 With clear question purposes, the process of establishing 'actual questions' can begin. At this point, the researcher has to put themselves 'in the shoes' of the potential respondent. It is fine to say that certain questions need to be answered, but this has to be balanced with an appreciation of whether respondents are able or indeed willing to answer particular questions. For example in the GlobalCash Project, banks wished to know which banks individual companies use in every European country. The purpose of such a question was to establish which banks were gaining market share and to see whether companies were concentrating their business with fewer banks. This seems a straightforward question until one realises the enormity of the task. There are over 20 European countries and in some of them the norm is to conduct business with a large number of banks. In Italy, for example, one company conducted business with 70 banks. Thus, the task of remembering all the banks and writing them down placed an enormous burden on the respondent. Making the question closed by listing all the banks does not help either; in Germany, for example, there are over 400 banks that could be listed. The trade-off or final compromise was to ask respon-

dents who their 'lead bank' was, i.e. with whom they did most business. Asking them to name just one bank ensured that the task was not too onerous for the respondent, enabling a more complete and accurate response. The trade-off was that it did not completely satisfy the set question purpose.

4 Deciding how the data collected is to be analysed does not happen when questionnaires have been returned from respondents. 'Question analyses' must be thought through from an early stage. The connections between questions and the appropriate statistical tests that fulfil the question purposes should be established as the questionnaire is designed. Again, trade-offs have to be considered. In Chapter 12, different scale types were linked to different statistical tests. As one progresses from nominal to ordinal to interval and then ratio scales, more powerful statistical analyses can be performed. However, as one progresses through these scale types, the task for respondents becomes more onerous. This trade-off can be illustrated again using questions from the GlobalCash Project. Companies were asked who they thought were the top four banks in their country (ordinal scale). Respondents were then asked why they thought the top two banks were perceived to be best and second-best. This could be completed in a number of ways. A list of characteristics could be given and respondents asked to tick those that they thought matched the bank. This would be easy for respondents to undertake and produce nominal data. The same set of characteristics could be listed with respondents asked to rank order them. This task requires more thought and effort, though now produces the more powerful ordinal data. The same list could have been presented and respondents asked to allocate 100 points using a constant sum scale. This would have been an even more onerous task but would have produced the more powerful interval scale. The questionnaire designer has to consider how onerous the task is for respondents, especially when set in the context of all the other questions the respondent is being asked, and trade this off against the understanding they get from the data.

5 The understanding that is taken from the data comes back to the 'source of idea'. By now they may have collected other data, interpreted existing data differently, or been exposed to new forces in the marketplace. They may even now see what questions they should have been asking!

There can be no theory to encapsulate the trade-offs illustrated in Figure 13.2. Each research project will have different demands and emphases. With the experience of designing a number of questionnaires, the 'craft' of questionnaire design is developed and the balance understood to meet different demands and emphases.

In order to develop a further understanding of questionnaire design, the process will be presented as a series of steps, as shown in Figure 13.3, and we present guidelines for each step. The process outlined in Figure 13.2 shows that in practice the steps are interrelated and the development of a questionnaire involves much iteration and interconnection between stages.[4]

Specify the information needed

The first step in questionnaire design is to specify the information needed. This is also the first step in the research design process. Note that, as the research project progresses, the information needed becomes more and more clearly defined. It is helpful to review the components of the problem and the approach, particularly the research questions, hypotheses and characteristics that influence the research design. To further ensure that the information obtained fully addresses all the components of the

Specify the information needed

↓

Specify the type of interviewing method

↓

Determine the content of individual questions

↓

Overcome the respondent's inability and unwillingness to answer

↓

Choose question structure

↓

Choose question wording

↓

Arrange the questions in proper order

↓

Identify the form and layout

↓

Reproduce the questionnaire

↓

Eliminate problems by pilot-testing
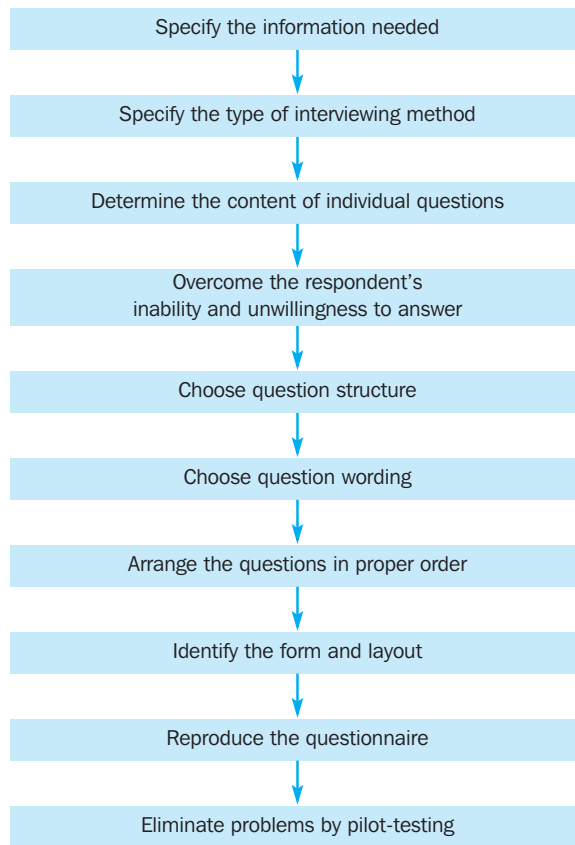
Figure 13.3
Questionnaire design process

problem, the researcher should prepare a set of dummy tables. A dummy table is a blank table used to catalogue data. It portrays how the analysis will be structured once the data have been collected.

It is also vital to have a clear idea of the target respondents. The characteristics of the respondent group have a great influence on questionnaire design. The wording and style of questions that may be appropriate for finance directors being surveyed about their IT needs may not be appropriate for retired persons being surveyed about their holiday needs. The more diversified the respondent group, the more difficult it is to design a single questionnaire appropriate for the entire group.

Specify the type of interviewing method

An appreciation of how the type of interviewing method influences questionnaire design can be obtained by considering how the questionnaire is administered under each method (see Chapter 10). In personal interviews, respondents see the questionnaire and interact face to face with the interviewer. Thus, lengthy, complex and varied questions can be asked. In telephone interviews the respondents interact with the interviewer, but they do not see the questionnaire. This limits the type of questions that can be asked to short and simple ones. Mail and electronic questionnaires are self administered, so the questions must be simple, and detailed instructions must be provided. In computer-assisted interviewing (CAPI and CATI), complex skip patterns and randomisation of questions to eliminate order bias can be easily accommodated. Questionnaires designed for personal and telephone interviews should be written in a conversational style.[5]

Determine the content of individual questions

Once the information needed is specified and the type of interviewing method decided, the next step is to determine individual question content: what to include in individual questions.

Is the question necessary?

Every question in a questionnaire should contribute to the information needed or serve some specific purpose. If there is no satisfactory use for the data resulting from a question, that question should be eliminated.

In certain situations, however, questions may be asked that are not directly related to the needed information. It is useful to ask some neutral questions at the beginning of the questionnaire to establish involvement and rapport, particularly when the topic of the questionnaire is sensitive or controversial. Sometimes filter questions are asked to disguise the purpose or sponsorship of the project. For example, rather than limiting the questions to the brand of interest, questions about competing brands may be included to disguise the sponsorship. Questions unrelated to the immediate problem may sometimes be included to generate client support for the project. At times, certain questions may be duplicated for the purpose of assessing reliability or validity.[6]

Are several questions needed instead of one?

Double-barrelled question
A single question that attempts to cover two issues. Such questions can be confusing to respondents and result in ambiguous responses.

Once we have ascertained that a question is necessary, we must make sure that it is sufficient to get the desired information. Sometimes several questions are needed to obtain the required information in an unambiguous manner. Consider the question 'Do you think Coca-Cola is a tasty and refreshing soft drink?' A yes answer will presumably be clear, but what if the answer is no? Does this mean that the respondent thinks that Coca-Cola is not tasty, that it is not refreshing, or that it is neither tasty nor refreshing? Such a question is called a **double-barrelled question**, because two or more questions are combined into one. To obtain the required information, two distinct questions should be asked: 'Do you think Coca-Cola is a tasty soft drink?' and 'Do you think Coca-Cola is a refreshing soft drink?'.

Another example of multiple questions embedded in a single question is the 'why' question. In the context of the GlobalCash study, consider the question 'Why do you bank at ABN AMRO?' The possible answers may include: 'for their banking software', 'their branches are more conveniently located', and 'they were recommended by a respected colleague at a conference'. Each answer relates to a different question embedded in the why question. The first tells why the respondent banks with the pan-European bank, the second reveals what the respondent likes about ABN AMRO compared with other banks, and the third tells how the respondent learned about ABN AMRO. The three answers are not comparable and any one answer may not be sufficient. Complete information may be obtained by asking two separate questions: 'What do you like about ABN AMRO compared with other banks?' and 'How did you first develop a relationship with ABN AMRO?' Most 'why' questions about the use of a product or choice alternative involve two aspects: (1) attributes of the product and (2) influences leading to knowledge of it.[7]

Overcoming the respondent's inability and unwillingness to answer

Researchers should not assume that respondents can provide accurate or reasonable answers to all questions (assuming that they are willing to!). The researcher should attempt to overcome the respondents' inability to answer. Certain factors limit the respondents' ability to provide the desired information. The respondents may not be informed, may not remember, or may be unable to articulate certain types of responses.

Is the respondent informed?

Respondents are often asked about topics on which they are not informed. A parent may not be informed about their child's daily purchasing or vice versa (especially when one considers that sons and daughters can have an age range from 0 to around 80!).

Filter question
An initial question in a questionnaire that screens potential respondents to ensure they meet the requirements of the sample.

In situations where not all respondents are likely to be informed about the topic of interest, **filter questions** that measure familiarity, product use and past experience should be asked before questions about the topics themselves.[8] Filter questions enable the researcher to filter out respondents who are not adequately informed. The GlobalCash questionnaire included questions that asked respondents about their experiences when working with banks in domestic markets and in international markets. A filter question was used to establish whether a respondent was informed about banking in international markets. If they had no experience of international banking, they could skip the 'international banking' section and focus upon questions that were relevant to them. The use of CATI, CAPI and Internet surveys allow extensive filtering to produce a variety of questionnaire formats that can be tailored to the familiarity, product use and past experiences of respondents.

As a general rule, a 'don't know' option appears to reduce uninformed responses without reducing the overall response rate or the response rate for questions about which the respondents have information. Hence, this option should be provided unless there are explicit reasons for not doing so.[9]

Can the respondent remember?

Many things that we might expect everyone to know are remembered by only a few. Test this on yourself. Can you remember the brand name of the socks you are wearing (presuming you are wearing socks), what you had for lunch a week ago, or what you were doing a month ago at noon? Further, do you know how many litres of soft drinks you consumed during the last four weeks? Evidence indicates that consumers are particularly poor at remembering quantities of products consumed. In situations where factual data were available for comparison, it was found that consumer reports of product usage exceeded actual usage by 100% or more.[10]

Telescoping
A psychological phenomenon that takes place when an individual telescopes or compresses time by remembering an event as occurring more recently than it actually occurred.

The inability to remember leads to errors of omission, telescoping and creation. Omission is the inability to recall an event that actually took place. **Telescoping** takes place when an individual telescopes or compresses time by remembering an event as occurring more recently than it actually occurred.[11] For example, a respondent reports three trips to the supermarket in the last two weeks when, in fact, one of these trips was made 18 days ago. Creation error takes place when a respondent 'remembers' an event that did not actually occur.

The ability to remember an event is influenced by (1) the event itself, (2) the time elapsed since the event, and (3) the presence or absence of things that would aid memory. We tend to remember events that are important or unusual or that occur frequently. People remember their wedding anniversary and birthday. Likewise, more recent events are remembered better. A grocery shopper is more likely to remember what he purchased on his last shopping trip than what he bought three shopping trips ago.

Research indicates that questions that do not provide the respondent with cues to the event, and that rely on unaided recall, can underestimate the actual occurrence of an event. For example, testing whether respondents were exposed to a beer commercial at the cinema could be measured in an unaided manner by questions like 'What brands of beer do you remember being advertised last night at the cinema?' (having established that the respondent was at a cinema last night). Naming a brand shows that they saw the advert, took in the brand name and could recall it – three different stages. An aided recall approach attempts to stimulate the respondent's memory by providing cues related to the event of interest. Thus, the important features to measure may be that they saw the advert and took in the brand name – the fact that they cannot say the brand name may not affect their purchasing intentions. The aided recall approach would list a number of beer brands and then ask 'Which of these brands were advertised last night at the cinema?' In presenting cues, the researcher must guard against biasing the responses by testing out several successive levels of stimulation. The influence of stimulation on responses can then be analysed to select an appropriate level of stimulation.

Is the respondent able to articulate?

Respondents may be unable to articulate certain types of responses. For example, if asked to describe the atmosphere of the bank branch they would prefer to patronise, most respondents may be unable to phrase their answers. On the other hand, if the respondents are provided with alternative descriptions of bank atmosphere, they will be able to indicate the one they like the best. If the respondents are unable to articulate their responses to a question, they are likely to ignore that question and refuse to respond to the rest of the questionnaire. Thus, respondents should be given aids such as pictures, maps and descriptions to help them articulate their responses.

Even if respondents are able to answer a particular question, they may be unwilling to do so, either because too much effort is required, the situation or context may not seem appropriate for disclosure, no legitimate purpose or need for the information requested is apparent, or the information requested is sensitive.

Effort required of the respondents

Most respondents are unwilling to devote much effort to providing information. Suppose that the researcher is interested in determining from which shops a respondent bought goods on their most recent shopping trip. This information can be obtained in at least two ways. The researcher could ask the respondent to list all the items purchased on their most recent shopping trip, or the researcher could provide a list of shops and ask the respondent to indicate the applicable ones. The second option is preferable, because it requires less effort from respondents.

Context

Some questions may seem appropriate in certain contexts but not in others. For example, questions about personal hygiene habits may be appropriate when asked in a survey sponsored by a health organisation but not in one sponsored by a breakfast cereal manufacturer. Respondents are unwilling to respond to questions they consider to be inappropriate for the given context. Sometimes, the researcher can manipulate the context in which the questions are asked so that the questions seem appropriate.

Legitimate purpose

Respondents are also unwilling to divulge information that they do not see as serving a legitimate purpose. Why should a firm marketing breakfast cereals want to know

their age, income and occupation? Explaining why the data are needed can make the request for the information seem legitimate and may increase the respondents' willingness to answer. A statement such as 'To determine how the preferences for cereal brands vary among people of different ages, we need information on . . . ' can make the request for information seem more legitimate.

Sensitive information

Respondents may be unwilling to disclose, at least accurately, sensitive information because this may cause embarrassment or threaten the respondent's prestige or self-image, or be seen as too personal and an invasion of privacy. If pressed for the answer, respondents may give biased responses, especially during personal interviews[12] (see Table 10.2). Sensitive topics include money, personal hygiene, family life, political and religious beliefs, involvement in accidents or crimes. In industrial surveys, sensitive questions may encompass much of what a company does, especially if it reveals strategic activities and plans. The techniques described in the following section can be adopted to increase the likelihood of obtaining information that respondents are unwilling to give.

Increasing the willingness of respondents

Respondents may be encouraged to provide information which they are unwilling to give by the following techniques.[13]

1   Place sensitive topics at the end of the questionnaire. By then, initial mistrust has been overcome, rapport has been created, legitimacy of the project has been established, and respondents are more willing to give information. In this context, consider how sensitive classification questions such as gender, age and income may be perceived.
2   Preface the question with a statement that the behaviour of interest is common. For example, before requesting information on credit card debt, say 'Recent studies show that most European consumers are in debt'. This technique describes the use of counter-biasing statements.
3   Ask the question using the third-person technique (see Chapter 7): phrase the question as if it referred to other people.
4   Hide the question in a group of other questions that respondents are willing to answer. The entire list of questions can then be asked quickly.
5   Provide response categories rather than asking for specific figures.[14] Do not ask 'What is your household's annual income?' Instead, ask the respondent to indicate an appropriate income category. In personal interviews, give the respondents cards that list the numbered choices. The respondents then indicate their responses by number.
6   Use randomised techniques. In these techniques, respondents are presented with two questions, one sensitive and the other a neutral question with a known probability of yes responses (e.g. 'Is your birthday in March?'). They are asked to select one question randomly by flipping a coin, for example. The respondent then answers the selected question yes or no, without telling the researcher which question is being answered.[15] Given the overall probability of a yes response, the probability of selecting the sensitive question, and the probability of a yes response to the neutral question, the researcher can determine the probability of a yes response to the sensitive question using the law of probability. The researcher cannot, however, determine which respondents have answered yes to the sensitive question.[16]

Choose question structure

A question may be unstructured or structured. We define unstructured questions and discuss their relative advantages and disadvantages and then consider the major types of structured questions: multiple-choice, dichotomous and scales.[17]

Unstructured questions

Unstructured questions are open-ended questions that respondents answer in their own words. They are also referred to as free-response or free-answer questions. The following are some examples:

- What is your occupation?
- What do you think of people who patronise secondhand clothes shops?
- Who is your favourite film personality?

Open-ended questions are good first questions on a topic. They enable the respondents to express general attitudes and opinions that can help the researcher interpret their responses to structured questions. They can also be useful as a final question in a questionnaire. After respondents have thought through and given all their answers in a questionnaire, there may be other issues that are important to them, that may not have been covered. Having an open-ended question at the end allows respondents to express these issues. As well as providing material to help the researcher interpret other responses, the respondent has the chance to express what they feel to be important. Unstructured questions have a much less biasing influence on response than structured questions. Respondents are free to express any views. Their comments and explanations can provide the researcher with rich insights.

A principal disadvantage is that potential for interviewer bias is high. Whether the interviewers record the answers verbatim or write down only the main points, the data depend on the skills of the interviewers. Tape recorders should be used if verbatim reporting is important.

Another major disadvantage of unstructured questions is that the coding of responses is costly and time-consuming, as illustrated in Chapter 9.[18] The coding procedures required to summarise responses in a format useful for data analysis and interpretation can be extensive. Implicitly, unstructured or open-ended questions give extra weight to respondents who are more articulate. Also, unstructured questions are not very suitable for self-administered questionnaires (mail and CAPI), because respondents tend to be briefer in writing than in speaking.

Pre-coding can overcome some of the disadvantages of unstructured questions. Expected responses are recorded in multiple-choice format, although the question is presented to the respondents as an open-ended question. Based on the respondent's reply, the interviewer selects the appropriate response category. Because the response alternatives are limited, this approach may be satisfactory when the respondent can easily formulate the response and when it is easy to develop pre-coded categories. In general, open-ended questions are useful in exploratory research and as opening or closing questions. They should be chosen with great care as their disadvantages can outweigh their advantages in a large survey.[19]

Structured questions

Structured questions specify the set of response alternatives and the response format. A structured question may be multiple-choice, dichotomous or a scale.

Multiple choice questions. In multiple-choice questions, the researcher provides a choice of answers and respondents are asked to select one or more of the alternatives given. Consider the following question:

Which of the following items have you purchased in the last two months?
Please tick as many as apply.

✓

1 Clothing	
2 Jewellery	
3 Cosmetics	
4 Shoes	
•	
•	
•	
16 Camera equipment	
17 Other (please specify)	

Of concern in designing multiple-choice questions are the number of alternatives that should be included and the order of potential responses, known as position bias. The response alternatives should include the set of all possible choices. The general guideline is to list all alternatives that may be of importance and to include an alternative labelled 'other (please specify)', as shown above. The response alternatives should be mutually exclusive. Respondents should also be able to identify one, and only one, alternative, unless the researcher specifically allows two or more choices (for example, 'Please indicate all the brands of soft drinks that you have consumed in the past week'). If the response alternatives are numerous, consider using more than one question to reduce the information processing demands on the respondents.

Order bias (position bias)

A respondent's tendency to choose an alternative merely because it occupies a certain position or is listed in a certain order.

Order bias or **position bias** is the respondents' tendency to tick an alternative merely because it occupies a certain position or is listed in a certain order. Respondents tend to tick the first or the last statement in a list, particularly the first.[20] For a list of numbers (quantities or prices), there is a bias towards the central value on the list. To control for order bias, several forms of the questionnaire should be prepared with the order in which the alternatives are listed varied from form to form. Each alternative should appear once in each of the extreme positions, once in the middle, and once somewhere in between.[21]

Multiple-choice questions overcome many of the disadvantages of open-ended questions because interviewer bias is reduced and these questions are administered quickly. Also, coding and processing of data are much less costly and time-consuming. In self-administered questionnaires, respondent cooperation is improved if the majority of the questions are structured.

Multiple-choice questions are not without disadvantages. Considerable effort is required to design effective multiple-choice questions. Qualitative techniques may be required to determine the appropriate response alternatives. It is difficult to obtain information on alternatives not listed. Even if an 'other (please specify)' category is included, respondents tend to choose among the listed alternatives. In addition, showing respondents the list of possible answers produces biased responses.[22] There is also the potential for order bias.

Dichotomous question

A structured question with only two response alternatives, such as yes and no.

Dichotomous questions. A **dichotomous question** has only two response alternatives, such as yes or no, or agree or disagree. Often, the two alternatives of interest are supplemented by a neutral alternative, such as 'no opinion', 'don't know', 'both', or 'none', as in this example.[23]

Do you intend to buy a new laptop computer within the next six months?

✓

Yes	
No	
Don't know	

Note that this question could also be framed as a multiple-choice question using response alternatives 'Definitely will buy', 'Probably will buy', 'Probably will not buy', and so forth. The decision to use a dichotomous question should be guided by whether the respondents approach the issue as a yes-or-no issue. Although decisions are often characterised as series of binary or dichotomous choices, the underlying decision-making process may reflect uncertainty that can best be captured by multiple-choice responses. For example, two individuals may be equally likely to buy a new laptop computer within the next six months if the economic conditions remain favourable. One individual, who is being optimistic about the economy, will answer yes, while the other, feeling pessimistic, will answer no.

Another issue in the design of dichotomous questions is whether to include a neutral response alternative. If it is not included, respondents are forced to choose between yes and no even if they feel indifferent. On the other hand, if a neutral alternative is included, respondents can avoid taking a position on the issue, thereby biasing the results. We offer the following guidelines. If a substantial proportion of the respondents can be expected to be neutral, include a neutral alternative. If the proportion of neutral respondents is expected to be small, avoid the neutral alternative.[24]

The general advantages and disadvantages of dichotomous questions are very similar to those of multiple-choice questions. Dichotomous questions are the easiest type of questions to code and analyse, but they have one acute problem. The response can be influenced by the wording of the question. To illustrate, the statement 'Individuals are more to blame than social conditions for crime and lawlessness in this country' produced agreement from 59.6% of the respondents. On a matched sample that responded to the opposite statement, 'Social conditions are more to blame than individuals for crime and lawlessness in this country', however, 43.2% (as opposed to 40.4%) agreed.[25] To overcome this problem, the question should be framed in one way on one-half of the questionnaires and in the opposite way on the other half. This is referred to as the split ballot technique.

Scales. Scales were discussed in detail in Chapter 12. To illustrate the difference between scales and other kinds of structural questions, consider the question about intentions to buy a new laptop computer. One way of framing this using a scale is as follows:

Do you intend to buy a new laptop computer within the next six months?

Definitely will not buy	Probably will not buy	Undecided	Probably will buy	Definitely will buy
1	2	3	4	5

This is only one of several scales that could be used to ask this question (see Chapter 12).

Choose question wording

Question wording is the translation of the desired question content and structure into words that respondents can clearly and easily understand. Deciding on question wording is perhaps the most critical and difficult task in developing a questionnaire. If a question is worded poorly, respondents may refuse to answer it or answer it incorrectly. The first condition, known as item non-response, can increase the complexity of data analysis.[26] The second condition leads to response error, discussed earlier. Unless the respondents and the researcher assign exactly the same meaning to the question, the results will be seriously biased.[27]

To avoid these problems, we offer the following guidelines: (1) define the issue, (2) use ordinary words, (3) use unambiguous words, (4) avoid leading or biasing questions, (5) avoid implicit alternatives, (6) avoid implicit assumptions, (7) avoid generalisations and estimates, and (8) use positive and negative statements.

Define the issue

A question should clearly define the issue being addressed. Consider the following question:

> Which brand of shampoo do you use?

On the surface, this may seem to be a well-defined question, but we may reach a different conclusion when we examine it in terms of 'who', 'what', 'when' and 'where'. 'Who' in this question refers to the respondent. It is not clear, though, whether the researcher is referring to the brand the respondent uses personally or the brand used by the household. 'What' is the brand of shampoo. But what if more than one brand of shampoo is being used? Should the respondent mention the most preferred brand, the brand used most often, the brand used most recently, or the brand that comes to mind first? 'When' is not clear; does the researcher mean last time, last week, last month, last year, or ever? As for 'where', it is implied that the shampoo is used at home, but this is not stated clearly. A better wording for this question would be:

> Which brand or brands of shampoo have you personally used at home during the last month? In the case of more than one brand, please list all the brands that apply.

Use ordinary words

Ordinary words should be used in a questionnaire, and they should match the vocabulary level of the respondents.[28] In other words, even though we may speak the same language as our potential respondents, there may be particular colloquialisms and ways of using words and terms they use which we should acquaint ourselves with. When choosing words, bear in mind the intellectual level of the target group of respondents, and how comfortable they are with technical terms related to any products or services we are measuring. Most respondents do not understand technical marketing words. For example, instead of asking 'Do you think the distribution of soft drinks is adequate?', ask 'Do you think soft drinks are readily available when you want to buy them?' Never forget that you are imposing your language upon respondents in the form of a questionnaire. Your language communicates and puts respondents in a particular frame of mind as they answer the questions you pose. Unless that language is meaningful to respondents, they will be put in a frame of mind that you may not intend, and be answering different questions from those you set.

Use unambiguous words

The words used in a questionnaire should have a single meaning that is known to the respondents.[29] A number of words that appear to be unambiguous have different meanings for different people.[30] These include 'usually', 'normally', 'frequently', 'often', 'regularly', 'occasionally' and 'sometimes'. Consider the following question:

In a typical month, how often do you visit a bank?

✓

Never	☐
Occasionally	☐
Sometimes	☐
Often	☐
Regularly	☐

The answers to this question are fraught with response bias, because the words used to describe category labels have different meanings for different respondents. Three respondents who visit a bank once a month may tick three different categories: occasionally, sometimes and often. A much better wording for this question would be the following:

In a typical month, how often do you visit a bank?

✓

Less than once	☐
1 or 2 times	☐
3 or 4 times	☐
More than 4 times	☐

Note that this question provides a consistent frame of reference for all respondents. Response categories have been objectively defined, and respondents are no longer free to interpret them in their own way.

In deciding on the choice of words, researchers should consult a dictionary and thesaurus and ask the following questions:

1 Does the word mean what we intend?
2 Does the word mean the same to our target respondents?
3 Does the word have any other meanings?
4 If so, does the context make the intended meaning clear?
5 Does the word have more than one pronunciation or similar pronunciations that may be confusing?
6 Is a simpler word or phrase suggested that may be more meaningful to our target respondents?

Avoid leading or biasing questions

Leading question
A question that gives the respondent a clue as to what the answer should be.

A **leading question** is one that clues the respondent to what the answer should be, as in the following:

Do you think that patriotic French people should buy imported cars when that would put French workers out of employment?

✓

Yes	☐
No	☐
Don't know	☐

Leading questions can take your respondents in unintended directions – a little like a wobbly wheel on your supermarket trolley.

This question would lead respondents to a 'No' answer. After all, how could patriotic French people put French people out of work? Therefore, this question would not help determine the preferences of French people for imported versus domestic cars.

Bias may also arise when respondents are given clues about the sponsor of the project. Respondents tend to respond favourably towards the sponsor. The question 'Is Colgate your favourite toothpaste?' is likely to bias the responses in favour of Colgate. A more unbiased way of obtaining this information would be to ask 'What is your favourite toothpaste brand?' Likewise, the mention of a prestigious or non-prestigious name can bias the response, as in 'Do you agree with the British Dental Association that Colgate is effective in preventing cavities?' An unbiased question would be to ask 'Is Colgate effective in preventing cavities?'[31]

Avoid implicit alternatives

Implicit alternative
An alternative that is not explicitly expressed.

An alternative that is not explicitly expressed in the options is an **implicit alternative**. Making an implied alternative explicit may increase the percentage of people selecting that alternative, as in the following two questions.

1 Do you like to fly when travelling short distances?
2 Do you like to fly when travelling short distances, or would you rather drive?

In the first question, the alternative of driving is only implicit, but in the second question it is explicit. The first question is likely to yield a greater preference for flying than the second question.

Questions with implicit alternatives should be avoided unless there are specific reasons for including them.[32] When the alternatives are close in preference or large in number, the alternatives at the end of the list have a greater chance of being selected. To overcome this bias, the split ballot technique should be used to rotate the order in which the alternatives appear.

Avoid implicit assumptions

Implicit assumptions
An assumption that is not explicitly stated in a question.

Questions should not be worded so that the answer is dependent on implicit assumptions about what will happen as a consequence. **Implicit assumptions** are assumptions that are not explicitly stated in the question, as in the following example.[33]

1 Are you in favour of a balanced national budget?
2 Are you in favour of a balanced national budget if it would result in an increase in personal income tax?

Implicit in question 1 are the consequences that will arise as a result of a balanced national budget. There might be a cut in defence expenditures, an increase in personal income tax, a cut in health spending, and so on. Question 2 is a better way to word this question. Question 1's failure to make its assumptions explicit would result in overestimating the respondents' support for a balanced national budget.

Avoid generalisations and estimates

Questions should be specific, not general. Moreover, questions should be worded so that the respondent does not have to make generalisations or compute estimates. Suppose that we were interested in households' annual per capita expenditure on groceries. If we asked respondents the question:

> What is the annual per capita expenditure on groceries in your household?

they would first have to determine the annual expenditure on groceries by multiplying the monthly expenditure on groceries by 12 or the weekly expenditure by 52. Then they would have to divide the annual amount by the number of persons in the household. Most respondents would be unwilling or unable to perform these calculations. A better way of obtaining the required information would be to ask the respondents two simple questions:

> What is the monthly (or weekly) expenditure on groceries in your household?

and

> How many members are there in your household?

The researcher can then perform the necessary calculations.

Use positive and negative statements

Many questions, particularly those measuring attitudes and lifestyles, are worded as statements to which respondents indicate their degree of agreement or disagreement. Evidence indicates that the response obtained is influenced by the directionality of the statements: whether they are stated positively or negatively. In these cases, it is better to use dual statements, some of which are positive and others negative. Two different questionnaires could be prepared. One questionnaire would contain half-negative and half-positive statements in an interspersed way. The direction of these statements would be reversed in the other questionnaire. An example of dual statements was provided in the summated Likert scale in Chapter 12 designed to measure attitudes towards Dresdner Bank.

Arrange the questions in proper order

The order of questions is of equal importance to the wording used in the questions. As noted in the last section, questions communicate and set respondents in a particular frame of mind. The frame of mind in which they are set affects how they perceive individual questions and respond to those questions. As well as understanding the characteristics of language in target respondents, questionnaire designers must be aware of the logical connections between questions – as perceived by target respondents. The following issues help to determine the order of questions.

Opening questions

The opening questions can be crucial in gaining the confidence and cooperation of respondents. They should be interesting, simple and non-threatening. Questions that ask respondents for their opinions can be good opening questions, because most people like to express their opinions. Sometimes such questions are asked although they are unrelated to the research problem and their responses are not analysed.[34] Though classification questions seem simple to start a questionnaire, issues like age, gender and income can be seen as sensitive. Opening a questionnaire with these questions tends to make respondents concerned about the purpose of these questions and indeed the whole survey.

Type of information

Classification information
Socio-economic and demographic characteristics used to classify respondents.

Identification information
A type of information obtained in a questionnaire that includes name, address and phone number.

The type of information obtained in a questionnaire may be classified as (1) basic information, (2) **classification information**, and (3) **identification information**. Basic information relates directly to the research problem. Classification information, consisting of socio-economic and demographic characteristics, is used to classify the respondents, understand the results and validate the sample (see Chapter 14). Identification information includes name, address and telephone number. Identification information may be obtained for a variety of purposes, including verifying that the respondents listed were actually interviewed and to send promised incentives or prizes. As a general guideline, basic information should be obtained first, followed by classification, and finally identification information. The basic information is of greatest importance to the research project and should be obtained first, before we risk alienating the respondents by asking a series of personal questions.

Difficult questions

Difficult questions or questions that are sensitive, embarrassing, complex or dull should be placed late in the sequence. After rapport has been established and the respondents become involved, they are less likely to object to these questions. Thus, in the GlobalCash Project, information about the banks that companies use and how they rated those banks was asked at the end of the section on basic information. Likewise, income should be the last question in the classification section (if it is to be used at all).

Effect on subsequent questions

Questions asked early in a sequence can influence the responses to subsequent questions. As a rule of thumb, general questions should precede specific questions. This prevents specific questions from biasing responses to the general questions. Consider the following sequence of questions:

Q1: *What considerations are important to you in selecting a bank?*

Q2: *In selecting a bank, how important is convenience of location?*

Note that the first question is general whereas the second is specific. If these questions were asked in the reverse order, respondents would be clued about convenience of location and would be more likely to give this response to the general question.

Going from general to specific is called the **funnel approach**. The funnel approach is particularly useful when information has to be obtained about respondents' general choice behaviour and their evaluations of specific products.[35] Sometimes the inverted funnel approach may be useful. In this approach, questioning starts with specific questions and concludes with the general questions. The respondents are compelled to provide specific information before making general evaluations. This approach is useful when respondents have no strong feelings or have not formulated a point of view.

Logical order

Questions should be asked in a logical order. This may seem a simple rule, but as the researcher takes time to understand respondents and how they use language, they should also take time to understand their logic, i.e. what 'logical order' means to target respondents. All questions that deal with a particular topic should be asked before beginning a new topic. When switching topics, brief transitional phrases should be used to help respondents switch their train of thought.

Branching questions should be designed carefully.[36] Branching questions direct respondents to different places in the questionnaire based on how they respond to the question at hand. These questions ensure that all possible contingencies are covered. They also help reduce interviewer and respondent error and encourage complete responses. Skip patterns based on the branching questions can become quite complex. A simple way to account for all contingencies is to prepare a flowchart of the logical possibilities and then develop branching questions and instructions based on it. A flowchart used to assess the use of electronic payments in clothes purchases via the Internet is shown in Figure 13.4.
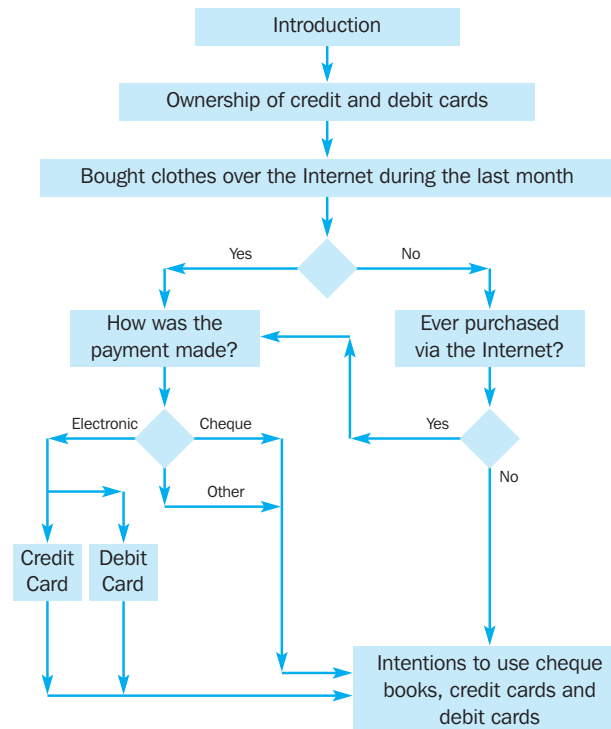
Figure 13.4
Flowchart for questionnaire design

Placement of branching questions is important and the following guidelines should be followed: (1) the question being branched (the one to which the respondent is being directed) should be placed as close as possible to the question causing the branching, and (2) the branching questions should be ordered so that the respondents cannot anticipate what additional information will be required. Otherwise, the respondents may discover that they can avoid detailed questions by giving certain answers to branching questions. For example, the respondents should first be asked if they have seen any of the listed commercials before they are asked to evaluate commercials. Otherwise, the respondents will quickly discover that stating that they have seen a commercial leads to detailed questions about that commercial and that they can avoid detailed questions by stating that they have not seen the commercial.

Identify the form and layout

The format, spacing and positioning of questions can have a significant effect on the results, particularly in self-administered questionnaires. It is good practice to divide a questionnaire into several parts. Several parts may be needed for questions pertaining to the basic information.

The questions in each part should be numbered, particularly when branching questions are used. Numbering of questions also makes the coding of responses easier. In addition, the questionnaires should preferably be pre-coded. In **pre-coding**, the codes to enter in the computer are printed on the questionnaire. Coding of questionnaires is explained in more detail in Chapter 17 on data preparation.

Pre-coding
In questionnaire design, assigning a code to every conceivable response before data collection.

The questionnaires themselves should be numbered serially. This facilitates the control of questionnaires in the field as well as the coding and analysis. Numbering makes it easy to account for the questionnaires and to determine whether any have been lost. A possible exception to this rule is mail questionnaires. If these are numbered, respondents assume that a given number identifies a particular respondent. Some respondents may refuse to participate or may answer differently under these conditions. However, recent research suggests that this loss of anonymity has little influence on the results.[37]

Reproduce the questionnaire

How a questionnaire is reproduced for administration can influence the results. For example, if the questionnaire is reproduced on poor-quality paper or is otherwise shabby in appearance, the respondents will think that the project is unimportant and the quality of response will be adversely affected. Therefore, the questionnaire should be reproduced on good-quality paper and have a professional appearance.

When a printed questionnaire runs to several pages, it should take the form of a booklet rather than a number of sheets of paper clipped or stapled together. Booklets are easier for the interviewer and the respondents to handle and do not easily come apart with use. They allow the use of a double-page format for questions and look more professional.

Each question should be reproduced on a single page (or double-page spread). A researcher should avoid splitting a question, including its response categories. Split questions can mislead the interviewer or the respondent into thinking that the question has ended at the end of a page. This will result in answers based on incomplete questions.

Vertical response columns should be used for individual questions. It is easier for interviewers and respondents to read down a single column rather than reading sideways across several columns. Sideways formatting and splitting, done frequently to conserve space, should be avoided.

The tendency to crowd questions together to make the questionnaire look shorter should be avoided. Overcrowded questions with little blank space between them can lead to errors in data collection and yield shorter and less informative replies. Moreover, they give the impression that the questionnaire is complex and can result in lower cooperation and completion rates. Although shorter questionnaires are more desirable than longer ones, the reduction in size should not be obtained at the expense of crowding.[38]

Directions or instructions for individual questions should be placed as close to the questions as possible. Instructions relating to how the question should be administered or answered by the respondent should be placed just before the question. Instructions concerning how the answer should be recorded or how the probing should be done should be placed after the question (for more information on probing and other interviewing procedures, see Chapter 16). It is common practice to distinguish instructions from questions by using different typefaces (such as capital or boldfaced letters).

Although colour does not influence response rates to questionnaires, it can be employed advantageously in some respects. Colour coding is useful for branching questions. The next question to which the respondent is directed is printed in a colour that matches the space in which the answer to the branching question was recorded. Surveys directed at different respondent groups can be reproduced on paper of a different colour. In the GlobalCash survey, the questionnaire was printed with different coloured covers to identify which country the response came from.

The questionnaire should be reproduced in such a way that it is easy to read and answer. The type should be large and clear. Reading the questionnaire should not impose a strain.

Eliminate problems by pilot-testing

Pilot-testing
Testing the questionnaire on a small sample of respondents for the purpose of improving the questionnaire by identifying and eliminating potential problems.

Pilot-testing refers to testing the questionnaire on a small sample of respondents to identify and eliminate potential problems.[39] Even the best questionnaire can be improved by pilot-testing. As a general rule, a questionnaire should not be used in the field survey without adequate pilot-testing.[40] A pilot-test should be extensive. All aspects of the questionnaire should be tested, including question content, wording, sequence, form and layout, question difficulty, and instructions. The respondents in the pilot-test should be similar to those who will be included in the actual survey in terms of background characteristics, familiarity with the topic, and attitudes and behaviours of interest.[41] In other words, respondents for the pilot-test and for the actual survey should be drawn from the same population.

Pilot-tests are best done by personal interviews, even if the actual survey is to be conducted by Internet, mail or telephone, because interviewers can observe respondents' reactions and attitudes. After the necessary changes have been made, another pilot-test could be conducted by Internet, mail or telephone if those methods are to be used in the actual survey. The latter pilot-tests should reveal problems peculiar to the interviewing method.

A variety of interviewers should be used for pilot-tests. The project director, the researcher who developed the questionnaire, and other key members of the research

team should conduct some pilot-test interviews. This will give them a good feel for potential problems and the nature of the expected data. Most of the pilot-test interviews, however, should be conducted by regular interviewers. It is good practice to employ both experienced and new interviewers. Experienced interviewers can easily perceive uneasiness, confusion and resistance in the respondents, and new interviewers can help the researcher identify interviewer-related problems. Ordinarily the pilot-test sample size is small, varying from 15 to 30 respondents for the initial testing, depending on the heterogeneity (e.g. a wide array of education levels) of the target population. The sample size can increase substantially if the pilot-testing involves several stages.

Protocol analysis and debriefing are two commonly used procedures in pilot-testing. In protocol analysis, the respondent is asked to 'think aloud' while answering the questionnaire, as explained in Chapter 12. Typically, the respondent's remarks are tape-recorded and analysed to determine the reactions invoked by different parts of the questionnaire. Debriefing occurs after the questionnaire has been completed. Respondents are told that the questionnaire they just completed was a pilot-test and the objectives of pilot-testing are described to them. They are then asked to describe the meaning of each question, to explain their answers, and to state any problems they encountered while answering the questionnaire.

Editing involves correcting the questionnaire for the problems identified during pilot-testing. After each significant revision of the questionnaire, another pilot-test should be conducted, using a different sample of respondents. Sound pilot-testing involves several stages. One pilot-test is a bare minimum. Pilot-testing should be continued until no further changes are needed.

Finally, the responses obtained from the pilot-test should be coded and analysed. The analysis of pilot-test responses can serve as a check on the adequacy of the problem definition and the data and analysis required to obtain the necessary information. The dummy tables prepared before developing the questionnaire will point to the need for the various sets of data. If the response to a question cannot be related to one of the pre-planned dummy tables, either those data are superfluous or some relevant analysis has not been foreseen. If part of a dummy table remains empty, a necessary question may have been omitted. Analysis of pilot-test data helps to ensure that all data collected will be utilised and that the questionnaire will obtain all the necessary data.[42]

Summarising the questionnaire design process

Table 13.1 summarises the questionnaire design process in the form of a checklist.

Table 13.1 Questionnaire design checklist

Step 1: Specify the Information Needed
1 Ensure that the information obtained fully addresses all the components of the problem. Review components of the problem and the approach, particularly the research questions, hypotheses and characteristics that influence the research design.
2 Prepare a set of dummy tables.
3 Have a clear idea of characteristics and motivations of the target population.

Step 2: Specify the Type of Interviewing Method
1 Review the type of interviewing method determined based on considerations discussed in Chapter 10.

Step 3: Determine the Content of Individual Questions
1 Is the question necessary?
2 Are several questions needed instead of one to obtain the required information in an unambiguous manner?
3 Do not use double-barrelled questions.

Step 4: Overcome the Respondent's Inability and Unwillingness to Answer

1 Is the respondent informed?

2 If the respondent is not likely to be informed, filter questions that measure familiarity, product use and past experience should be asked before questions about the topics themselves.

3 Can the respondent remember?

4 Avoid errors of omission, telescoping and creation.

5 Questions that do not provide the respondent with cues can underestimate the actual occurrence of an event.

6 Can the respondent articulate?

7 Minimise the effort required of the respondent.

8 Is the context in which the questions are asked appropriate?

9 Make the request for information seem legitimate.

10 If the information is sensitive:

    (a) Place sensitive topics at the end of the questionnaire.
    (b) Preface the question with a statement that the behaviour of interest is common.
    (c) Ask the question using the third-person technique.
    (d) Hide the question in a group of other questions that respondents are willing to answer.
    (e) Provide response categories rather than asking for specific figures.
    (f) Use randomised techniques, if appropriate.

Step 5: Choose Question Structure

1 Open-ended questions are useful in exploratory research and as opening questions.

2 Use structured questions whenever possible.

3 In multiple-choice questions, the response alternatives should include the set of all possible choices and should be mutually exclusive.

4 In a dichotomous question, if a substantial proportion of the respondents can be expected to be neutral, include a neutral alternative.

5 Consider the use of the split ballot technique to reduce order bias in dichotomous and multiple-choice questions.

6 If the response alternatives are numerous, consider using more than one question to reduce the information processing demands on the respondents.

Step 6: Choose Question Wording

1 Define the issue in terms of 'who', 'what', 'when' and 'where'.

2 Use ordinary words. Words should match the vocabulary level of the respondents.

3 Avoid ambiguous words: usually, normally, frequently, often, regularly, occasionally, sometimes, etc.

4 Avoid leading or biasing questions that cue the respondent to what the answer should be.

5 Avoid implicit alternatives that are not explicitly expressed in the options.

6 Avoid implicit assumptions.

7 Respondent should not have to make generalisations or compute estimates.

8 Use positive and negative statements.

Step 7: Arrange the Questions in the Proper Order

1 The opening questions should be interesting, simple and non-threatening.

2 Qualifying questions should serve as the opening questions.

3 Basic information should be obtained first, followed by classification and finally identification information.

4 Difficult, sensitive or complex questions should be placed late in the sequence.

5 General questions should precede specific questions.

6 Questions should be asked in a logical order.

7 Branching questions should be designed carefully to cover all possible contingencies.

8 The question being branched should be placed as close as possible to the question causing the branching, and the branching questions should be ordered so that the respondents cannot anticipate what additional information will be required.

Step 8: Identify the Form and Layout

1 Divide a questionnaire into several parts.

2 Questions in each part should be numbered.

3 The questionnaire should be pre-coded.

4 The questionnaires themselves should be numbered serially.

Step 9: Reproduce the Questionnaire

1 The questionnaire should have a professional appearance.

2 A booklet format should be used for long questionnaires.

3 Each question should be reproduced on a single page (or double-page spread).

4 Vertical response columns should be used.

5 Grids are useful when there are a number of related questions that use the same set of response categories.

6 The tendency to crowd questions to make the questionnaire look shorter should be avoided.

7 Directions or instructions for individual questions should be placed as close to the questions as possible.

Step 10: Eliminate Problems by Pilot-testing

1 Pilot-testing should always be done.

2 All aspects of the questionnaire should be tested, including question content, wording, sequence, form and layout, question difficulty and instructions.

3 The respondents in the pilot-test should be similar to those who will be included in the actual survey.

4 Begin the pilot-test by using personal interviews.

5 The pilot-test should also be conducted by mail or telephone if those methods are to be used in the actual survey.

6 A variety of interviewers should be used for pilot-tests.

7 The pilot-test sample size should be small, varying from 15 to 30 respondents for the initial testing.

8 Use protocol analysis and debriefing to identify problems.

9 After each significant revision of the questionnaire, another pilot-test should be conducted, using a different sample of respondents.

10 The responses obtained from the pilot-test should be coded and analysed.

International marketing research

The questionnaire or research instrument should be adapted to the specific cultural environment and should not be biased in terms of any one culture. This requires careful attention to each step of the questionnaire design process. The information needed should be clearly specified. It is important to take into account any differences in underlying consumer behaviour, decision-making processes, psychographics, lifestyles and demographic variables. In the context of demographic characteristics, information on marital status, education, household size, occupation, income and dwelling unit may have to be specified differently for different countries, as these variables may not be directly comparable across countries. For example, household definition and size varies greatly, given the extended family structure in some countries and the practice of two or even three families living under the same roof.

Although personal interviewing may dominate as a survey method in many Western countries, different survey methods may be favoured in different countries. Hence, the questionnaire may have to be suitable for administration by more than one method. For ease of comprehension and translation, it is desirable to have two or more simple questions rather than a single complex question. In overcoming the inability to answer, the variability in the extent to which respondents in different cultures are informed about the subject matter of the survey should be taken into account. Respondents in some parts of the world may not be as well informed on many issues as people in Europe.

The use of unstructured or open-ended questions may be desirable if the researcher lacks knowledge about the determinants of response in other countries. Because they do not impose any response alternatives, unstructured questions also

reduce cultural bias, but they are more affected by differences in educational levels than structured questions. They should be used with caution in countries with low literacy levels.

The questionnaire may have to be translated for administration in different cultures. The researcher must ensure that the questionnaires in different languages are equivalent. The special procedures designed for this purpose are discussed in Chapter 26. The following example illustrates the problems of translation.

e x a m p l e

If you board the Asian 'bus, better mind your language[43]

On the surface it would appear that the omnibus service could be made very standardised in Asia, even more so than in Europe where one must deal with a different language in almost every country. One might conclude that a questionnaire in Chinese could be used with little or no modification in a number of countries such as China, Hong Kong, Taiwan and Malaysia, thus avoiding the problems of timing, cost and inaccuracies generally associated with translations. However, due to vast differences in the region in terms of language, culture and geography, the 'standard' omnibus survey becomes less standard than would first meet the eye. While the omnibus is a very cost-effective and efficient way of conducting research in Asia, great care must be taken in order to make best use of this service, particularly by companies thinking of entering these markets. Countries such as Japan, Korea and Thailand are typical of South East Asia in that they each possess a single culture and a single language. But, in both Singapore and Malaysia, for example, omnibus questionnaires are always printed in three languages: English, Mandarin and Malay. In Malaysia, in addition to Mandarin which is spoken by all 'Chinese-literate' consumers, there are three other commonly spoken dialects which have to be dealt with at the respondent level: Cantonese, Hokkien and Hakka. In Singapore the commonly spoken dialects are Cantonese, Hokkien and Teochew. ■

Pilot-testing the questionnaire is complicated in international research because linguistic equivalence must be pilot-tested. Two sets of pilot-tests are recommended. The translated questionnaire should be pilot-tested on monolingual subjects in their native language, and the original and translated versions should also be administered to bilingual subjects. The pilot-test data from administration of the questionnaire in different countries or cultures should be analysed and the pattern of responses compared to detect any cultural biases.

Ethics in marketing research

The researcher must be mindful of the demands placed on the respondents when designing questionnaires. Because the administration of the questionnaire is a substantial intrusion by the researcher, several ethical concerns arise pertaining to the researcher–respondent relationship. Ethical issues impinging on the researcher–marketing decision-maker relationship may also have to be addressed.

In consideration of the respondents, exceedingly long questionnaires should be avoided. As a general guideline, the following are generally considered 'overly long': a personal interview in-home over 60 minutes, a telephone interview over 30 minutes and a street interview over 30 minutes.[44] Excessively long questionnaires are burdensome on the respondents and adversely affect the quality of responses. Similarly, questions that are confusing, exceed the respondents' ability, are difficult

or are otherwise improperly worded, should be avoided. Professional Perspective 12 'The field-good factor' on the Companion Website addresses many of these issues.

Overly sensitive questions deserve special attention. A real ethical dilemma exists for researchers investigating social problems such as poverty, drug use and sexually transmitted diseases like AIDS, or conducting studies of highly personal products like feminine hygiene products or financial products.[45] Candid and truthful responses are needed to generate meaningful results. But how do we obtain such data without asking sensitive questions that invade respondents' privacy? When asking sensitive questions, researchers should attempt to minimise the discomfort of the respondents. It should be made clear at the beginning of the questionnaire that respondents are not obligated to answer any question that makes them uncomfortable.[46]

One researcher–marketing decision-maker issue worth mentioning is piggybacking, which occurs when a questionnaire contains questions pertaining to more than one sponsoring organisation. One sponsor's questions may take up a part of the questionnaire, while a second sponsor's study takes up the rest. Although there is some risk that one study will contaminate the other or that the questionnaire may not be very coherent, piggybacking can substantially reduce the cost. Thus, it can be a good way for sponsors with limited research budgets to collect primary data they would not be able to afford otherwise. In these cases all sponsors must be aware of and consent to the arrangement. Unfortunately, piggybacking is sometimes used without disclosure to the sponsors for the sole purpose of increasing the researcher's profit. This is unethical.

Finally, the researcher has the ethical responsibility of designing the questionnaire so as to obtain the required information in an unbiased manner. Deliberately biasing the questionnaire in a desired direction – for example, by asking leading questions – cannot be condoned. In deciding the question structure, the most appropriate rather than the most convenient option should be adopted. Also, the questionnaire should be thoroughly pilot-tested before fieldwork begins, or an ethical breach has occurred.

Internet and computer applications

Word processing packages/software, such as Microsoft Word, are widely used for designing questionnaires, particularly if they are paper based. Researchers are familiar with their operation as they typically use this type of software to write project briefs and reports. However, although the finished result may be sufficiently presentable, the underlying structures associated with a survey are missing. These include elements such as single or multiple response on any tick box question, data validation and routing.

Even though word processing software is becoming increasingly more sophisticated, desktop publishing packages, such as QuarkXpress, provide added graphical functionality and are widely used within the design and print industry. Consequently, some researchers are advised to use this type of software. However, the learning curve is fairly steep and researchers rarely undertake design tasks more complex than can be achieved in a standard word processing package.

One word of caution, though – the print industry is very often Apple Mac based and the marketing research industry is very often PC based. Although software such

as QuarkXpress is available for both platforms, and questionnaires can be designed on a PC and then transferred to a Mac for typesetting and printing, the results are not always identical. Do be extremely careful in the proof checking process. If questionnaires are to be scanned, where the positioning of boxes is even more critical, then the process of proof checking will be even more important and consequently time consuming.

Software such as SNAP enables questionnaires to be designed and printed for all survey formats (paper, Web, CATI, CAPI, PDA) and can be likened to a word processor with an intelligent template. The processes of selecting boxes, single or multiple responses, routing, etc., are automated, and entire libraries (see the example of SurveyPaks on www.snapsurveys.com) exist to select individual questions or entire questionnaires. These include topics such as Customer Satisfaction, Human Resources, Travel and Tourism, Healthcare and Best Value. Questions can be selected, pasted in a survey and amendments made where necessary to remove or add new codes. The formatted questionnaire is then ready for the next stages of data collection and data analysis. The SNAP software is described in more detail and reviewed by Tim Macer in Professional Perspective 20 on the Companion Website. Visit www.camsp.com to download a demo of the software KEYPOINT2, which again allows for the design of questionnaires in all formats, the handling of survey responses and the analysis and presentation of results. Another Website that should be familiar from the review of qualitative data analysis software in Chapter 10 is www.scolari.co.uk. Click on the SphinxSurvey to evaluate the survey software.

➤➤➤
See Professional Perspective 20.

Internet questionnaires share many of the features of CAPI questionnaires. The questionnaire can be designed using a wide variety of stimuli such as graphics, pictures, advertisements, animations, sound clips and full-motion video. Moreover, the researcher can control the amount of time for which the stimuli are available to the respondents, and the number of times a respondent can access each stimulus. This greatly increases the range and complexity of questionnaires that can be administered over the Internet. As in the case of CATI and CAPI, complicated skip patterns can be programmed into the questionnaire. The questions can be personalised and answers to previous questions can be inserted into subsequent questions. The various types of scales, such as ordinal ranking scales, Likert scales, semantic differential scales and Stapel scales, can be utilised. Open-ended questions can also be posed using the Internet. The data collected may need coding after the survey process, as with all open-ended questions.

As the Internet becomes a more accepted method of administering surveys, the nature and expectations of the sponsors of marketing research increases. Their expectations can include the need to do the following.

- Allow respondents to take a break part way through completing a questionnaire and then return to it later.
- Include a password to control access to a survey.
- Include a user ID to restrict access and avoid multiple completions by the same respondent as well as accessing demographic data of participants on a panel.
- Log information on the time spent by each respondent on each individual question (known as paradata).

➤➤➤
See Professional Perspective 4.

All of these capabilities are becoming more and more standard.

To see a more detailed evaluation of the power and potential of Internet surveys, go to the Companion Website and read Professional Perspective 4 by Peter Wills.

Summary

A questionnaire has three objectives. It must translate the information needed into a set of specific questions the respondents can and will answer. It must motivate respondents to complete the interview. It must also minimise response error.

Designing a questionnaire is more of a craft than a science. This is primarily caused by the interrelationship of stages and the trade-offs that questionnaire designers make in balancing the source of ideas, question purposes, actual questions and question analyses. The steps involved in the questionnaire design process involve:

1 Specifying the information needed. Understanding what information decision-makers need.
2 Specifying the type of interviewing method. Understanding which means of eliciting the information will work best, given the research design constraints that the researcher has to work with.
3 Determining the content of individual questions. Understanding the purpose of each question and working out how a posed question may fulfil that purpose.
4 Overcoming the respondents' inability and unwillingness to answer questions. Understanding the process of approaching and questioning respondents – from their perspective. Knowing what benefits they get from taking part in the survey process.
5 Choosing the question structure. Understanding how individual questions help to elicit information from respondents and help them to express their feelings.
6 Choosing the question wording. Understanding the meaning of words from the perspective of the respondent.
7 Arranging the questions in a proper order. Understanding what 'proper' means from the perspective of the respondent. Recognising that, as each question is posed to a respondent and they think about their response, the respondent changes. Information is not only drawn out of respondents, it is communicated to them as each question is tackled.
8 Identifying the form and layout of the questionnaire. Understanding how in a self-completion scenario the form and layout motivate and help the respondent to answer the questions properly and honestly. Understanding how the form and layout help the interviewer to conduct and record the interview.
9 Reproducing the questionnaire. Understanding how the professional appearance of a questionnaire affects the perceived credibility and professional ability of researchers.
10 Eliminating problems by pilot-testing. Understanding that no matter how much experience the researcher has in designing questionnaires – the issues, respondent characteristics and context of questioning make each survey unique – pilot-testing is vital.

Questions

1 What is the purpose of the questionnaire?

2 What expectations does the marketing researcher have of potential questionnaire respondents – in terms of how they will react to the experience of completing a questionnaire?

3 What does the marketing researcher have to offer potential questionnaire respondents? Why should this question be considered?

4 How would you determine whether a specific question should be included in a questionnaire?

5 What are the reasons why respondents may be (a) unable to answer and (b) unwilling to answer the question asked?

6 Explain the errors of omission, telescoping and creation. What can be done to reduce such errors?

7 Explain the concepts of aided and unaided recall.

8 What can a researcher do to make the request for information seem legitimate?

9 What are the advantages and disadvantages of unstructured questions?

10 What are the issues involved in designing multiple-choice questions?

11 What are the guidelines available for deciding on question wording?

12 What is a leading question? Give an example.

13 What is the proper order for questions intended to obtain basic, classification and identification information?

14 What guidelines are available for deciding on the form and layout of a questionnaire?

15 Describe the issues involved in pilot-testing a questionnaire.

Notes

1 Dowding, P., Research, Fieldwork Supplement, July 2000, 4.

2 Abel, S., 'Finding the elusive young through co-operative parents', *ResearchPlus* (February 1996).

3 The founding reference to this subject is Payne, S.L., *The Art of Asking Questions* (Princeton, NJ: Princeton University Press, 1951).

4 These guidelines are drawn from several books on questionnaire design: Schuman, H. and Presser, S., *Questions and Answers in Attitude Surveys* (Thousand Oaks, CA: Sage, 1996); Fink, A., *How to Ask Survey Questions* (Thousand Oaks, CA: Sage, 1995); Sudman, S. and Bradburn, N.M., *Asking Questions* (San Francisco, CA: Jossey-Bass, 1983); Erdos, P.L., *Professional Mail Surveys* (Malabar, FL: Robert E. Krieger, 1983); Fowler Jr, F.J., Backstrom, C.H. and Hursh-Csar, G., *Survey Research* (Cambridge, MA: Wiley, 1981); Labau, P., *Advanced Questionnaire Design* (Orlando, FL: Abt Books, 1981); Dillman, D., *Mail and Telephone Surveys: The Total Design Method* (New York: Wiley, 1978); Blankenship, A.B., *Professional Telephone Surveys* (New York: McGraw-Hill, 1977), 94–5; Korhauser, A. and Sheatsley, P.B., 'Questionnaire construction and interview procedure', in Selltiz, C., Wrightsman, L.S. and Cook, S.W. (eds), *Research Methods in Social Relations*, 3rd edn (New York: Holt, Rinehart & Winston, 1976), 541–73.

5 Bourque, L.B. and Fielder, E.P., *How to Conduct Self-Administered and Mail Surveys* (Thousand Oaks, CA: Sage, 1995); Frey, J.H. and Oishi, S.M., *How to Conduct Interviews by Telephone and in Person* (Thousand Oaks, CA: Sage, 1995).

6 Semon, T.T., 'Asking "how important" is not enough', *Marketing News* 31(16) (4 August 1997), 19; Hague, P., 'Good and bad in questionnaire design', *Industrial Marketing Digest*, 12, Third Quarter 1987, 161–70.

7 Boyd Jr, H.W., Westfall, R. and Stasch, S.E., *Marketing Research: Text and Cases*, 7th edn (Homewood, IL: Irwin, 1989), 277.

8 Stapel, J., 'Observations: a brief observation about likability and interestingness of advertising', *Journal of Advertising Research* 34(2) (March/April 1994), 79–80; Bishop, G.E., Oldendick, R.W. and Tuchfarber, A.J., 'Effects of filter questions in public opinion surveys', *Public Opinion Quarterly* 46 (Spring 1982), 66–85.

9 Schneider, K.C. and Johnson, J.C., 'Link between response-inducing strategies and uninformed response', *Marketing Intelligence and Planning* 12(1) (1994), 29–36.

10 Dutka, S. and Frankel, L.R., 'Measuring response error', *Journal of Advertising Research* 37(1) (January/February 1997), 33–9; Haller, T., *Danger: Marketing Researcher at Work* (Westport, CT: Quotum Books, 1983), 149.

11 Menon, G., Raghubir, P. and Schwarz, N., 'Behavioural frequency judgments: an accessibility-diagnosticity framework', *Journal of Consumer Research* 22(2) (September 1995), 212–28; Cook, W.A., 'Telescoping and memory's other tricks', *Journal of Advertising Research* (February–March 1987), 5–8; Sudman, S., Finn, A. and Lannom, L., 'The use of bounded recall procedures in single interviews', *Public Opinion Quarterly* (Summer 1984), 520–4.

12 Hill, R.P., 'Researching sensitive topics in marketing – the special case of vulnerable populations', *Journal of Public Policy and Marketing* 1(1) (Spring 1995), 143–8.

13 Tourangeau, R. and Smith, T.W., 'Asking sensitive questions: the impact of data collection mode, question format, and question context', *Public Opinion Quarterly* 60(20) (Summer 1996), 275–304; Marquis, K.H., Marquis, M.S. and Polich, M.J., 'Response bias and reliability in sensitive topic surveys', *Journal of the American Statistical Association* (June 1986), 381–9.

14 Peterson, R.A., 'Asking the age question: a research note', *Public Opinion Quarterly* (Spring 1984), 379–83; Sheth, J.N., LeClaire Jr, A. and Wachspress, D., 'Impact of asking race information in mail surveys', *Journal of Marketing* (Winter 1980), 67–70.

15 For a recent application, see Burton, B.K. and Near, J.P., 'Estimating the incidence of wrongdoing and whistle-blowing: results of a study using randomized response technique', *Journal of Business Ethics* 14 (January 1995), 17–30.

16 Mukhopadhyay, P., 'A note on UMVU-estimation under randomized-response model', *Communications in Statistics – Theory and Methods* 26(10) (1997), 2415–20; Stem Jr, D.E. and Steinhorst, R.K., 'Telephone interview and mail questionnaire applications of the randomized response model', *Journal of the American Statistical Association* (September 1984), 555–64.

17 Newman, L.M., 'That's a good question', *American Demographics* (Marketing Tools) (June 1995), 10–13.

18 Luyens, S., 'Coding verbatims by computers', *Marketing Research: A Magazine of Management and Applications* 7(2) (Spring 1995), 20–5.

19 Mossholder, K.W., Settoon, R.P., Harris, S.G. and Armenakis, A.A., 'Measuring emotion in open-ended survey responses: an application of textual data analysis', *Journal of Management* 21(2) (1995), 335–55.

20 Fowler Jr, F.J., *Improving Survey Questions* (Thousand Oaks, CA: Sage 1995); Krosnick, J.A. and Alwin, D.E., 'An evaluation of a cognitive theory of response-order effects in survey measurement', *Public Opinion Quarterly* (Summer 1987).

21 Blunch, N.J., 'Position bias in multiple-choice questions', *Journal of Marketing Research* 21 (May 1984), 216–20, has argued that position bias in multiple-choice questions cannot be eliminated by rotating the order of the alternatives. This viewpoint is contrary to the common practice.

22 Schuman, H. and Presser, S., *Questions and Answers in Attitude Surveys* (Thousand Oaks, CA: Sage, 1996).

23 Herriges, J.A. and Shogren, J.F., 'Starting point bias in dichotomous choice valuation with follow-up questioning', *Journal of Environmental Economics and Management* 30(1) (January 1996), 112–31; Mizerski, R.W., Freiden, J.B. and Green Jr, R.C., 'The Effect of the "don't know" option on TV ad claim recognition tests', in *Advances in Consumer Research* 10 (Association for Consumer Research, 1983), 283–7.

24 Kalton, G. and Schuman, H., 'The effect of the question on survey responses: a review', *Journal of the Royal Statistical Society Series A*, 145, Part 1 (1982), 44–5.

25 McBurnett, M., 'Wording of questions affects responses to gun control issue', *Marketing News* 31(1) (6 January 1997), 12; Wanke, M., Schwarz, N. and Noelle-Neumann, E., 'Asking comparative questions: the impact of the direction of comparison', *Public Opinion Quarterly* 59(3) (Fall 1995), 347–72.

26 Etter, J.F. and Perneger, T.V., 'Analysis of nonresponse bias in a mailed health survey', *Journal of Clinical Epidemiology* 50(10) (October 1997), 1123–8; Omura, G.S., 'Correlates of item non-response', *Journal of the Market Research Society* (October 1983), 321–30; Presser, S., 'Is inaccuracy on factual survey items item-specific or respondent-specific?', *Public Opinion Quarterly* (Spring 1984), 344–55.

27 Stout, N.J., 'Questionnaire design workshop helps market researchers build better surveys', *Health Care Strategic Management* 12(7) (July 1994), 10–11.

28 Edmondson, B., 'How to spot a bogus poll', *American Demographics* 8(10) (October 1996), 10–15; O'Brien, J., 'How do market researchers ask questions?', *Journal of the Market Research Society* 26 (April 1984).

29 Semon, T.T., 'Ask simple questions to improve analysis of value perception', *Marketing News* 29(5) (27 February 1995), 32.

30 Abramson, P.R. and Ostrom, C.W., 'Question wording and partisanship', *Public Opinion Quarterly* 58(1) (Spring 1994), 21–48.

31 'Don't lead: you may skew poll results', *Marketing News* 30(12) (3 June 1996), H37.

32 Adamek, R.J., 'Public opinion and Roe v. Wade: measurement difficulties', *Public Opinion Quarterly* 58(3) (Fall 1994), 409–18; Neumann, E.N. and Worcester, B., 'International opinion research', *European Research* (July 1984), 124–31.

33 Jacoby, J. and Szybillo, G.J., 'Consumer research in FTC versus Kraft (1991): a case of heads we win, tails you lose?', *Journal of Public Policy and Marketing* 14(1) (Spring 1995), 1–14; Jaffe, E.D. and Nebenzahl, I.D., 'Alternative questionnaire formats for country image studies', *Journal of Marketing Research* (November 1984), 463–71.

34 Schuman, H. and Presser, S., *Questions and Answers in Attitude Surveys* (Thousand Oaks, CA: Sage, 1996); Krosnick, J.A. and Alwin, D.E., 'An evaluation of a cognitive theory of response-order effects in survey measurement', *Public Opinion Quarterly* (Summer 1987), 201–19.

35 Rating a brand on specific attributes early in a survey may affect responses to a later overall brand evaluation. For example, see Bickart, B.A., 'Carryover and backfire effects in marketing research', *Journal of Marketing Research* 30 (February 1993), 52–62. See also McAllister, I. and Wattenberg, M.P., 'Measuring levels of party identification: does question order matter?' *Public Opinion Quarterly* 59(2) (Summer 1995), 259–68.

36 Willits, F.K. and Ke, B., 'Part-whole question order effects: views of rurality', *Public Opinion Quarterly* 59(3) (Fall 1995), 392–403; Messmer, D.J. and Seymour, D.J., 'The effects of branching on item non-response', *Public Opinion Quarterly* 46 (Summer 1982), 270–7.

37 Milne, G.R., 'Consumer participation in mailing lists: a field experiment', *Journal of Public Policy and Marketing* 16(2) (Fall 1997), 298–309.

38 Dickinson, S.R. and Kirzner, E., 'Questionnaire item omission as a function of within-group question position', *Journal of Business Research* (February 1985), 71–5; Herzog, A.R. and Bachman, J.G., 'Effects of questionnaire length on response quality', *Public Opinion Quarterly* 45 (Winter 1981), 549–59.

39 Martin, E. and Polivka, A.E., 'Diagnostics for redesigning survey questionnaires – measuring work in the current population survey', *Public Opinion Quarterly* 59(4) (Winter 1995), 547–67.

40 Mohrle, M.G., 'Empirical testing of a computer-based dialog questionnaire – 11 design rules for successful usage', *Wirtschaftsinformatik* 39(5) (October 1997), 461.

41 Diamantopoulos, A., Schlegelmilch, B.B. and Reynolds, N., 'Pre-testing in questionnaire design: the impact of respondent characteristics on error detection', *Journal of the Market Research Society* 36 (October 1994), 295–314.

42 Reynolds, N., Diamantopoulos, A. and Schlegelmilch, B.B., 'Pre-testing in questionnaire design: a review of the literature and suggestions for further research', *Journal of the Market Research Society* 35 (April 1993), 171–82.

43 Hutton, G., 'If you board the Asian bus, better mind your language', *Research Plus* (February 1996), 7.

44 *Rules of Conduct and Good Practice of the Professional Marketing Research Society of Canada* (1984).

45 Laczniak, G.R. and Murphy, P.E., *Ethical Marketing Decisions: the Higher Road* (Needham Heights, MA: Allyn and Bacon, 1993).

46 Morris, M.H., Marks, A.S., Allen, J.A. and Peery, N.S., 'Modeling ethical attitudes and behaviours under conditions of environmental turbulence – case of South Africa', *Journal of Business Ethics* 15(10) (October 1996), 1119–30; Laczniak, G.R. and Murphy, P.E., *Ethical Marketing Decisions: the Higher Road* (Needham Heights, MA: Allyn and Bacon, 1993).

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Sampling: design and procedures

Objectives

After reading this chapter, you should be able to:

1 differentiate a sample from a census and identify the conditions that favour the use of a sample versus a census;

2 discuss the sampling design process: definition of the target population, determination of the sampling frame, selection of sampling technique(s), determination of sample size, execution of the sampling process and validating the sample;

3 classify sampling techniques as non-probability and probability sampling techniques;

4 describe the non-probability sampling techniques of convenience, judgemental, quota and snowball sampling;

5 describe the probability sampling techniques of simple random, systematic, stratified and cluster sampling;

6 identify the conditions that favour the use of non-probability sampling versus probability sampling;

7 understand the sampling design process and the use of sampling techniques in international marketing research;

8 identify the ethical issues related to the sampling design process and the use of appropriate sampling techniques.

There is no hope of making scientific statements about a population based on the knowledge obtained from a sample, unless we are circumspect in choosing a sampling method.

Overview

Sampling is a key component of any research design. Sampling design involves several basic questions:

1 Should a sample be taken?
2 If so, what process should be followed?
3 What kind of sample should be taken?
4 How large should it be?
5 What can be done to control and adjust for non-response errors?

This chapter introduces the fundamental concepts of sampling and the qualitative considerations necessary to answer these questions. We address the question of whether or not to sample and describe the steps involved in sampling. Next, we present non-probability and probability sampling techniques. We discuss the use of sampling techniques in international marketing research and identify the relevant ethical issues. Statistical determination of sample size, and the causes for, control of and adjustments for non-response error are discussed in Chapter 15.

We begin with the following example, which illustrates the usefulness of sampling.

example

Can football give brands 110% recall?[1]

The European Football Championship held in England in 1996 proved that football is the world's number one sport. More than 1.3 million tickets were sold and 250,000 overseas visitors followed their teams to England. The 31 matches were watched in more than 190 countries around the world by a global cumulative television audience of 6.7 billion; 445 million tuned in to watch the final itself.

Commercial support for the event came from a group of 11 official sponsors: Canon, Carlsberg, Coca-Cola, Fuji Film, General Motors, JVC, McDonald's, MasterCard, Philips, Snickers and Umbro. Such a weight of marketing activity represented a major investment for the sponsors and it was vital that they were able to evaluate its effectiveness. Many individual sponsors conducted their own studies, but a study was conducted for the sponsor group to evaluate the impact of EURO 96 as a sponsorship opportunity.

The survey to evaluate sponsorship effectiveness was a two-wave study, the first taking place eight months prior to the tournament, the second immediately after the final. The research was carried out in four countries, the national football teams of each one having successfully qualified for the finals, thus ensuring a consistent level of interest in the tournament across the sample:

UK: Host nation
France: Host to the 98 World Cup
Germany: Tournament favourites
Russia: East European nation

In each market a sample of 500 individuals aged 12 to 65 was interviewed at each wave, quota controls being imposed on sex, age and socio-economic group to match the incidence of each within the general population. ■

This example illustrates the various aspects of sampling design, clearly founded in a marketing and research problem. The target population is defined, the sampling technique is established with criteria for quotas, and the sample size for each country and survey wave is set.

Imagine if the researchers in the above example had attempted to survey every viewer, i.e. to conduct a census. The cost involved and the time needed to complete the task would have been astronomical. Even if the sponsors could afford to undertake a census, would it give them a more accurate view of the television audience? In

certain circumstances administering a survey to the whole of a target makes sense; to an audience of 6.7 billion it does not.

Sample or census

Population
The aggregate of all the elements, sharing some common set of characteristics, that comprise the universe for the purpose of the marketing research problem.

Census
A complete enumeration of the elements of a population or study objects.

Sample
A subgroup of the elements of the population selected for participation in the study.

The objective of most marketing research projects is to obtain information about the characteristics or parameters of a **population**. A population is the aggregate of all the elements that share some common set of characteristics and that comprise the universe for the purpose of the marketing research problem. The population parameters are typically numbers, such as the proportion of consumers who are loyal to a particular brand of toothpaste. Information about population parameters may be obtained by taking a **census** or a **sample**. A census involves a complete enumeration of the elements of a population. The population parameters can be calculated directly in a straightforward way after the census is enumerated. A sample, on the other hand, is a subgroup of the population selected for participation in the study. Sample characteristics, called statistics, are then used to make inferences about the population parameters. The inferences that link sample characteristics and population parameters are estimation procedures and tests of hypotheses. These inference procedures are considered in Chapters 18 to 24.

Table 14.1 **Sample versus census**

Factors	Conditions favouring the use of	
	Sample	*Census*
1 Budget	Small	Large
2 Time available	Short	Long
3 Population size	Large	Small
4 Variance in the characteristic	Small	Large
5 Cost of sampling errors	Low	High
6 Cost of non-sampling errors	High	Low
7 Nature of measurement	Destructive	Non-destructive
8 Attention to individual cases	Yes	No

Table 14.1 summarises the conditions favouring the use of a sample versus a census. Budget and time limits are obvious constraints favouring the use of a sample. A census is both costly and time-consuming to conduct. A census is unrealistic if the population is large, as it is for most consumer products. In the case of many industrial products, however, the population is small, making a census feasible as well as desirable. For example, in investigating the use of certain machine tools by Italian car manufacturers, a census would be preferred to a sample. Another reason for preferring a census in this case is that variance in the characteristic of interest is large. For example, machine tool usage of Fiat may vary greatly from the usage of Ferrari. Small population sizes as well as high variance in the characteristic to be measured favour a census.

If the cost of sampling errors is high (e.g. if the sample omitted a major manufacturer like Ford, the results could be misleading) a census, which eliminates such errors, is desirable. If the cost of non-sampling errors is high (e.g. interviewers incorrectly questioning target respondents) a sample, where fewer resources would have been spent, would be favoured.

A census can greatly increase non-sampling error to the point that these errors exceed the sampling errors of a sample. Non-sampling errors are found to be the major contributor to total error, whereas random sampling errors have been relatively small in magnitude.[2] Hence, in most cases, accuracy considerations would favour a sample over a census.

A sample may be preferred if the measurement process results in the destruction or contamination of the elements sampled. For example, product usage tests result in the consumption of the product. Therefore, taking a census in a study that requires households to use a new brand of photographic film would not be feasible. Sampling may also be necessary to focus attention on individual cases, as in the case of depth interviews. Finally, other pragmatic considerations, such as the need to keep the study secret, may favour a sample over a census.

The sampling design process

The sampling design process includes six steps, which are shown sequentially in Figure 14.1. These steps are closely interrelated and relevant to all aspects of the marketing research project, from problem definition to the presentation of the results. Therefore, sample design decisions should be integrated with all other decisions in a research project.[3]

Define the target population

Target population
The collection of elements or objects that possess the information sought by the researcher and about which inferences are to be made.

Element
An object that possesses the information sought by the researcher and about which inferences are to be made.

Sampling unit
An element, or a unit containing the element, that is available for selection at some stage of the sampling process.

Sampling design begins by specifying the **target population**. This is the collection of elements or objects that possess the information sought by the researcher and about which inferences are to be made. The target population must be defined precisely. Imprecise definition of the target population will result in research that is ineffective at best and misleading at worst. Defining the target population involves translating the problem definition into a precise statement of who should and should not be included in the sample.

The target population should be defined in terms of elements, sampling units, extent and time. An **element** is the object about which or from which the information is desired. In survey research, the element is usually the respondent. A **sampling unit** is an element, or a unit containing the element, that is available for selection at some stage of the sampling process. Suppose that Clinique wanted to assess consumer

Define the target population

Determine the sampling frame

Select sampling techniques(s)

Determine the sample size

Execute the sampling process

Validate the sample
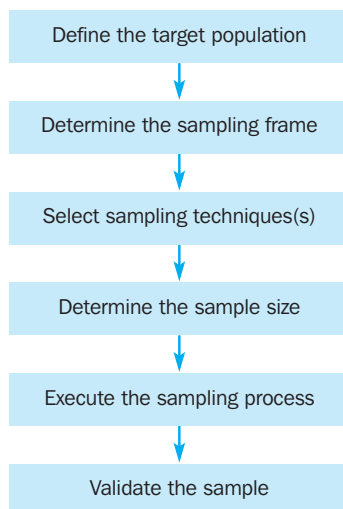
Figure 14.1
The sampling design process

response to a new line of lipsticks and wanted to sample females over 25 years of age. It may be possible to sample females over 25 directly, in which case a sampling unit would be the same as an element. Alternatively, the sampling unit might be households. In the latter case, households would be sampled and all females over 25 in each selected household would be interviewed. Here, the sampling unit and the population element are different. Extent refers to the geographical boundaries of the research, and the time refers to the period under consideration. We use the GlobalCash Project to illustrate.

Target population

The target population for the GlobalCash Project was defined as follows:

Elements: managers responsible for cash management decisions
Sampling units: the largest companies and non-banking financial institutions in a country
Extent: 19 European countries plus the centres of taxation concessions in Ireland and Belgium
Time: 2002

Defining the target population may not be as easy as it was in this example. Consider a marketing research project assessing consumer response to a new brand of men's moisturiser. Who should be included in the target population? All men? Men who have used a moisturiser during the last month? Men of 17 years of age or older? Should females be included, because some women buy moisturiser for men whom they know? These and similar questions must be resolved before the target population can be appropriately defined.[4]

Determine the sampling frame

Sampling frame
A representation of the elements of the target population that consists of a list or set of directions for identifying the target population.

A **sampling frame** is a representation of the elements of the target population. It consists of a list or set of directions for identifying the target population. Examples of a sampling frame include the telephone book, an association directory listing the firms in an industry, a customer database, a mailing list on a database purchased from a commercial organisation, a city directory, or a map.[5] If a list cannot be compiled, then at least some directions for identifying the target population should be specified, such as random-digit dialling procedures in telephone surveys.

Often it is possible to compile or obtain a list of population elements, but the list may omit some elements of the population or may include other elements that do not belong. Therefore, the use of a list will lead to sampling frame error, which was discussed in Chapter 3.[6]

In some instances, the discrepancy between the population and the sampling frame is small enough to ignore. In most cases, however, the researcher should recognise and attempt to treat the sampling frame error. The main approach is to redefine the population in terms of the sampling frame. For example, if a telephone directory is used as a sampling frame, the population of households could be redefined as those with a correct listing in a given area. Although this approach is simplistic, it does prevent the researcher from being misled about the actual population being investigated.[7]

Ultimately, the major drawback of redefining the population based upon available sampling frames is that the nature of the research problem may be compromised. Who is being measured and ultimately to whom the research findings may be generalised may not match the target group of individuals identified in a research problem definition. Evaluating the accuracy of sampling frames matches the issues of evaluating the quality of secondary data (see Chapter 4).

Select a sampling technique

Selecting a sampling technique involves several decisions of a broader nature. The researcher must decide whether to use a Bayesian or traditional sampling approach, to sample with or without replacement, and to use non-probability or probability sampling.

In the **Bayesian approach**, the elements are selected sequentially. After each element is added to the sample, the data are collected, sample statistics computed and sampling costs determined. The Bayesian approach explicitly incorporates prior information about population parameters as well as the costs and probabilities associated with making wrong decisions. This approach is theoretically appealing. Yet it is not used widely in marketing research because much of the required information on costs and probabilities is not available. In the traditional sampling approach, the entire sample is selected before data collection begins. Because the traditional approach is the most common approach used, it is assumed in the following sections.

In **sampling with replacement**, an element is selected from the sampling frame and appropriate data are obtained. Then the element is placed back in the sampling frame. As a result, it is possible for an element to be included in the sample more than once. In **sampling without replacement**, once an element is selected for inclusion in the sample, it is removed from the sampling frame and therefore cannot be selected again. The calculation of statistics is done somewhat differently for the two approaches, but statistical inference is not very different if the sampling frame is large relative to the ultimate sample size. Thus, the distinction is important only when the sampling frame is small compared with the sample size.

The most important decision about the choice of sampling technique is whether to use non-probability or probability sampling. Non-probability sampling relies on the judgement of the researcher, while probability sampling relies on chance. Given its importance, the issues involved in this decision are discussed in detail below, in the section headed 'A classification of sampling techniques'.

If the sampling unit is different from the element, it is necessary to specify precisely how the elements within the sampling unit should be selected. With in-home personal interviews and telephone interviews, merely specifying the address or the telephone number may not be sufficient. For example, should the person answering the doorbell or the telephone be interviewed, or someone else in the household? Often, more than one person in a household may qualify. For example, both the male and female head of household, or even their children, may be eligible to participate in a study examining family leisure-time activities. When a probability sampling technique is being employed, a random selection must be made from all the eligible persons in each household. A simple procedure for random selection is the 'next birthday' method. The interviewer asks which of the eligible persons in the household has the next birthday and includes that person in the sample.

Determine the sample size

Sample size refers to the number of elements to be included in the study. Determining the sample size involves several qualitative and quantitative considerations. The qualitative factors are discussed in this section, and the quantitative factors are considered in Chapter 15. Important qualitative factors to be considered in determining the sample size include (1) the importance of the decision, (2) the nature of the research, (3) the number of variables, (4) the nature of the analysis, (5) sample sizes used in similar studies, (6) incidence rates, (7) completion rates, and (8) resource constraints.

In general, for more important decisions, more information is necessary, and that information should be obtained very precisely. This calls for larger samples, but as the

Bayesian approach
A selection method where the elements are selected sequentially. The Bayesian approach explicitly incorporates prior information about population parameters as well as the costs and probabilities associated with making wrong decisions.

Sampling with replacement
A sampling technique in which an element can be included in the sample more than once.

Sampling without replacement
A sampling technique in which an element cannot be included in the sample more than once.

Sample size
The number of elements to be included in a study.

sample size increases, each unit of information is obtained at greater cost. The degree of precision may be measured in terms of the standard deviation of the mean, which is inversely proportional to the square root of the sample size. The larger the sample, the smaller the gain in precision by increasing the sample size by one unit.

The nature of the research also has an impact on the sample size. For exploratory research designs, such as those using qualitative research, the sample size is typically small. For conclusive research, such as descriptive surveys, larger samples are required. Likewise, if data are being collected on a large number of variables, i.e. many questions are asked in a survey, larger samples are required. The cumulative effects of sampling error across variables are reduced in a large sample.

If sophisticated analysis of the data using multivariate techniques is required, the sample size should be large. The same applies if the data are to be analysed in great detail. Thus, a larger sample would be required if the data are being analysed at the subgroup or segment level than if the analysis is limited to the aggregate or total sample.

Sample size is influenced by the average size of samples in similar studies. Table 14.2 gives an idea of sample sizes used in different marketing research studies. These sample sizes have been determined based on experience and can serve as rough guidelines, particularly when non-probability sampling techniques are used.

Table 14.2 Usual sample sizes used in marketing research studies

Type of study	Minimum size	Typical range
Problem identification research (e.g. market potential)	500	1000–2500
Problem-solving research (e.g. pricing)	200	300–500
Product tests	200	300–500
Test marketing studies	200	300–500
TV, radio or print advertising (per commercial/ad tested)	150	200–300
Test-market audits	10 stores	10–20 stores
Focus groups	2 groups	4–12 groups

➤➤➤
See Professional Perspective 12.

The sample size required should be adjusted for the incidence of eligible respondents and the completion rate (see Professional Perspective 12 by Pat Dowding on the Companion Website for a fuller discussion and statistics to support these points). Finally, the sample size decision should be guided by a consideration of the resource constraints. In any marketing research project, money and time are limited. The decisions involved in determining the sample size are covered in detail in the next chapter.

Execute the sampling process

Execution of the sampling process requires a detailed specification of how the sampling design decisions with respect to the population, sampling unit, sampling frame, sampling technique and sample size are to be implemented. Whilst an individual researcher may know how they are going to execute their sampling process, once more than one individual is involved, a specification for execution is needed to ensure that the process is conducted in a consistent manner. For example, if households are the sampling unit, an operational definition of a household is needed. Procedures should be specified for empty housing units and for call-backs in case no one is at home.

Validate the sample

Sample validation aims to account for sampling frame error by screening the respondents in the data collection phase. Respondents can be screened with respect to demographic characteristics, familiarity, product usage and other characteristics to

ensure that they satisfy the criteria for the target population. Screening can eliminate inappropriate elements contained in the sampling frame, but it cannot account for elements that have been omitted. The success of the validation process depends upon the accuracy of base statistics that describe the structure of a target population.

Once data are collected from a sample, comparisons between the structure of the sample and the target population should be made. Once data have been collected and it is found that the structure of a sample does not match the target population, a weighting scheme can be used (this is discussed in Chapter 17).

The steps involved in the sampling design process are illustrated in the following example based upon a continuous tracking survey.

e x a m p l e

Taking the people's temperature – right across Europe[8]

Since the earliest days of the European Economic Community, the Commission's information, communication, culture and audio-visual directorate, Directorate General X, has conducted regular polls across Europe. In 1996 the need for speed and flexibility in EU surveys led to the introduction of the Continuous Tracking Survey or CTS. The origin of the CTS was a policy initiative on information and communication, established in 1993 as a result of the decline in support for the EU in the period leading up to the Single Market and the ratification of the Maastricht Treaty.

The sample design for this study involved:

1 *Target population*: adults aged 18 years and over (element) in a household with a working telephone number (sampling unit) in individual EU countries (extent) during the survey period (time).
2 *Sampling frame*: computer program for generating random digit dialling (except in Germany where the code of practice forbids this approach).
3 *Sampling unit*: working telephone numbers.
4 *Sampling technique*: random sampling.
5 *Sample size*: 800 in each of 16 sampling areas making a total of 12,800 interviews in each four-week wave of interviewing – a total of 140,000 interviews each year.
6 *Execution*: CATI, random digit dialling followed by the random selection of individuals in households. 19 variants of the questionnaire to include people living in countries where more than one national language is common.
7 *Validation*: Sample characteristics compared with census statistics in each country. ■

A classification of sampling techniques

Sampling techniques may be broadly classified as non-probability and probability (see Figure 14.2).

Non-probability sampling
Sampling techniques that do not use chance selection procedures but rather rely on the personal judgement of the researcher.

Non-probability sampling relies on the personal judgement of the researcher rather than on chance to select sample elements. The researcher can arbitrarily or consciously decide what elements to include in the sample. Non-probability samples may yield good estimates of the population characteristics, but they do not allow for objective evaluation of the precision of the sample results. Because there is no way of determining the probability of selecting any particular element for inclusion in the sample, the estimates obtained are not statistically projectable to the population. Commonly used non-probability sampling techniques include convenience sampling, judgemental sampling, quota sampling and snowball sampling.

Probability sampling
A sampling procedure in which each element of the population has a fixed probabilistic chance of being selected for the sample.

In **probability sampling**, sampling units are selected by chance. It is possible to pre-specify every potential sample of a given size that could be drawn from the population, as well as the probability of selecting each sample. Every potential sample need not have the same probability of selection, but it is possible to specify the probability
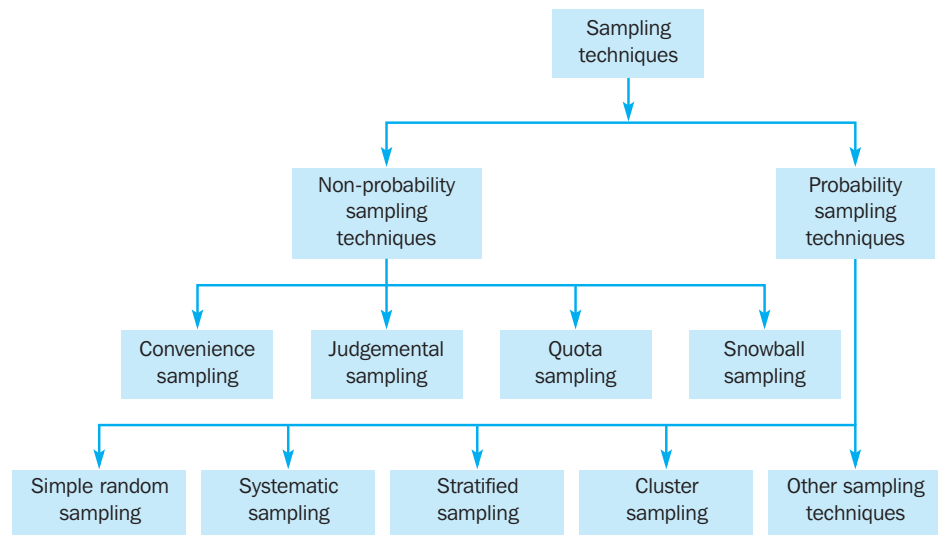
Figure 14.2 A classification of sampling techniques

of selecting any particular sample of a given size. This requires not only a precise definition of the target population but also a general specification of the sampling frame. Because sample elements are selected by chance, it is possible to determine the precision of the sample estimates of the characteristics of interest. **Confidence intervals**, which contain the true population value with a given level of certainty, can be calculated. This permits the researcher to make inferences or projections about the target population from which the sample was drawn. Classification of probability sampling techniques is based on:

Confidence intervals
The range into which the true population parameter will fall, assuming a given level of confidence.

- Element versus cluster sampling
- Equal unit probability versus unequal unit probability
- Unstratified versus stratified selection
- Random versus systematic selection
- One-stage versus multistage techniques.

All possible combinations of these five aspects result in 32 different probability sampling techniques. Of these techniques, we consider simple random sampling, systematic sampling, stratified sampling and cluster sampling in depth and briefly touch on some others. First, however, we discuss non-probability sampling techniques.

Non-probability sampling techniques

Convenience sampling

Convenience sampling
A non-probability sampling technique that attempts to obtain a sample of convenient elements. The selection of sampling units is left primarily to the interviewer.

Convenience sampling attempts to obtain a sample of convenient elements. The selection of sampling units is left primarily to the interviewer. Often, respondents are selected because they happen to be in the right place at the right time. Examples of convenience sampling include: (1) use of students, church groups and members of social organisations, (2) street interviews without qualifying the respondents, (3) some forms of email and Internet survey, (4) tear-out questionnaires included in a newspaper or magazine, and (5) journalists interviewing 'people on the street'.[9]

Convenience sampling is the least expensive and least time-consuming of all sampling techniques. The sampling units are accessible, easy to measure and cooperative. Despite these advantages, this form of sampling has serious limitations. Many potential sources of selection bias are present, including respondent self-selection.

363

Convenience samples are not representative of any definable population. Hence, it is not theoretically meaningful to generalise to any population from a convenience sample, and convenience samples are not appropriate for marketing research projects involving population inferences. Convenience samples are not recommended for descriptive or causal research, but they can be used in exploratory research for generating ideas, insights or hypotheses. Convenience samples can be used for pre-testing questionnaires, or pilot studies. Even in these cases, caution should be exercised in interpreting the results.

Judgemental sampling

Judgemental sampling
A form of convenience sampling in which the population elements are purposely selected based on the judgement of the researcher.

Judgemental sampling is a form of convenience sampling in which the population elements are selected based on the judgement of the researcher. The researcher, exercising judgement or expertise, chooses the elements to be included in the sample because he or she believes that they are representative of the population of interest or are otherwise appropriate. Common examples of judgemental sampling include: (1) test markets selected to determine the potential of a new product, (2) purchase engineers selected in industrial marketing research because they are considered to be representative of the company, (3) product testing with individuals who may be particularly fussy or who hold extremely high expectations, (4) expert witnesses used in court, and (5) supermarkets selected to test a new merchandising display system. The use of this technique is illustrated in the context of the GlobalCash Project.

e x a m p l e
GlobalCash Project

Judgemental sampling

In the GlobalCash study, at least two cash managers in every country surveyed were selected for an additional interview. The selection of these managers was based on the judgement of members of the research team. Many of the companies and indeed cash manager respondents were known to the research team. If the individual manager was not known, characteristics of the respondent organisations were known. The purpose of the interviews was to explore the reasons behind why cash managers carry out specific activities or plan certain events. To be able to fulfil this purpose managers were needed who:

■ were willing to give up at least two hours to discuss issues
■ were able to articulate the reasons for their behaviour
■ were 'sophisticated' in their approach to cash management
■ had a wide array of experience of pan-European banks.

Perhaps the most subjective element of this list is the issue of 'sophistication'. This does not necessarily mean the most technically complex or the use of state-of-the-art technology, but the use of creative solutions to cash management problems. In this example, judgement was used to select specific managers who met the above criteria. ■

Judgemental sampling is inexpensive, convenient and quick, yet it does not allow direct generalisations to a specific population, usually because the population is not defined explicitly. Judgemental sampling is subjective and its value depends entirely on the researcher's judgement, expertise and creativity. It can be useful if broad population inferences are not required. As in the GlobalCash example, judgement samples are frequently used in business-to-business marketing research projects.

Quota sampling

Quota sampling
A non-probability sampling technique that is a two-stage restricted judgemental sampling. The first stage consists of developing control categories or quotas of population elements. In the second stage, sample elements are selected based on convenience or judgement.

Quota sampling may be viewed as two-stage restricted judgemental sampling that is used extensively in street interviewing. The first stage consists of developing control characteristics, or quotas, of population elements such as age or gender. To develop these quotas, the researcher lists relevant control characteristics and determines the

distribution of these characteristics in the target population, such as Males 49%, Females 51% (resulting in 490 men and 510 women being selected in a sample of 1,000 respondents). Often, the quotas are assigned so that the proportion of the sample elements possessing the control characteristics is the same as the proportion of population elements with these characteristics. In other words, the quotas ensure that the composition of the sample is the same as the composition of the population with respect to the characteristics of interest.

In the second stage, sample elements are selected based on convenience or judgement. Once the quotas have been assigned, there is considerable freedom in selecting the elements to be included in the sample. The only requirement is that the elements selected fit the control characteristics. This technique is illustrated with the following example.

_e x a m p l e_

How is epilepsy perceived?

A study was undertaken by the Scottish Epilepsy Association to determine the perceptions of the condition of epilepsy by the adult population in the city of Glasgow. A quota sample of 500 adults was selected. The control characteristics were gender, age and propensity to donate to a charity. Based on the composition of the adult population of the city, the quotas assigned were as follows:

		Male		Female		Totals
		48%		52%		
Propensity to donate		_Have a flag_	_No flag_	_Have a flag_	_No flag_	
Age		_50%_	_50%_	_50%_	_50%_	
18 to 30	25%	30	30	33	32	**125**
31 to 45	40%	48	48	52	52	**200**
46 to 60	15%	18	18	19	20	**75**
Over 60	20%	24	24	26	26	**100**
Totals		**120**	**120**	**130**	**130**	
Totals		**240**		**260**		**500**

Note that the percentages of gender and age within the target population can be taken from local census statistics. The percentages of 'propensity to donate' could not be gleaned from secondary data sources and so were split on a 50/50 basis. The interviews were conducted on a Saturday when it was customary to see charity 'flag sellers' operating. One of the hypotheses to be tested in the study was the extent to which those who donated to charities on flag days were more aware of the condition of epilepsy and how to treat epileptic sufferers. Thus the instruction to interviewers was to split interviews between those who wore the 'flag' that they had bought from a street collector and those who had not bought a flag. It was recognised that this was a crude measure of propensity to donate to a charity but was the only tangible clue that could be consistently observed. ■

In this example, quotas were assigned such that the composition of the sample mirrored the population. In certain situations, however, it is desirable to either under- or over-sample elements with certain characteristics. To illustrate, it may be desirable to over-sample heavy users of a product so that their behaviour can be examined in detail. Although this type of sample is not representative, it may nevertheless be very relevant to allow a particular group of individuals to be broken down into sub-categories and analysed in depth.

Even if the sample composition mirrors that of the population with respect to the control characteristics, there is no assurance that the sample is representative. If a characteristic that is relevant to the problem is overlooked, the quota sample will not be representative. Relevant control characteristics are often omitted because there are practical difficulties associated with including certain control characteristics. For example, suppose a sample was sought that was representative of the different strata of socio-economic classes in a population. Imagine street interviewers approaching potential respondents who they believe would fit into the quota they have been set. Could an interviewer 'guess' which potential respondents fit into different classes in the same way that they may guess the gender and age of respondents? The initial questions of a street interview could establish characteristics of potential respondents to see whether they fit a set quota. But given the levels of non-response and ineligibility levels found by such an approach, this is not an ideal solution.

Because the elements within each quota are selected based on convenience or judgement, many sources of selection bias are potentially present. The interviewers may go to selected areas where eligible respondents are more likely to be found. Likewise, they may avoid people who look unfriendly or are not well dressed or those who live in undesirable locations. Quota sampling does not permit assessment of sampling error.[10]

Quota sampling attempts to obtain representative samples at a relatively low cost. Its advantages are the lower costs and greater convenience to the interviewers in selecting elements for each quota. Under certain conditions, quota sampling obtains results close to those for conventional probability sampling. [11]

Snowball sampling

Snowball sampling
A non-probability sampling technique in which an initial group of respondents is selected randomly. Subsequent respondents are selected based on the referrals or information provided by the initial respondents. By obtaining referrals from referrals, this process may be carried out in waves.

In **snowball sampling**, an initial group of respondents is selected, sometimes on a random basis, but more typically targeted at a few individuals who are known to possess the desired characteristics of the target population. After being interviewed, these respondents are asked to identify others who also belong to the target population of interest. Subsequent respondents are selected based on the referrals. By obtaining referrals from referrals, this process may be carried out in waves, thus leading to a snowballing effect. Even though probability sampling can be used to select the initial respondents, the final sample is a non-probability sample. The referrals will have demographic and psychographic characteristics more similar to the persons referring them than would occur by chance.[12]

The main objective of snowball sampling is to estimate characteristics that are rare in the wider population. Examples include users of particular government or social services, such as food stamps, whose names cannot be revealed; special census groups, such as widowed males under 35; and members of a scattered minority ethnic group. Another example is research in industrial buyer–seller relationships, using initial contacts to identify buyer–seller pairs and then subsequent 'snowballed' pairs. The major advantage of snowball sampling is that it substantially increases the likelihood of locating the desired characteristic in the population. It also results in relatively low sampling variance and costs.[13] Snowball sampling is illustrated by the following example.

example

Sampling horse owners

Dalgety Animal Feeds wished to question horse owners about the care and feeding of their horses. They could not locate any sampling frame that listed all horse owners, with the exception of registers of major racing stables. However, they wished to contact owners who had one or two horses as they believed this group was not well understood and held great marketing potential. Their initial approach involved locating interviewers at horse feed outlets. The interviewers ascertained basic characteristics of horse owners but more importantly they invited them along to focus groups. When the focus groups were conducted, issues of horse care

Snowball sampling – to help target respondents who do not usually display their unique characteristics so clearly.

and feeding were developed in greater detail to allow the construction of a meaningful postal questionnaire. As a rapport and trust was built up with those that attended the focus groups, names as referrals were given that allowed a sampling frame for the first wave of respondents to the subsequent postal survey. The process of referrals continued, allowing a total of four waves and a response of 800 questionnaires. ■

In this example, note the non-random selection of the initial group of respondents through focus group invitations. This procedure was more efficient than random selection, which given the absence of an appropriate sampling frame would be very cumbersome. In other cases where an appropriate sampling frame exists (appropriate in terms of identifying the desired characteristics in a number of respondents, not in terms of being exhaustive – if it were exhaustive, a snowball sample would not be needed), random selection of respondents through probability sampling techniques may be more appropriate.

Probability sampling techniques

Probability sampling techniques vary in terms of sampling efficiency. Sampling efficiency is a concept that reflects a trade-off between sampling cost and precision. Precision refers to the level of uncertainty about the characteristic being measured. Precision is inversely related to sampling errors but positively related to cost. The greater the precision, the greater the cost, and most studies require a trade-off. The researcher should strive for the most efficient sampling design, subject to the budget allocated. The efficiency of a probability sampling technique may be assessed by comparing it with that of simple random sampling.

Simple random sampling

Simple random sampling (SRS)

A probability sampling technique in which each element has a known and equal probability of selection. Every element is selected independently of every other element, and the sample is drawn by a random procedure from a sampling frame.

In **simple random sampling (SRS)**, each element in the population has a known and equal probability of selection. Furthermore, each possible sample of a given size (n) has a known and equal probability of being the sample actually selected. This implies that every element is selected independently of every other element. The sample is

367

drawn by a random procedure from a sampling frame. This method is equivalent to a lottery system in which names are placed in a container, the container is shaken and the names of the winners are then drawn out in an unbiased manner.

To draw a simple random sample, the researcher first compiles a sampling frame in which each element is assigned a unique identification number. Then random numbers are generated to determine which elements to include in the sample. The random numbers may be generated with a computer routine or a table (see Table 1 in the Appendix of Statistical Tables). Suppose that a sample of size 10 is to be selected from a sampling frame containing 800 elements. This could be done by starting with row 1 and column 1 of Table 1, considering the three rightmost digits, and going down the column until 10 numbers between 1 and 800 have been selected. Numbers outside this range are ignored. The elements corresponding to the random numbers generated constitute the sample. Thus, in our example, elements 480, 368, 130, 167, 570, 562, 301, 579, 475 and 553 would be selected. Note that the last three digits of row 6 (921) and row 11 (918) were ignored, because they were out of range. Using these tables is fine for small samples, but can be very tedious. A more pragmatic solution is to turn to random number generators in most data analysis packages. For example, in Excel, the Random Number Generation Analysis Tool allows you to set a number of characteristics of your target population, including the nature of distribution of the data, and to create a table of random numbers on a separate worksheet.

SRS has many desirable features. It is easily understood, the sample results may be projected to the target population, and most approaches to statistical inference assume that the data have been collected by simple random sampling. SRS suffers from at least four significant limitations, however. First, it is often difficult to construct a sampling frame that will permit a simple random sample to be drawn. Second, SRS can result in samples that are very large or spread over large geographic areas, thus increasing the time and cost of data collection. Third, SRS often results in lower precision with larger standard errors than other probability sampling techniques. Fourth, SRS may or may not result in a representative sample. Although samples drawn will represent the population well on average, a given simple random sample may grossly misrepresent the target population. This is more likely if the size of the sample is small. For these reasons, SRS is not widely used in marketing research. Procedures such as systematic sampling are more popular.

Systematic sampling

Systematic sampling
A probability sampling technique in which the sample is chosen by selecting a random starting point and then picking every *i*th element in succession from the sampling frame.

In **systematic sampling**, the sample is chosen by selecting a random starting point and then picking every *i*th element in succession from the sampling frame. The sampling interval, *i*, is determined by dividing the population size *N* by the sample size *n* and rounding to the nearest whole number. For example, there are 100,000 elements in the population and a sample of 1000 is desired. In this case, the sampling interval, *i*, is 100. A random number between 1 and 100 is selected. If, for example, this number is 23, the sample consists of elements 23, 123, 223, 323, 423, 523, and so on.[14]

Systematic sampling is similar to SRS in that each population element has a known and equal probability of selection. It is different from SRS, however, in that only the permissible samples of size *n* that can be drawn have a known and equal probability of selection. The remaining samples of size *n* have a zero probability of being selected.

For systematic sampling, the researcher assumes that the population elements are ordered in some respect. In some cases, the ordering (for example, alphabetical listing in a telephone book) is unrelated to the characteristic of interest. In other instances, the ordering is directly related to the characteristic under investigation. For example, credit card customers may be listed in order of outstanding balance, or firms in a given industry may be ordered according to annual sales. If the population elements

are arranged in a manner unrelated to the characteristic of interest, systematic sampling will yield results quite similar to SRS.

On the other hand, when the ordering of the elements is related to the characteristic of interest, systematic sampling increases the representativeness of the sample. If firms in an industry are arranged in increasing order of annual sales, a systematic sample will include some small and some large firms. A simple random sample may be unrepresentative because it may contain, for example, only small firms or a disproportionate number of small firms. If the ordering of the elements produces a cyclical pattern, systematic sampling may decrease the representativeness of the sample. To illustrate, consider the use of systematic sampling to generate a sample of monthly department store sales from a sampling frame containing monthly sales for the last 60 years. If a sampling interval of 12 is chosen, the resulting sample would not reflect the month-to-month variation in sales.[15]

Systematic sampling is less costly and easier than SRS because random selection is done only once to establish a starting point. Moreover, random numbers do not have to be matched with individual elements as in SRS. Because some lists contain millions of elements, considerable time can be saved, which reduces the costs of sampling. If information related to the characteristic of interest is available for the population, systematic sampling can be used to obtain a more representative and reliable (lower sampling error) sample than SRS. Another relative advantage is that systematic sampling can even be used without knowledge of the elements of the sampling frame. For example, every ith person leaving a shop or passing a point in the street can be intercepted (provided very strict control of the flow of potential respondents is exercised). For these reasons, systematic sampling is often employed in consumer mail, telephone and street interviews, as illustrated by the following example.

e x a m p l e

Tennis's systematic sampling returns a smash[16]

Tennis magazine conducted a postal survey of its subscribers to gain a better understanding of its market. Systematic sampling was employed to select a sample of 1,472 subscribers from the publication's domestic circulation list. If we assume that the subscriber list had 1,472,000 names, the sampling interval would be 1,000 (1,472,000/1,472). A number from 1 to 1,000 was drawn at random. Beginning with that number, every subsequent 1,000th was selected.

An 'alert' postcard was mailed one week before the survey. A second, follow-up, questionnaire was sent to the whole sample 10 days after the initial questionnaire. There were 76 post office returns, so the net effective mailing was 1,396. Six weeks after the first mailing, 778 completed questionnaires were returned, yielding a response rate of 56%. ■

Stratified sampling

Stratified sampling
A probability sampling technique that uses a two-step process to partition the population into subsequent sub-populations, or strata. Elements are selected from each stratum by a random procedure.

Stratified sampling is a two-step process in which the population is partitioned into sub-populations, or strata. The strata should be mutually exclusive and collectively exhaustive in that every population element should be assigned to one and only one stratum and no population elements should be omitted. Next, elements are selected from each stratum by a random procedure, usually SRS. Technically, only SRS should be employed in selecting the elements from each stratum. In practice, sometimes systematic sampling and other probability sampling procedures are employed. Stratified sampling differs from quota sampling in that the sample elements are selected probabilistically rather than based on convenience or judgement. A major objective of stratified sampling is to increase precision without increasing cost.[17]

The variables used to partition the population into strata are referred to as stratification variables. The criteria for the selection of these variables consist of homogeneity, heterogeneity, relatedness and cost. The elements within a stratum should be as homogeneous as possible, but the elements in different strata should be as

369

heterogeneous as possible. The stratification variables should also be closely related to the characteristic of interest. The more closely these criteria are met, the greater the effectiveness in controlling extraneous sampling variation. Finally, the variables should decrease the cost of the stratification process by being easy to measure and apply. Variables commonly used for stratification include demographic characteristics (as illustrated in the example for quota sampling), type of customer (e.g. credit card versus non-credit card), size of firm, or type of industry. It is possible to use more than one variable for stratification, although more than two are seldom used because of pragmatic and cost considerations. Although the number of strata to use is a matter of judgement, experience suggests the use of no more than six. Beyond six strata, any gain in precision is more than offset by the increased cost of stratification and sampling.

Another important decision involves the use of proportionate or disproportionate sampling. In proportionate stratified sampling, the size of the sample drawn from each stratum is proportionate to the relative size of that stratum in the total population. In disproportionate stratified sampling, the size of the sample from each stratum is proportionate to the relative size of that stratum and to the standard deviation of the distribution of the characteristic of interest among all the elements in that stratum. The logic behind disproportionate sampling is simple. First, strata with larger relative sizes are more influential in determining the population mean, and these strata should also exert a greater influence in deriving the sample estimates. Consequently, more elements should be drawn from strata of larger relative size. Second, to increase precision, more elements should be drawn from strata with larger standard deviations and fewer elements should be drawn from strata with smaller standard deviations. (If all the elements in a stratum are identical, a sample size of one will result in perfect information.) Note that the two methods are identical if the characteristic of interest has the same standard deviation within each stratum.

Disproportionate sampling requires that some estimate of the relative variation, or standard deviation of the distribution of the characteristic of interest, within strata be known. As this information is not always available, the researcher may have to rely on intuition and logic to determine sample sizes for each stratum. For example, large retail stores might be expected to have greater variation in the sales of some products as compared with small stores. Hence, the number of large stores in a sample may be disproportionately large. When the researcher is primarily interested in examining differences between strata, a common sampling strategy is to select the same sample size from each stratum.

Stratified sampling can ensure that all the important sub-populations are represented in the sample. This is particularly important if the distribution of the characteristic of interest in the population is skewed. For example, very few households have annual incomes that allow them to own a second home overseas. If a simple random sample is taken, households that have a second home overseas may not be adequately represented. Stratified sampling would guarantee that the sample contains a certain number of these households. Stratified sampling combines the simplicity of SRS with potential gains in precision. Therefore, it is a popular sampling technique.

Cluster sampling

In **cluster sampling**, the target population is first divided into mutually exclusive and collectively exhaustive sub-populations. These sub-populations or clusters are assumed to contain the diversity of respondents held in the target population. A random sample of clusters is selected, based on a probability sampling technique such as SRS. For each selected cluster, either all the elements are included in the sample or a sample of elements is drawn probabilistically. If all the elements in each selected cluster are included in the sample, the procedure is called one-stage cluster sampling. If a sample of elements is drawn probabilistically from each selected cluster, the pro-

Cluster sampling
A two-step probability sampling technique where the target population is first divided into mutually exclusive and collectively exhaustive sub-populations called clusters, and then a random sample of clusters is selected based on a probability sampling technique such as simple random sampling. For each selected cluster, either all the elements are included in the sample, or a sample of elements is drawn probabilistically.

cedure is two-stage cluster sampling. As shown in Figure 14.3, two-stage cluster sampling can be either simple two-stage cluster sampling involving SRS or probability proportionate to size (PPS) sampling. Furthermore, a cluster sample can have multiple (more than two) stages, as in multistage cluster sampling.

The key distinction between cluster sampling and stratified sampling is that in cluster sampling only a sample of sub-populations (clusters) is chosen, whereas in stratified sampling all the sub-populations (strata) are selected for further sampling. The objectives of the two methods are also different. The objective of cluster sampling is to increase sampling efficiency by decreasing costs, but the objective of stratified sampling is to increase precision. With respect to homogeneity and heterogeneity, the criteria for forming clusters are just the opposite of those for strata. Elements within a cluster should be as heterogeneous as possible, but clusters themselves should be as homogeneous as possible. Ideally, each cluster should be a small-scale representation of the population. In cluster sampling, a sampling frame is needed only for those clusters selected for the sample.

Area sampling
A common form of cluster sampling in which the clusters consist of geographic areas such as counties, housing tracts, blocks or other area descriptions.

A common form of cluster sampling is **area sampling**, in which the clusters consist of geographic areas, such as counties, housing districts or residential blocks. If only one level of sampling takes place in selecting the basic elements (for example, if the researcher samples blocks and then all the households within the selected blocks are included in the sample), the design is called one-stage area sampling. If two or more levels of sampling take place before the basic elements are selected (if the researcher samples blocks and then samples households within the sampled blocks), the design is called two-stage (or multistage) area sampling. The distinguishing feature of one-stage area sampling is that all the households in the selected blocks (or geographic areas) are included in the sample.

There are two types of two-stage cluster sampling designs, as shown in Figure 14.3. Simple two-stage cluster sampling involves SRS at the first stage (e.g. sampling blocks) as well as the second stage (e.g. sampling households within blocks). In this design the number of elements (e.g. households) selected at the second stage is the same for each sample cluster (e.g. selected blocks). This design is appropriate when the clusters are equal in size, that is, when the clusters contain approximately the same number of sampling units. If they differ greatly in size, however, simple two-stage cluster sampling can lead to biased estimates. Sometimes the clusters can be made of equal size by combining clusters. When this option is not feasible, probability proportionate to size (PPS) sampling can be used.

Probability proportionate to size (PPS)
A selection method where the probability of selecting a sampling unit in a selected cluster varies inversely with the size of the cluster. Therefore, the size of all the resulting clusters is approximately equal.

In **probability proportionate to size (PPS)** sampling, the clusters are sampled with probability proportional to size. The size of a cluster is defined in terms of the number of sampling units within that cluster. Thus, in the first stage, large clusters are more likely to be included than small clusters. In the second stage, the probability of selecting a sampling unit in a selected cluster varies inversely with the size of the cluster. Thus, the probability that any particular sampling unit will be included in the
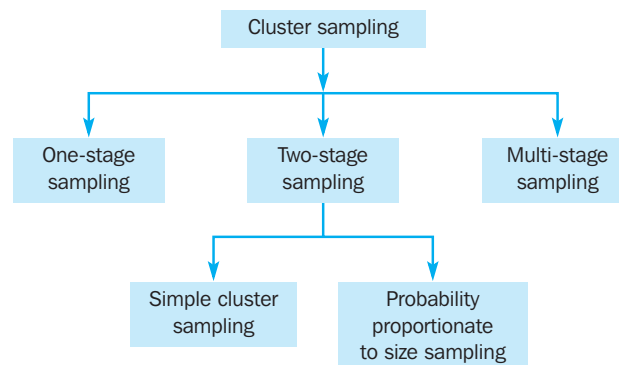
Figure 14.3 Types of cluster sampling

sample is equal for all units, because the unequal first stage probabilities are balanced by the unequal second stage probabilities. The numbers of sampling units included from the selected clusters are approximately equal.

Cluster sampling has two major advantages: feasibility and low cost. These advantages are illustrated in the following example where alternative means of drawing a sample were severely restricted without incurring great costs.

example

Sport and leisure demands in schools

The Sports Council has supported marketing research in the South West of England that has facilitated new facilities and sports development in a range of cities. The methodology involved using a postal survey which was sent to individuals in households. The electoral register acted as the sampling frame and a systematic sampling method was used. A major problem lay in sampling younger members of the community who were not named on the electoral register and had no known sampling frame that marketing researchers could access.

The solution lay in the use of a simple two-stage cluster sampling design. All schools and colleges within a target district could be identified; a complete and current sampling frame of schools and colleges was available. A simple random sample of all the schools and colleges constituted the first stage. Then a random sample of classes within the schools and colleges was taken (between the ages of 12 and 18). In this design, the number of elements (classes) selected at the second stage was the same for each sample cluster (school or college).

With the selected classes, permission was gained to administer the questionnaire to the whole class. This resulted in a very cost-effective means of data collection. As all the class completed the task together, a consistent means to motivate respondents and instructions could be given. ■

In many situations the only sampling frames readily available for the target population are clusters, not population elements. In the above example, the schools and their classes are known but not the pupils. It is often impossible to compile a list of all consumers in a population, given the resources and constraints. Lists of geographical areas, telephone exchanges and other clusters of consumers, however, can be constructed relatively easily. Cluster sampling is the most cost-effective probability sampling technique. This advantage must be weighed against several limitations. Cluster sampling results in relatively imprecise samples, and it is difficult to form clusters in which the elements are heterogeneous, because, for example, households in a block tend to be similar rather than dissimilar.[18] It can be difficult to compute and interpret statistics based on clusters.

Other probability sampling techniques

In addition to the four basic probability sampling techniques, there are a variety of other sampling techniques. Most of these may be viewed as extensions of the basic techniques and were developed to address complex sampling problems. Two techniques with some relevance to marketing research are sequential sampling and double sampling.

Sequential sampling
A probability sampling technique in which the population elements are sampled sequentially, data collection and analysis are done at each stage, and a decision is made as to whether additional population elements should be sampled.

In **sequential sampling**, the population elements are sampled sequentially, data collection and analysis are done at each stage, and a decision is made as to whether additional population elements should be sampled. The sample size is not known in advance, but a decision rule is stated before sampling begins. At each stage, this rule indicates whether sampling should be continued or whether enough information has been obtained. Sequential sampling has been used to determine preferences for two competing alternatives. In one study, respondents were asked which of two alternatives they preferred, and sampling was terminated when sufficient evidence was accumulated to validate a preference. It has also been used to establish the price differential between a standard model and a deluxe model of a consumer durable.[19]

Double sampling
A sampling technique in which certain population elements are sampled twice.

In **double sampling**, also called two-phase sampling, certain population elements are sampled twice. In the first phase, a sample is selected and some information is col-

lected from all the elements in the sample. In the second phase, a sub-sample is drawn from the original sample and additional information is obtained from the elements in the sub-sample. The process may be extended to three or more phases, and the different phases may take place simultaneously or at different times. Double sampling can be useful when no sampling frame is readily available for selecting final sampling units but when the elements of the frame are known to be contained within a broader sampling frame. For example, a researcher wants to select households in a given city that consume apple juice. The households of interest are contained within the set of all households, but the researcher does not know which ones they are. In applying double sampling, the researcher would obtain a sampling frame of all households in the first phase. This would be constructed from the city directory or purchased. Then a sample of households would be drawn, using systematic random sampling to determine the amount of apple juice consumed. In the second phase, households that consume apple juice would be selected and stratified according to the amount of apple juice consumed. Then a stratified random sample would be drawn and detailed questions regarding apple juice consumption asked.[20]

Choosing non-probability versus probability sampling

The choice between non-probability and probability samples should be based on considerations such as the nature of the research, relative magnitude of non-sampling versus sampling errors, and variability in the population, as well as statistical and operational considerations (see Table 14.3).

Table 14.3 Choosing non-probability vs. probability sampling

Factors	Conditions favouring the use of	
	Non-probability sampling	Probability sampling
Nature of research	Exploratory	Conclusive
Relative magnitude of sampling and non-sampling errors	Non-sampling errors are larger	Sampling errors are larger
Variability in the population	Homogeneous (low)	Heterogeneous (high)
Statistical considerations	Unfavourable	Favourable
Operational considerations	Favourable	Unfavourable

For example, in exploratory research, the judgement of the researchers in selecting respondents with particular qualities may be far more effective than any form of probability sampling. On the other hand, in conclusive research where the researcher wishes to use the results to estimate overall market shares or the size of the total market, probability sampling is favoured. Probability samples allow statistical projection of the results to a target population.

For some research problems, highly accurate estimates of population characteristics are required. In these situations, the elimination of selection bias and the ability to calculate sampling error make probability sampling desirable. Probability sampling will not always result in more accurate results, however. If non-sampling errors are likely to be an important factor, then non-probability sampling may be preferable because the use of judgement may allow greater control over the sampling process.

Another consideration is the homogeneity of the population with respect to the variables of interest. A heterogeneous population would favour probability sampling because it would be more important to secure a representative sample. Probability sampling is preferable from a statistical viewpoint, as it is the basis of most common statistical techniques.

Probability sampling generally requires statistically trained researchers, generally costs more and takes longer than non-probability sampling, especially in the establishment of accurate sampling frames. In many marketing research projects, it is difficult to justify the additional time and expense. Therefore, in practice, the objectives of the study dictate which sampling method will be used.

Non-probability sampling is used in concept tests, package tests, name tests and copy tests where projections to the populations are usually not needed. In such studies, interest centres on the proportion of the sample that gives various responses or expresses various attitudes. Samples for these studies can be drawn using methods such as street interviewing and quota sampling. On the other hand, probability sampling is used when there is a need for highly accurate estimates of market share or sales volume for the entire market. National market tracking studies, which provide information on product category and brand usage rates as well as psychographic and demographic profiles of users, use probability sampling. Studies that use probability sampling generally employ telephone interviews. Stratified and systematic sampling are combined with some form of random-digit dialling to select the respondents.

Summary of sampling techniques

The strengths and weaknesses of cluster sampling and the other basic sampling techniques are summarised in Table 14.4. Table 14.5 describes the procedures for drawing probability samples.

Table 14.4 Strengths and weaknesses of sampling techniques

Technique	Strengths	Weaknesses
Non-probability sampling		
Convenience sampling	Least expensive, least time consuming, most convenient	Selection bias, sample not representative, not recommended for descriptive or causal research
Judgemental sampling	Low cost, convenient, not time consuming.Ideal for exploratory research designs	Does not allow generalisation, subjective
Quota sampling	Sample can be controlled for certain characteristics	Selection bias, no assurance of representativeness
Snowball sampling	Can estimate rare characteristics	Time consuming
Probability sampling		
Simple random sampling (SRS)	Easily understood, results projectable	Difficult to construct sampling frame, expensive, lower precision, no assurance of representativeness
Systematic sampling	Can increase representativeness, easier to implement than SRS, sampling frame not always necessary	Can decrease representativeness
Stratified sampling	Includes all important sub-populations, precision	Difficult to select relevant stratification variables, not feasible to stratify on many variables, expensive
Cluster sampling	Easy to implement, cost-effective	Imprecise, difficult to compute and interpret results

Table 14.5 Procedures for drawing probability samples

Simple random sampling

1   Select a suitable sampling frame.

2   Each element is assigned a number from 1 to N (population size).

3   Generate n (sample size) different random numbers between 1 and N using a software package or a table of simple random numbers (Table 1 in the Appendix of Statistical Tables). To use Table 1, select the appropriate number of digits (e.g. if $N = 900$, select three digits). Arbitrarily select a beginning number. Then proceed up or down until n different numbers between 1 and N have been selected. Discard 0, duplicate numbers, and numbers greater than N.

4   The numbers generated denote the elements that should be included in the sample.

Systematic sampling

1   Select a suitable sampling frame.

2   Each element is assigned a number from 1 to N (population size).

3   Determine the sampling interval i, where $i = N/n$. If i is a fraction, round to the nearest whole number.

4   Select a random number, r, between 1 and i, as explained in simple random sampling.

5   The elements with the following numbers will comprise the systematic random sample:

$$r, r + i, r + 2i, r + 3i, r + 4i \ldots r + (n-1)i$$

Stratified sampling

1   Select a suitable sampling frame.

2   Select the stratification variable(s) and the number of strata, H.

3   Divide the entire population into H strata. Based on the classification variable, each element of the population is assigned to one of the H strata.

4   In each stratum, number the elements from 1 to N_h (the population size of stratum h).

5   Determine the sample size of each stratum, n_h, based on proportionate or disproportionate stratified sampling, where

$$\sum_{h=1}^{H} n_h = n$$

6   In each stratum, select a simple random sample of size n_h.

Cluster sampling

We describe the procedure for selecting a two-stage PPS sample, because this represents the most commonly used general case.

1   Assign a number from 1 to N to each element in the population.

2   Divide the population into C clusters of which c will be included in the sample.

3   Calculate the sampling interval i, where $i = N/c$. If i is a fraction, round to the nearest whole number.

4   Select a random number, r, between 1 and i, as explained in simple random sampling.

5   Identify elements with the following numbers: $r, r + i, r + 2i, r + 3i \ldots, r + (c-1)i$.

6   Select the clusters that contain the identified elements.

7   Select sampling units within each selected cluster based on SRS or systematic sampling. The number of sampling units selected from each sample cluster is approximately the same and equal to n/c.

8   If the population of the cluster exceeds the sampling interval i, that cluster is selected with certainty. That cluster is removed from further consideration. Calculate the new proportion size, N^*, the number of clusters to be selected, $c^* (= c - 1)$, and the new sampling interval i^*. Repeat this process until each of the remaining clusters has a population less than the relevant sampling interval. If b clusters have been selected with certainty, select the remaining $c - b$ clusters according to steps 1 to 7. The fraction of units to be sampled from each cluster selected with certainty is the overall sampling fraction n/N. Thus, for clusters selected with certainty, we would select $n_s = (n/N)(N_1 + N_2 + \ldots + N_b)$ units. The units selected from clusters selected under PPS sampling will therefore be $n^* = n - n_s$.

International marketing research

Implementing the sampling design process in international marketing research is seldom easy. Several factors should be considered in defining the target population. The relevant element (respondent) may differ from country to country. In Europe, children play an important role in the purchase of children's cereals and may be seen as target respondents. In countries with authoritarian child-rearing practices, however, the mother or father may be the relevant target respondents. Accessibility also varies across countries. In Mexico, 'upper class' houses cannot be entered by strangers because of boundary walls and servants. Additionally, dwelling units may be unnumbered and streets unidentified, making it difficult to locate designated households.[21]

Developing an appropriate sampling frame is a difficult task. In many countries, particularly developing countries, reliable information about the target population may not be available from secondary sources. Government data may be unavailable or highly biased. Population lists may not be available commercially. The time and money required to compile these lists may be prohibitive. For example, in Saudi Arabia, there is no officially recognised census of population, no elections and hence no voter registration records, and no accurate maps of population centres. In this situation, the interviewers could be instructed to begin at specified starting points and to sample every nth dwelling until the specified number of units has been sampled.

Given the lack of suitable sampling frames, the inaccessibility of certain respondents, such as women in some cultures, and the dominance of personal interviewing, probability sampling techniques are uncommon in international marketing research. Imagine the problems involved in tracking down an accurate sampling frame in the following example.

e x a m p l e

Post-Deng China with a new 'middle class' of 35 million households[22]

China has been transformed from a centralised state system offering only two imported items (cigarettes and soft drinks) into a socialist market economy where consumers can buy Rolex watches, Burberry raincoats, Cadbury's chocolate, Kentucky Fried Chicken, Colgate toothpaste and Nike sports shoes and other international brands. With a population of 1.2 billion it is not surprising that companies are keen to enter China, where even niche markets can be huge.

While 'middle class' is a Western concept and as such does not exist in China, there are 'Xiao Kang' or Little Rich households. Xiao Kang is a state of society in Confucian ideology where people live and work happily, which in today's context means that the people eat well, dress smartly and live in nicely furnished homes equipped with consumer durables. There are estimated to be in the region of 35 million Xiao Kang families in China, a large and lucrative segment of population that totals 1.2 billion. ■

Quota sampling has been used widely in the developed and developing countries in both consumer and industrial surveys. Quota sampling has a long history of working well in Britain, France and Germany, but is a sampling method that is seen as 'unthinkable' for many US marketing researchers.[23] Snowball sampling is also appealing when the characteristic of interest is rare in the target population or when respondents are hard to reach. For example, it has been suggested that in Saudi Arabia graduate students be employed to hand-deliver questionnaires to relatives and friends.[24] These initial respondents can be asked for referrals to other potential respondents and so on. This approach would result in a large sample size and a high response rate.

Sampling techniques and procedures vary in accuracy, reliability and cost from country to country. If the same sampling procedures are used in each country, the results may not be comparable.[25] To achieve comparability in sample composition and representativeness, it may be desirable to use different sampling techniques in different countries.

Ethics in marketing research

The researcher has several ethical responsibilities to both the client and the respondents pertaining to sampling. With regard to the client, the researcher must develop a sampling design that best fits the project in an effort to minimise the sampling and non-sampling errors (see Chapter 3). When probability sampling can be used it should be.

When non-probability design such as convenience sampling is used, the limitations of the design should be explicit in any findings that are presented. It is unethical and misleading to treat non-probability samples as probability samples and to project the results to a target population. Appropriate definition of the population and the sampling frame, and application of the correct sampling techniques, are essential if the research is to be conducted and the findings used ethically.

Researchers must be extremely sensitive to preserving the anonymity of the respondents when conducting business-to-business research with small populations, particularly when reporting the findings to the client. When the population size is small, it is easier to discern the identities of the respondents than when the samples are drawn from a large population. Special care must be taken when sample details are too revealing and when using verbatim quotations in reports to the client. This problem is acute in areas such as employee research. Here a breach of a respondent's anonymity can cost the respondent a pay rise, a promotion, or even their employment. In such situations, special effort should be made to protect the identities of the respondents.

Internet and computer applications

Sampling potential respondents who are surfing the Internet is meaningful if the sample generated is representative of the target population. More and more industries are meeting this criterion. In software, computers, networking, technical publishing, semiconductors and graduate education, it is rapidly becoming feasible to use the Internet for sampling respondents for quantitative research, such as surveys. For internal customer surveys, where the client's employees share a corporate email system, an intranet survey is practical, even if workers have no access to the external Internet. Look at Professional Perspective 5 written by Ron Whelan on the Companion Website to see these issues discussed in more detail.

To avoid sampling errors, the researcher must be able to control the pool from which the respondents are selected. Also, it must be ensured that the respondents do not respond more than once. These requirements are met by email surveys, in which the researcher selects specific respondents. Furthermore, the surveys can be encoded to match the returned surveys with their corresponding outbound emailings. This can also be accomplished with Web surveys by emailing invitations to selected respondents and asking them to visit the Website on which the survey is posted. In this case, the survey is posted in a hidden location on the Web, which is protected by a password. Hence, non-invited Web surfers are unable to access it.

Non-probability as well as probability sampling techniques can be implemented on the Internet. Moreover, the respondents can be pre-recruited or tapped online. Tapping visitors to a Website is an example of convenience sampling. Based on the researcher's judgement, certain qualifying criteria can be introduced to pre-screen the respondents. Even quotas can be imposed. However, the extent to which quotas will be met is limited by the number as well as the characteristics of visitors to the site.

Likewise, simple random sampling is commonly used. To prevent gathering information from the same professional respondents (professional in this context meaning respondents who take part in many surveys for their own enjoyment), some companies use a 'click-stream intercept', which randomly samples online users and gives them the opportunity to participate or decline.

Microcomputers and mainframes can make the sampling design process more effective and efficient. Random number generators are available in most data analysis packages. For example, in Excel, the Random Number Generation Analysis Tool allows you to set a number of characteristics of your target population, including the nature of distribution of the data, and to create a table of random numbers on a separate worksheet. Computers can also be used in the specification of the sampling frame. Geodemographic information systems such as Experian (www.experian.com) handle lists of population elements as well as geographical maps. Database packages can also be used to store and manipulate sampling frames, especially when the sampling frame is built up from multiple sources and duplicates need to be identified and eliminated. Once the sampling frame has been determined, simulations can be used to generate random numbers and select the sample directly from the database.

➤➤➤
See Professional Perspectives 5, 20.

Go to the Companion Website and read Professional Perspective 5 from Tim Macer. In 'Playing the Internet number game', Tim reviews an online sample ordering system, SSI-SNAP (www.surveysampling.com). See also Professional Perspective 20 'Online methodological meditations' by Ron Whelan. He tackles the issues of how representative Internet samples are, and the concerns about the self-completion nature of Internet interviews.

To experience how you can precisely define a target population and create a distinctive sampling frame, based upon an array of demographic and lifestyle characteristics, look at www.prospectlocator.com. Another good reference to evaluate sampling techniques can be found on http://trochim.human.cornell.edu/kb/sampprob.htm.

Summary

Information about the characteristics of a population may be obtained by conducting either a sample or a census. Budget and time limits, large population size, and small variance in the characteristic of interest favour the use of a sample. Sampling is also preferred when the cost of sampling error is low, the cost of non-sampling error is high, the nature of measurement is destructive, and attention must be focused on individual cases. The opposite set of conditions favours the use of a census.

Sampling design begins by defining the target population in terms of elements, sampling units, extent and time. Then the sampling frame should be determined. A sampling frame is a representation of the elements of the target population. It consists of a list of directions for identifying the target population. At this stage, it is important to recognise any sampling frame errors that may exist. The next step involves selecting a sampling technique and determining the sample size. In addition to quantitative analysis, several qualitative considerations should be taken into account in determining the sample size. Execution of the sampling process requires detailed specifications for each step in the sampling process. Finally, the selected sample should be validated by comparing characteristics of the sample with known characteristics of the target population.

Sampling techniques may be classified as non-probability and probability techniques. Non-probability sampling techniques rely on the researcher's judgement.

Consequently, they do not permit an objective evaluation of the precision of the sample results, and the estimates obtained are not statistically projectable to the population. The commonly used non-probability sampling techniques include convenience sampling, judgemental sampling, quota sampling and snowball sampling.

In probability sampling techniques, sampling units are selected by chance. Each sampling unit has a non-zero chance of being selected, and the researcher can pre-specify every potential sample of a given size that could be drawn from the population as well as the probability of selecting each sample. It is also possible to determine the precision of the sample estimates and inferences and make projections to the target population. Probability sampling techniques include simple random sampling, systematic sampling, stratified sampling, cluster sampling, sequential sampling and double sampling. The choice between probability and non-probability sampling should be based on the nature of the research, degree of error tolerance, relative magnitude of sampling and non-sampling errors, variability in the population, and statistical and operational considerations.

When conducting international marketing research, it is desirable to achieve comparability in sample composition and representativeness even though this may require the use of different sampling techniques in different countries. It is unethical and misleading to treat non-probability samples as probability samples and to project the results to a target population.

Questions

1 Under what conditions would a sample be preferable to a census? A census preferable to a sample?

2 Describe the sampling design process.

3 How should the target population be defined? How does this definition link with the definition of a marketing research problem?

4 What is a sampling unit? How is it different from the population element?

5 To what extent may the availability of sampling frames determine the definition of a population?

6 What qualitative factors should be considered in determining the sample size?

7 How do probability sampling techniques differ from non-probability sampling techniques? What factors should be considered in choosing between probability and non-probability sampling?

8 What is the least expensive and least time-consuming of all sampling techniques? What are the major limitations of this technique?

9 What is the major difference between judgemental and convenience sampling? Give examples of where each of these techniques may be successfully applied.

10 Describe snowball sampling. How may the technique be supported by qualitative research techniques?

11 What are the distinguishing features of simple random sampling?

12 Describe the procedure for selecting a systematic random sample.

13 Describe stratified sampling. What are the criteria for the selection of stratification variables?

14 What are the differences between proportionate and disproportionate stratified sampling?

15 Describe the cluster sampling procedure. What is the key distinction between cluster sampling and stratified sampling?

Notes

1 Easton, S. and Mackie, P., 'Can football give brands 110% recall?', *ResearchPlus* (June 1997), 10.

2 Verma, V. and Le, T., 'An analysis of sampling errors for the demographic and health surveys', *International Statistical Review* 64(3) (December 1966), 265–94; Assael, H. and Keon, J., 'Non-sampling vs. sampling errors in sampling research', *Journal of Marketing* (Spring 1982), 114–23.

3 Fink, A., *How to Sample in Surveys* (Thousand Oaks, CA: Sage, 1995); Frankel, M.R., 'Sampling theory', in Rossi, P.H., Wright, J.D. and Anderson, A.B. (eds), *Handbook of Survey Research* (Orlando, FL: Academic Press, 1983), 21–67; Jaeger, R.M., *Sampling in Education and the Social Sciences* (New York: Longman, 1984) 28–9; Kalron, G., *Introduction to Survey Sampling* (Beverly Hills, CA: Sage, 1982).

4 Henry, G.T., *Practical Sampling* (Thousand Oaks, CA: Sage, 1995); Sudman, S., 'Applied sampling', in Rossi, P.H., Wright, J.D. and Anderson, A.B. (eds), *Handbook of Survey Research* (Orlando, FL: Academic Press, 1983), 145–94.

5 Cage, R., 'New methodology for selecting CPI outlet samples', *Monthly Labor Review* 119(12) (December 1996), 49–83.

6 Smith, W., Mitchell, P., Attebo, K. and Leeder, S., 'Selection bias from sampling frames: telephone directory and electoral roll compared with door-to-door population census: results from the Blue Mountain eye study', *Australian and New Zealand Journal of Public Health* 21(2) (April 1997), 127–33.

7 For the effect of sample frame error on research results, see Fish, K.E., Barnes, J.H. and Banahan III, B.F., 'Convenience or calamity', *Journal of Health Care Marketing* 14 (Spring 1994), 45–9.

8 Phillips, A. 'Taking the people's temperature – right across Europe', *ResearchPlus* (November 1996), 6.

9 For an application of convenience sampling, see Ho, F., Ong, B.S. and Seonsu, A., 'A multicultural comparison of shopping patterns among Asian consumers', *Journal of Marketing Theory and Practice* 5(1) (Winter 1997), 42–51.

10 Curtice, J. and Sparrow, N. 'How accurate are traditional quota opinion polls', *Journal of the Market Research Society* 39(3) (July 1997), 433–48.

11 'Public opinion: polls apart', *Economist* 336(7927) (12 August 1995), 48; Kalton, G., *Introduction to Survey Sampling* (Beverly Hills, CA: Sage, 1982); Sudman, S., 'Improving the quality of shopping center sampling', *Journal of Marketing Research* 17 (November 1980), 423–31.

12 For an application of snowball sampling, see Frankwick, G.L., Ward, J.C., Hutt, M.D. and Reingen, P.H., 'Evolving patterns of organisational beliefs in the formation of strategy', *Journal of Marketing* 58 (April 1994), 96–110.

13 If certain procedures for listing members of the rare population are followed strictly, the snowball sample can be treated as a probability sample. See Kalton, G. and Anderson, D.W., 'Sampling rare populations', *Journal of the Royal Statistical Association* (1986), 65–82; Biemacki, P. and Waldorf, D., 'Snowball sampling: problems and techniques of chain referred sampling', *Sociological Methods and Research* 10 (November 1981), 141–63; Rothbart, G.S., Fine, M. and Sudman, S., 'On finding and interviewing the needles in the haystack: the use of multiplicity sampling', *Public Opinion Quarterly* 46 (Fall 1982), 408–21.

14 When the sampling interval, *i*, is not a whole number, the easiest solution is to use as the interval the nearest whole number below or above *i*. If rounding has too great an effect on the sample size, add or delete the extra cases.

15 For an application of systematic random sampling, see Qu, H. and Li, I., 'The characteristics and satisfaction of mainland Chinese visitors to Hong Kong', *Journal of Travel Research* 35(4) (Spring 1997), 37–41; Chakraborty, G., Ettenson, R. and Gaeth, G., 'How consumers choose health insurance', *Journal of Health Care Marketing* 14 (Spring 1994), 21–33.

16 Adams, M., 'Court Marshall', *Mediaweek* 6(12) (18 March 1996), 22; 'Readership survey serves *Tennis* magazine's marketing needs', *Quirk's Marketing Research Review* (May 1988), 75–6.

17 For an application of stratified random sampling, see Weerahandi, S. and Moitra, S., 'Using survey data to predict adoption and switching for services', *Journal of Marketing Research* 32 (February 1995), 85–96.

18 Geographic clustering of rare populations, however, can be an advantage. See Raymondo, J.C., 'Confessions of a Nielsen Housechild', *American Demographics* 19(3) (March 1997), 24–7; Sudman, S., 'Efficient screening methods for the sampling of geographically clustered special populations', *Journal of Marketing Research* 22 (February 1985), 20–9.

19 Park, J.S., Peters, M. and Tang, K., 'Optimal inspection policy in sequential screening', *Management Science* 37(8) (August 1991), 1058–61; Anderson, E.J., Gorton, K. and Tudor, R., 'The application of sequential analysis in market research', *Journal of Marketing Research* 17 (February 1980), 97–105.

20 For more discussion of double sampling, see Baillie, D.H., 'Double sampling plans for inspection by variables when the process standard deviation is unknown', *International Journal of Quality & Reliability Management* 9(5) (1992), 59–70; Frankel, M.R. and Frankel, L.R., 'Probability sampling', in Ferber, R. (ed.), *Handbook of Marketing Research* (New York: McGraw-Hill, 1974), 2-230–2-246.

21 Murphy, S., 'Moving targets', *Business Latin America* 31(13) (1 April 1996), 4–5. For the use of different non-probability and probability sampling techniques in cross-cultural research, see Saeed, S. and Jeong, I., 'Cross-cultural research in advertising: an assessment of methodologies', *Journal of the Academy of Marketing Science* 22 (Summer 1994), 205–15.

22 Hutton, G., 'The land where the little rich have big three aspirations', *ResearchPlus* (March 1997), 12.

23 Taylor, H. 'Horses for courses: how survey firms in different countries measure public opinion with very different methods', *Journal of the Market Research Society* 37(3) (July 1995), 218.

24 Tuncalp, S., 'The marketing research scene in Saudi Arabia', *European Journal of Marketing* 22(5) (1988), 15–22.

25 Grosh, M.E. and Glewwe, P., 'Household survey data from developing countries: progress and prospects', *American Economic Review* 86(2) (May 1996), 15–19.

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Sampling: final and initial sample size determination

Objectives

After reading this chapter, you should be able to:

1 define key concepts and symbols pertinent to sampling;

2 understand the concepts of the sampling distribution, statistical inference and standard error;

3 discuss the statistical approach to determining sample size based on simple random sampling and the construction of confidence intervals;

4 derive the formulas to determine statistically the sample size for estimating means and proportions;

5 discuss the non-response issues in sampling and the procedures for improving response rates and adjusting for non-response;

6 understand the difficulty of statistically determining the sample size in international marketing research;

7 identify the ethical issues related to sample size determination, particularly the estimation of population variance.

Making a sample too big wastes resources, making it too small diminishes the value of findings – a dilemma resolved only with the judicious use of sampling theory.

Overview

This chapter focuses on the determination of sample size in simple random sampling. We define various concepts and symbols and discuss the properties of the sampling distribution. Additionally, we describe statistical approaches to sample size determination based on confidence intervals. We present the formulas for calculating the sample size with these approaches and illustrate their use. We briefly discuss the extension to determining sample size in other probability sampling designs. The sample size determined statistically is the final or net sample size; that is, it represents the completed number of interviews or observations. To obtain this final sample size, however, a much larger number of potential respondents have to be contacted initially. We describe the adjustments that need to be made to the statistically determined sample size to account for incidence and completion rates and calculate the initial sample size. We also cover the non-response issues in sampling, with a focus on improving response rates and adjusting for non-response. We discuss the difficulty of statistically determining the sample size in international marketing research and identify the relevant ethical issues.

Statistical determination of sample size requires knowledge of the normal distribution and the use of normal probability tables. The normal distribution is bell-shaped and symmetrical. Its mean, median and mode are identical (see Chapter 18). Information on the normal distribution and the use of normal probability tables is presented in Appendix 15A. The following example illustrates the statistical aspects of sampling.

_e x a m p l e_   **Has there been a shift in opinion?**

The sample size used in opinion polls commissioned and published by most national newspapers is influenced by statistical considerations. The allowance for sampling error may be limited to around three percentage points.

The table that follows can be used to determine the allowances that should be made for sampling error. These intervals indicate the range (plus or minus the figure shown) within which the results of repeated samplings in the same time period could be expected to vary, 95% of the time, assuming that the sample procedure, survey execution and questionnaire used were the same.

Recommended allowance for sampling error of a percentage

In percentage points (at 95% confidence level for a sample size of 385)	
Percentage near 10	3
Percentage near 20	4
Percentage near 30	4
Percentage near 40	5
Percentage near 50	5
Percentage near 60	5
Percentage near 70	4
Percentage near 80	4
Percentage near 90	3

The table should be used as follows. If a reported percentage is 43 (e.g. 43% of Norwegian Chief Executives believe their company will suffer from staff shortages in the next 12 months), look at the row labelled 'percentages near 40'. The number in this row is 5, so the 43% obtained in the sample is subject to a sampling error of ±5 percentage points. Another way of saying this is that very probably (95 times out of 100) the average of repeated samplings would be somewhere between 38% and 48%. The reader can be 95% confident

that in the total population of Norwegian Chief Executives between 38% and 48% believe their company will suffer from staff shortages in the next 12 months, with the most likely figure being 43%.

The fortunes of political parties measured through opinion polls are regularly reported in newspapers throughout Europe. The next time that you read a report of a political opinion poll, examine the sample size used, the confidence level assumed and the stated margin of error. When comparing the results of a poll with a previous poll, consider whether a particular political party or politician has *really* grown or slumped in popularity, or the reported changes can be accounted for within the set margin of error as summarised in this example. ■

Definitions and symbols

Confidence intervals and other statistical concepts that play a central role in sample size determination are defined in the following list.

■ *Parameter*. A parameter is a summary description of a fixed characteristic or measure of the target population. A parameter denotes the true value that would be obtained if a census rather than a sample was undertaken.
■ *Statistic*. A statistic is a summary description of a characteristic or measure of the sample. The sample statistic is used as an estimate of the population parameter.
■ *Finite population correction*. The finite population correction (fpc) is a correction for overestimation of the variance of a population parameter – for example, a mean or proportion – when the sample size is 10% or more of the population size.
■ *Precision level*. When estimating a population parameter by using a sample statistic, the precision level is the desired size of the estimating interval. This is the maximum permissible difference between the sample statistic and the population parameter.
■ *Confidence interval*. The confidence interval is the range into which the true population parameter will fall, assuming a given level of confidence.
■ *Confidence level*. The confidence level is the probability that a confidence interval will include the population parameter.

The symbols used in statistical notation for describing population and sample characteristics are summarised in Table 15.1.

Table 15.1 Symbols for population and sample variables

Variable	Population	Sample
Mean	μ	$\bar{X}$
Proportion	π	p
Variance	σ^2	s^2
Standard deviation	σ	s
Size	N	n
Standard error of the mean	$\sigma_{\bar{x}}$	$S_{\bar{x}}$
Standard error of the proportion	σ_p	S_p
Standardised variate (z)	$\dfrac{X - \mu}{\sigma}$	$\dfrac{X - \bar{X}}{S}$
Coefficient of variation (C)	$\dfrac{\sigma}{\mu}$	$\dfrac{S}{X}$

The sampling distribution

Sampling distribution
The distribution of the values of a sample statistic computed for each possible sample that could be drawn from the target population under a specified sampling plan.

The **sampling distribution** is the distribution of the values of a sample statistic computed for each possible sample that could be drawn from the target population under a specified sampling plan.[1] Suppose that a simple random sample of five hospitals is to be drawn from a population of 20 hospitals. There are $(20 \times 19 \times 18 \times 17 \times 16)/(1 \times 2 \times 3 \times 4 \times 5)$, or 15,504 different samples of size 5 that can be drawn. The relative frequency distribution of the values of the mean of these 15,504 different samples would specify the sampling distribution of the mean.

An important task in marketing research is to calculate statistics, such as the sample mean and sample proportion, and use them to estimate the corresponding true population values. This process of generalising the sample results to a target population is referred to as **statistical inference**. In practice, a single sample of predetermined size is selected, and the sample statistics (such as mean and proportion) are computed. Theoretically, to estimate the population parameter from the sample statistic, every possible sample that could have been drawn should be examined. If all possible samples were actually to be drawn, the distribution of the statistic would be the sampling distribution. Although in practice only one sample is actually drawn, the concept of a sampling distribution is still relevant. It enables us to use probability theory to make inferences about the population values.

Statistical inference
The process of generalising the sample results to a target population.

The important properties of the sampling distribution of the mean, and the corresponding properties for the proportion, for large samples (30 or more) are as follows:

1 The sampling distribution of the mean is a normal distribution (see Appendix 15A). Strictly speaking, the sampling distribution of a proportion is a binomial. For large samples ($n = 30$ or more), however, it can be approximated by the normal distribution.

2 The mean of the sampling distribution of the mean $\left(\overline{X} = \left(\sum_{i=1}^{n} X_i \right)/n \right)$ or of the proportion (p) equals the corresponding population parameter value, μ or π, respectively.

Standard error
The standard deviation of the sampling distribution of the mean or proportion.

3 The standard deviation is called the **standard error** of the mean or the proportion to indicate that it refers to a sampling distribution of the mean or the proportion and not to a sample or a population. The formulas are:

<table>
<tr><td align="center">Mean</td><td align="center">Proportion</td></tr>
<tr><td align="center">$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$</td><td align="center">$\sigma_p = \sqrt{\dfrac{\pi(1-\pi)}{n}}$</td></tr>
</table>

4 Often the population standard deviation, σ, is not known. In these cases, it can be estimated from the sample by using the following formula:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(Xi - \overline{X})^2}{n-1}}$$

or

$$s = \sqrt{\frac{\sum_{i=1}^{n} X_i^2 - \dfrac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}{n-1}}$$

In cases where σ is estimated by s, the standard error of the mean becomes

$$\text{est. } \sigma \overline{X} = \frac{s}{\sqrt{n}}$$

where 'est.' denotes that s has been used as an estimate of σ.

Assuming no measurement error, the reliability of an estimate of a population parameter can be assessed in terms of its standard error.

5 Likewise, the standard error of the proportion can be estimated by using the sample proportion p as an estimator of the population proportion, π, as

$$\text{est. } S_p = \sqrt{\frac{p(1-p)}{n}}$$

z value
The number of standard errors a point is away from the mean.

6 The area under the sampling distribution between any two points can be calculated in terms of **z values** . The z value for a point is the number of standard errors a point is away from the mean. The z values may be computed as follows:

$$z = \frac{\overline{X}\mu}{\sigma \overline{X}}$$

For example, the areas under one side of the curve between the mean and points that have z values of 1.0, 2.0 and 3.0 are, respectively, 0.3413, 0.4772 and 0.4986. (See Table 2 in the Appendix of Statistical Tables.) In the case of proportion, the computation of z values is similar.

7 When the sample size is 10% or more of the population size, the standard error formulas will overestimate the standard deviation of the population mean or proportion. Hence, these should be adjusted by a finite population correction factor defined by

$$\sqrt{\frac{N-n}{N-1}}$$

In this case,

$$\sigma \overline{X} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Statistical approaches to determining sample size

Several qualitative factors should also be taken into consideration when determining the sample size (see Chapter 14). These include the importance of the decision, the nature of the research, the number of variables, the nature of the analysis, sample sizes used in similar studies, incidence rates (the occurrence of behaviour or characteristics in a population), completion rates and resource constraints. The statistically determined sample size is the net or final sample size: the sample remaining after eliminating potential respondents who do not qualify or who do not complete the interview. Depending on incidence and completion rates, the size of the initial sample may have to be much larger. In commercial marketing research, limits on time, money and expert resources can exert an overriding influence on sample size determination. In the GlobalCash Project, the sample size was determined based on these considerations.

The statistical approach to determining sample size that we consider is based on traditional statistical inference.[2] In this approach the precision level is specified in advance. The confidence interval approach to sample size determination is based on the construction of confidence intervals around the sample means or proportions

Ensuring a 99% confidence level in the probability of a tossed pancake returning to the pan.

using the standard error formula. As an example, suppose that a researcher has taken a simple random sample of 300 households to estimate the monthly amount invested in savings schemes and found that the mean household monthly investment for the sample is €182. Past studies indicate that the population standard deviation σ can be assumed to be €55.

We want to find an interval within which a fixed proportion of the sample means would fall. Suppose that we want to determine an interval around the population mean that will include 95% of the sample means, based on samples of 300 households. The 95% could be divided into two equal parts, half below and half above the mean, as shown in Figure 15.1.

Calculation of the confidence interval involves determining a distance below $(\bar{X}_L)$ and above $(\bar{X}_U)$ the population mean $(\bar{X})$, which contains a specified area of the normal curve.

The z values corresponding to $\bar{X}_L$ and $\bar{X}_U$ may be calculated as

$$z_L = \frac{\bar{X}_L - \mu}{\sigma_{\bar{X}}}$$

$$z_U = \frac{\bar{X}_U - \mu}{\sigma_{\bar{X}}}$$
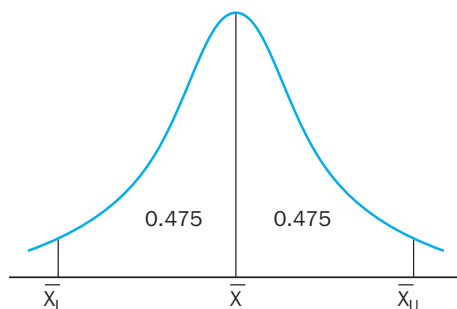
Figure 15.1
The 95% confidence interval

| 0.475 | 0.475 |
| $\bar{X}_L$ | $\bar{X}$ | $\bar{X}_U$ |

where $z_L = -z$ and $z_U = +z$. Therefore, the lower value of $\bar{X}$ is

$$\bar{X}_L = \mu - z\sigma_{\bar{X}}$$

and the upper value of $\bar{X}$ is

$$\bar{X}_U = \mu + z\sigma_{\bar{X}}$$

Note that μ is estimated by $\bar{X}$. The confidence interval is given by

$$\bar{X} \pm z\sigma_{\bar{X}}$$

We can now set a 95% confidence interval around the sample mean of €182. As a first step, we compute the standard error of the mean:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{55}{\sqrt{300}} = 3.18$$

From Table 2 in the Appendix of Statistical Tables, it can be seen that the central 95% of the normal distribution lies within ±1.96 z values. The 95% confidence interval is given by

$$\bar{X} \pm 1.96\sigma_{\bar{X}}$$
$$= 182.00 \pm 1.96\ (3.18)$$
$$= 182.00 \pm 6.23$$

Thus, the 95% confidence interval ranges from €175.77 to €188.23. The probability of finding the true population mean to be within €175.77 and €188.23 is 95%.

Sample size determination: means

The approach used here to construct a confidence interval can be adapted to determine the sample size that will result in a desired confidence interval.[3] Suppose that the researcher wants to estimate the monthly household savings investment more precisely so that the estimate will be within ± €5.00 of the true population value. What should be the size of the sample? The following steps, summarised in Table 15.2, will lead to an answer.

1 Specify the level of precision. This is the maximum permissible difference (D) between the sample mean and the population mean. In our example, $D = \pm$ €5.00.
2 Specify the level of confidence. Suppose that a 95% confidence level is desired.
3 Determine the z value associated with the confidence level using Table 2 in the Appendix of Statistical Tables. For a 95% confidence level, the probability that the population mean will fall outside one end of the interval is 0.025 (0.05/2). The associated z value is 1.96.
4 Determine the standard deviation of the population. This may be known from secondary sources. lf not, it might be estimated by conducting a pilot study. Alternatively, it might be estimated on the basis of the researcher's judgement. For example, the range of a normally distributed variable is approximately equal to ± 3 standard deviations, and one can thus estimate the standard deviation by dividing the range by 6. The researcher can often estimate the range based on knowledge of the phenomenon.
5 Determine the sample size using the formula for the standard error of the mean.

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

$$= \frac{D}{\sigma_{\bar{X}}}$$

or

$$\sigma_{\overline{X}} = \frac{D}{z}$$

or

$$\frac{\sigma}{\sqrt{n}} = \frac{D}{z}$$

or

$$n = \frac{\sigma^2 \times z^2}{D^2}$$

In our example,

$$n = \frac{55^2(1.96)^2}{5^2}$$

$$= 464.83$$
$$= 465 \text{ (rounded to the next highest integer)}$$

It can be seen from the formula for sample size that sample size increases with an increase in the population variability, the degree of confidence, and the precision level required of the estimate.

6 If the resulting sample size represents 10% or more of the population, the finite population correction (fpc) should be applied. The required sample size should then be calculated from the formula

$$n_c = \frac{nN}{N + n - 1}$$

where n = sample size without fpc
        n_c = sample size with fpc

7 If the population standard deviation, σ, is unknown and an estimate is used, it should be re-estimated once the sample has been drawn. The sample standard deviation, s, is used as an estimate of σ. A revised confidence interval should then be calculated to determine the precision level actually obtained.

   Suppose that the value of 55.00 used for σ was an estimate because the true value was unknown. A sample of n = 465 is drawn, and these observations generate a mean $\overline{X}$ of 180.00 and a sample standard deviation s of 50.00. The revised confidence interval is then

$$\overline{X} \pm z s_{\overline{X}} = 180.00 \pm 1.96 \times \frac{50.0}{\sqrt{465}}$$

$$= 180.00 \pm 4.55$$

or $175.45 \leq \mu \leq 184.55$. Note that the confidence interval obtained is narrower than planned, because the population standard deviation was overestimated, as judged by the sample standard deviation.

8 In some cases, precision is specified in relative rather than absolute terms. In other words, it may be specified that the estimate be within plus or minus R percentage points of the mean. Symbolically,

$$D = R\mu$$

In these cases, the sample size may be determined by

$$n = \frac{\sigma^2 \times z^2}{D^2}$$

$$= \frac{C^2 \times z^2}{R^2}$$

where the coefficient of variation $C = \sigma/\mu$ would have to be estimated.

The population size, N, does not directly affect the size of the sample, except when the finite population correction factor has to be applied. Although this may be counter-intuitive, upon reflection it makes sense. For example, if all the population elements are identical on the characteristics of interest, then a sample size of one will be sufficient to estimate the mean perfectly. This is true whether there are 50, 500, 5,000 or 50,000 elements in the population. What directly affects the sample size is the variability of the characteristic in the population. This variability enters into the sample size calculation by way of population variance σ^2 or sample variance s^2.

Sample size determination: proportions

If the statistic of interest is a proportion rather than a mean, the approach to sample size determination is similar. Suppose that the researcher is interested in estimating the proportion of households possessing a debit card. The following steps should be followed.[4]

1 Specify the level of precision. Suppose that the desired precision is such that the allowable interval is set as $D = p - \pi = \pm 0.05$.
2 Specify the level of confidence. Suppose that a 95% confidence level is desired.
3 Determine the z value associated with the confidence level. As explained in the case of estimating the mean, this will be $z = 1.96$.
4 Estimate the population proportion π. As explained earlier, the population proportion may be estimated from secondary sources, or from a pilot study, or may be based on the judgement of the researcher. Suppose that based on secondary data the researcher estimates that 64% of the households in the target population possess a debit card. Hence, $\pi = 0.64$.
5 Determine the sample size using the formula for the standard error of the proportion.

$$\sigma_p = \frac{p - \pi}{z}$$

$$= \frac{D}{z}$$

$$= \sqrt{\frac{\pi(1-\pi)}{n}}$$

or

$$n = \frac{\pi(1-\pi)z^2}{D^2}$$

In our example,

$$n = \frac{0.64(1 - 0.64)(1.96)^2}{(0.05)^2}$$

$$= 354.04$$

$$= 355 \text{ (rounded to the next highest integer)}$$

6 If the resulting sample size represents 10% or more of the population, the finite population correction (fpc) should be applied. The required sample size should then be calculated from the formula

$$n_c = \frac{nN}{N + n - 1}$$

where n = sample size without fpc
n_c = sample size with fpc

7 If the estimate of π turns out to be poor, the confidence interval will be more or less precise than desired. Suppose that after the sample has been taken, the proportion p is calculated to have a value of 0.55. The confidence interval is then re-estimated by employing s_p to estimate the unknown σ_p as

$$p \pm zs_p$$

where

$$S_p = \sqrt{\frac{p(1-p)}{n}}$$

In our example

$$S_p = \sqrt{\frac{0.55(1 - 0.55)}{355}}$$

$$= 0.0264$$

The confidence interval, then, is

$$0.55 \pm 1.96(0.0264) = 0.55 \pm 0.052$$

which is wider than that specified. This is because the sample standard deviation based on $p = 0.55$ was larger than the estimate of the population standard deviation based on $\pi = 0.64$.

    If a wider interval than specified is unacceptable, the sample size can be determined to reflect the maximum possible variation in the population. This occurs when the product is the greatest, which happens when π is set at 0.5. This result can also be seen intuitively. Since one half of the population has one value of the characteristic and the other half the other value, more evidence would be required to obtain a valid inference than if the situation was more clear cut and the majority had one particular value. In our example, this leads to a sample size of

$$n = \frac{0.5(0.5)(1.96)^2}{(0.05)^2}$$

$$= 384.16$$
$$= 385 \text{ (rounded to the next higher integer)}$$

8 Sometimes, precision is specified in relative rather than absolute terms. In other words, it may be specified that the estimate be within plus or minus R percentage points of the population proportion. Symbolically,

$$D = R\pi$$

In such a case, the sample size may be determined by

$$n = \frac{z^2(1 - \pi)}{R^2\pi}$$

Table 15.2 Summary of sample size determination for means and proportions

Steps		Means	Proportions
1	Specify the level of precision.	$D = \pm €5.00$	$D = p - \pi = \pm 0.05$
2	Specify the confidence level (CL).	CL = 95%	CL = 95%
3	Determine the z value associated with the CL.	z value is 1.96	z value is 1.96
4	Determine the standard deviation of the population.	Estimate σ: $\sigma = 55$	Estimate π: $\pi = 0.64$
5	Determine the sample size using the formula for the standard error.	$n = \dfrac{\sigma^2 z^2}{D^2}$ $n = \dfrac{55^2(1.96)^2}{5^2}$ $= 465$	$n = \dfrac{\pi(1-\pi)z^2}{D^2}$ $n = \dfrac{0.64(1-0.64)(1.96)^2}{(0.05)^2}$ $= 355$
6	If the sample size represents 10% of the population, apply the finite population correction (fpc).	$n_c = \dfrac{nN}{N + n - 1}$ $= \bar{X} + zs_{\bar{X}}$	$n_c = \dfrac{nN}{N + n - 1}$ $p \pm zs_p$
7	If necessary, re-estimate the confidence interval by employing s to estimate σ.	$D = R\mu$	$D = R\pi$
8	If precision is specified in relative rather than absolute terms, determine the sample size by substituting for D.	$n = \dfrac{C^2 z^2}{R^2}$	$n = \dfrac{z^2(1-\pi)}{R^2\pi}$

Multiple characteristics and parameters

In the preceding examples, we focused on the estimation of a single parameter. In most marketing research projects, several characteristics, not just one, are of interest. The researcher is required to estimate several parameters, not just one. The calculation of sample size in these cases should be based on a consideration of all the parameters that must be estimated.

For example, suppose that in addition to the mean household spend at a supermarket, it was decided to estimate the mean household spend on clothes and on gifts. The sample sizes needed to estimate each of the three mean monthly expenses are given in Table 15.3 and are 465 for supermarket shopping, 246 for clothes and 217 for gifts. If all three variables were equally important, the most conservative approach would be to select the largest value of $n = 465$ to determine the sample size. This will lead to each variable being estimated at least as precisely as specified. If the researcher was most concerned with the mean household monthly expense on clothes, however, a sample size of $n = 246$ could be selected.

Table 15.3 Sample size for estimating multiple parameters

	Variable		
	Monthly household spend on		
	Supermarket	Clothes	Gifts
Confidence level	95%	95%	95%
z value	1.96	1.96	1.96
Precision level (D)	€5	€5	€4
Standard deviation of the population (σ)	€55	€40	€30
Required sample size (n)	465	246	217

Other probability sampling techniques

So far, the discussion of sample size determination has been based on the methods of traditional statistical inference and has assumed simple random sampling. Next, we discuss the determination of sample size when other sampling techniques are used. The determination of sample size for other probability sampling techniques is based on the same underlying principles. The researcher must specify the level of precision and the degree of confidence and estimate the sampling distribution of the test statistic.

In simple random sampling, cost does not enter directly into the calculation of sample size. In the case of stratified or cluster sampling, however, cost has an important influence. The cost per observation varies by strata or cluster, and the researcher needs some initial estimates of these costs.[5] In addition, the researcher must take into account within-strata variability or within- and between-cluster variability. Once the overall sample size is determined, the sample is apportioned among strata or clusters. This increases the complexity of the sample size formulae. The interested reader is referred to standard works on sampling theory for more information.[6] In general, to provide the same reliability as simple random sampling, sample sizes are the same for systematic sampling, smaller for stratified sampling, and larger for cluster sampling.

Adjusting the statistically determined sample size

The sample size determined statistically represents the final or net sample size that must be achieved to ensure that the parameters are estimated with the desired degree of precision and the given level of confidence. In surveys, this represents the number of interviews that must be completed. To achieve this final sample size, a much greater number of potential respondents have to be contacted. In other words, the initial sample size has to be much larger because typically the incidence rates and completion rates are less than 100%.[7]

Incidence rate
The rate of occurrence of persons eligible to participate in the study expressed as a percentage.

Incidence rate refers to the rate of occurrence or the percentage of persons eligible to participate in the study. Incidence rate determines how many contacts need to be screened for a given sample size requirement.[8] For example, suppose that a study of book purchasing targets a sample of female heads of households aged 25 to 55. Of the women between the ages of 20 and 60 who might reasonably be approached to see if they qualify, approximately 75% are heads of households aged 25 to 55. This means that, on average, 1.33 women would be approached to obtain one qualified respondent. Additional criteria for qualifying respondents (for example, product usage behaviour) will further increase the number of contacts. Suppose that an added eligibility requirement is that the women should have bought a book during the last two months. It is estimated that 60% of the women contacted would meet this criterion. Then the incidence rate is $0.75 \times 0.6 = 0.45$. Thus the final sample size will have to be increased by a factor of (1/0.45) or 2.22.

Completion rate
The percentage of qualified respondents who complete the interview. It enables researchers to take into account anticipated refusals by people who qualify.

Similarly, the determination of sample size must take into account anticipated refusals by people who qualify. The **completion rate** denotes the percentage of qualified respondents who complete the interview. If, for example, the researcher expects an interview completion rate of 80% of eligible respondents, the number of contacts should be increased by a factor of 1.25. The incidence rate and the completion rate together imply that the number of potential respondents contacted – that is, the initial sample size – should be 2.22×1.25 or 2.77 times the sample size required. In

general, if there are c qualifying factors with an incidence of $Q_1 \times Q_2 \times Q_3 \times \ldots \times Q_c$ each expressed as a proportion, the following are true:

$$\text{Incidence rate} = Q_1 \times Q_2 \times Q_3 \times \ldots \times Q_c$$

$$\text{Initial sample size} = \frac{\text{final sample size}}{\text{incidence rate} \times \text{completion rate}}$$

Non-response issues in sampling

The two major non-response issues in sampling are improving response rates and adjusting for non-response. Non-response error arises when some of the potential respondents included in the sample do not respond (see Chapter 3). This is one of the most significant problems in survey research. Non-respondents may differ from respondents in terms of demographic, psychographic, personality, attitudinal, motivational and behavioural variables.[9] Evaluating these differences was detailed in Chapter 14 in the process of sample validation. For a given study, if the non-respondents differ from the respondents on the characteristics of interest, the sample estimates can be seriously biased. Higher response rates, in general, imply lower rates of non-response bias, yet response rate may not be an adequate indicator of non-response bias. Response rates themselves do not indicate whether the respondents are representative of the original sample.[10] Increasing the response rate may not reduce non-response bias if the additional respondents are no different from those who have already responded but do differ from those who still do not respond. As low response rates increase the probability of non-response bias, an attempt should be made to improve the response rate.[11] This is not an issue that should be considered after a survey approach has been decided and a questionnaire designed. Factors that improve response rates are integral to survey and questionnaire design. As detailed in Chapter 13, the marketing researcher should build up an awareness of what motivates their target respondents to participate in a research study. They should ask themselves what their target respondents get in return for spending time and effort, answering set questions in a full and honest manner. The following section details the techniques involved in improving response rates and adjusting for non-response.

Improving the response rates

The primary causes of low response rates are refusals and not-at-homes, as shown in Figure 15.2.
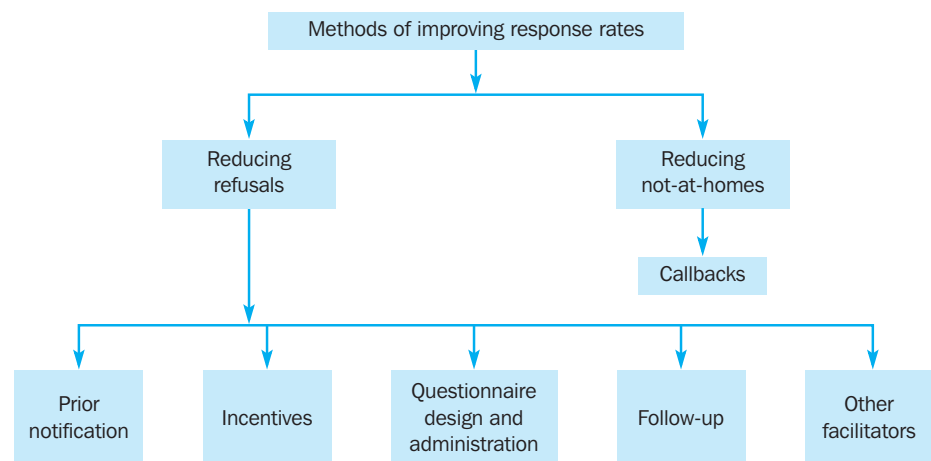
Figure 15.2
Improving response rates

Refusals. Refusals, which result from the unwillingness or inability of people included in the sample to participate, result in lower response rates and increased potential for non-response bias. Given the potential differences between respondents and non-respondents, researchers should attempt to lower refusal rates. This can be done by prior notification, incentives, good questionnaire design and administration, follow-up, and other facilitators.

- *Prior notification.* In prior notification, potential respondents are telephoned, sent a letter or email notifying them of the imminent mail, telephone personal or Internet survey. Prior notification increases response rates, as the respondent's attention is drawn to the purpose of a study and the potential benefits, without the apparent 'chore' of the questionnaire. With the potential respondent's attention focused upon the purpose and benefits, the chances increase for a greater reception when approached to actually complete a survey.[12]

- *Incentives.* Response rates can be increased by offering monetary as well as non-monetary incentives to potential respondents. Monetary incentives can be pre-paid or promised. The pre-paid incentive is included with the survey or questionnaire. The promised incentive is sent to only those respondents who complete the survey. The most commonly used non-monetary incentives are premiums and rewards, such as pens, pencils, books, and offers of survey results.[13] Pre-paid incentives have been shown to increase response rates to a greater extent than promised incentives. The amount of incentive can vary from trivial amounts to tens of euros. The amount of incentive has a positive relationship with response rate, but the cost of large monetary incentives may outweigh the value and quality of additional information obtained.

- *Questionnaire design and administration.* A well-designed questionnaire can decrease the overall refusal rate as well as refusals to specific questions (see Chapter 13). If the questionnaire and experience of answering the questions are interesting for the respondent, using words and logic that are meaningful to them, the response rate can improve. Likewise, the skill used to administer the questionnaire in telephone and personal interviews can increase the response rate. Trained interviewers are skilled in refusal conversion or persuasion. They do not accept a no response without an additional plea. The additional plea might emphasise the brevity of the questionnaire or importance of the respondent's opinion. Skilled interviewers can decrease refusals by about 7% on average. Interviewing procedures are discussed in more detail in Chapter 16.

- *Follow-up.* Follow-up, or contacting the non-respondents periodically after the initial contact, is particularly effective in decreasing refusals in mail and Internet surveys. The researcher might send a reminder to non-respondents to complete and return the questionnaire. Two or three mailings may be needed in addition to the original one. With proper follow-up, the response rate in mail surveys can be increased to 80% or more. Follow-ups can be done by postcard, letter, telephone, email or personal contacts.

- *Other facilitators.* Personalisation, or sending letters addressed to specific individuals, is effective in increasing response rates.[14] The following example illustrates the procedure employed by *Bicycling* magazine to increase its response rate.[15]

***Bicycling* magazine's procedure for increasing response to traditional mail surveys**

Bicycling magazine conducts a semi-annual survey of individual bicycle dealers. The following procedure is used to increase the response to the survey:

1 An 'alert' letter is sent to advise the respondent that a questionnaire is coming.
2 A questionnaire package is posted five days after the 'alert' letter. The package contains a cover letter, a five-page questionnaire, a new $1 bill, and a stamped addressed envelope.
3 A second package containing a reminder letter, a questionnaire and a stamped return envelope is posted five days after the first package.
4 A follow-up postcard is sent a week after the second package.
5 A second follow-up postcard is sent a week after the first.

In a recent survey, 1,000 questionnaires were posted to bicycle dealers, and 68% of these were returned. This represents a good response rate in a mail survey. ■

Not-at-homes. The second major cause of low response rates is not-at-homes. In telephone and in-home personal interviews, low response rates can result if the potential respondents are not at home when contact is attempted. A study analysing 182 commercial telephone surveys involving a total sample of over one million consumers revealed that a large percentage of potential respondents were never contacted. The median non-contact rate was 40%. In nearly 40% of the surveys, only a single attempt was made to contact potential respondents. The results of 259,088 first-call attempts using a sophisticated random-digit dialling system show that less than 10% of the calls resulted in completed interviews, and 14.3% of those contacted refused to participate.[16]

The likelihood that potential respondents will not be at home varies with several factors. People with small children are more likely to be at home. Consumers are more likely to be at home at weekends than on weekdays and in the evening as opposed to during the afternoon. Pre-notification and appointments increase the likelihood that the respondent will be at home when contact is attempted.

The percentage of not-at-homes can be substantially reduced by employing a series of call-backs, or periodic follow-up attempts to contact non-respondents. The decision about the number of call-backs should weigh the benefits of reducing non-response bias against the additional costs. As call-backs are completed, the call-back respondents should be compared with those who have already responded to determine the usefulness of making further call-backs. In most consumer surveys, three or four call-backs may be desirable. Although the first call yields the most responses, the second and third calls have higher response per call. It is important that call-backs be made and controlled according to a prescribed plan.

Adjusting for non-response

Low response rates increase the probability that non-response bias will be substantial. Response rates should always be reported, and whenever possible, the effects of non-response should be estimated. This can be done by linking the non-response rate to estimated differences between respondents and non-respondents. Information on differences between the two groups may be obtained from the sample itself. For example, differences found through call-backs could be extrapolated, or a concen-

trated follow-up could be conducted on a sub-sample of the non-respondents. Alternatively, it may be possible to estimate these differences from other sources.[17] To illustrate, in a survey of owners of vacuum cleaners, demographic and other information may be obtained for respondents and non-respondents from their guarantee cards. For a mail panel, a wide variety of information is available for both groups from syndicate organisations. If the sample is supposed to be representative of the general population, then comparisons can be made with census figures. Even if it is not feasible to estimate the effects of non-response, some adjustments can still be made during data analysis and interpretation.[18] The strategies available to adjust for non-response error include sub-sampling of non-respondents, replacement, substitution, subjective estimates, trend analysis, simple weighting and imputation.

Sub-sampling of non-respondents. Sub-sampling of non-respondents, particularly in the case of mail surveys, can be effective in adjusting for non-response bias. In this technique, the researcher contacts a sub-sample of the non-respondents, usually by means of telephone or personal interviews. This often results in a high response rate within that sub-sample. The values obtained for the sub-sample are then projected to all the non-respondents, and the survey results are adjusted to account for non-response. This method can estimate the effect of non-response on the characteristic of interest.

Replacement. In replacement, the non-respondents in the current survey are replaced with non-respondents from an earlier, similar survey. The researcher attempts to contact these non-respondents from the earlier survey and administer the current survey questionnaire to them, possibly by offering a suitable incentive. It is important that the nature of non-response in the current survey be similar to that of the earlier survey. The two surveys should use similar kinds of respondents, and the time interval between them should be short. As an example, as the GlobalCash survey is repeated two years later, the non-respondents in the present survey may be replaced by the non-respondents in the original survey.

Substitution
A procedure that substitutes for non-respondents other elements from the sampling frame who are expected to respond.

Substitution. In **substitution**, the researcher substitutes for non-respondents other elements from the sampling frame who are expected to respond. The sampling frame is divided into subgroups that are internally homogeneous in terms of respondent characteristics but heterogeneous in terms of response rates. These subgroups are then used to identify substitutes who are similar to particular non-respondents but dissimilar to respondents already in the sample. Note that this approach would not reduce non-response bias if the substitutes are similar to respondents already in the sample.

Subjective estimates. When it is no longer feasible to increase the response rate by sub-sampling, replacement or substitution, it may be possible to arrive at subjective estimates of the nature and effect of non-response bias. This involves evaluating the likely effects of non-response based on experience and available information. For example, married adults with young children are more likely to be at home than single or divorced adults or than married adults with no children. This information provides a basis for evaluating the effects of non-response due to not-at-homes in personal or telephone surveys.

Trend analysis
A method of adjusting for non-response in which the researcher tries to discern a trend between early and late respondents. This trend is projected to non-respondents to estimate their characteristic of interest.

Trend analysis. Trend analysis is an attempt to discern a trend between early and late respondents. This trend is projected to non-respondents to estimate where they stand on the characteristic of interest. For example, Table 15.4 presents the results of several waves of a mail survey. The characteristic of interest is money spent on shop-

Table 15.4 Use of trend analysis in adjusting for non-response

	Percentage response	Average euro expenditure	Percentage of previous wave's response
First mailing	12	412	–
Second mailing	18	325	79
Third mailing	13	277	85
Non-response	(57)	(230)	91
Total	100	275	

ping in supermarkets during the last two months. The known value of the characteristic for the total sample is given at the bottom of the table. The value for each successive wave of respondents becomes closer to the value for non-respondents. For example, those responding to the second mailing spent 79% of the amount spent by those who responded to the first mailing. Those responding to the third mailing spent 85% of the amount spent by those who responded to the second mailing. Continuing this trend, one might estimate that those who did not respond spent 91% [85 + (85 – 79)] of the amount spent by those who responded to the third mailing. This results in an estimate of €252 (277 × 0.91) spent by non-respondents and an estimate of €88 for the average amount spent in shopping at supermarkets during the last two months for the overall sample. Note that the actual amount spent by the respondents was €230 rather than the €252, and that the actual sample average was €275 rather than the €288 estimated by trend analysis. Although the trend estimates are wrong, the error is smaller than the error that would have resulted from ignoring the non-respondents. Had the non-respondents been ignored, the average amount spent would have been estimated at €335 for the sample.

Weighting. **Weighting** attempts to account for non-response by assigning differential weights to the data depending on the response rates.[19] For example, in a survey on personal computers, the sample was stratified according to income. The response rates were 85%, 70% and 40%, respectively, for the high-, medium- and low-income groups. In analysing the data, these subgroups are assigned weights inversely proportional to their response rates. That is, the weights assigned would be 100/85, 100/70 and 100/40, respectively, for the high-, medium- and low-income groups. Although weighting can correct for the differential effects of non-response, it destroys the self-weighting nature of the sampling design and can introduce complications. Weighting is further discussed in Chapter 17 on data preparation.

Weighting
A statistical procedure that attempts to account for non-response by assigning differential weights to the data depending on the response rates.

Imputation. **Imputation** involves imputing, or assigning, the characteristic of interest to the non-respondents based on the similarity of the variables available for both non-respondents and respondents.[20] For example, a respondent who does not report brand usage may be imputed based on the usage of a respondent with similar demographic characteristics. Often there is a high correlation between the characteristic of interest and some other variables. In such cases, this correlation can be used to predict the value of the characteristic for the non-respondents (see Chapter 14).

Imputation
A method to adjust for non-response by assigning the characteristic of interest to the non-respondents based on the similarity of the variables available for both non-respondents and respondents.

International marketing research

When conducting marketing research in foreign countries, statistical estimation of sample size may be difficult because estimates of the population variance may be unavailable. Hence, the sample size is often determined by qualitative considerations, as discussed in Chapter 14: (1) the importance of the decision, (2) the nature of the research, (3) the number of variables, (4) the nature of the analysis, (5) sample sizes used in similar studies, (6) incidence rates, (7) completion rates, and (8) resource constraints. If statistical estimation of sample size is at all attempted, it should be realised that the estimates of the population variance may vary from country to country. For example, in measuring consumer preferences, a greater degree of heterogeneity may be encountered in countries where consumer preferences are not well developed. Thus, it may be a mistake to assume that the population variance is the same or to use the same sample size in different countries.

example

The Chinese take to the sky [21]

The airline industry seems to have a strong and promising market potential in China. The airline market in China is growing rapidly. With billions of euros spent, China is trying to satisfy surging demand and to catch up with the rest of the world. The domestic airline traffic is growing at a rate of up to 30% a year. Strong economic growth, surging foreign trade, and a revival in tourism as the memory of the massacre in Tiananmen Square recedes, have helped to fuel the boom. China is making rapid progress in increasing its fleet and training pilots. For millions of Chinese, air travel is a relatively new experience and many more millions have never flown. Hence, Chinese preferences for air travel are likely to exhibit much more variability compared with Europeans. In a survey to compare attitudes towards air travel in China and European countries, the sample size of the Chinese survey would have to be larger than the European survey in order for the two survey estimates to have comparable precision. ■

Ethics in marketing research

As discussed in this chapter, statistical methods can be used to determine the sample size and, therefore, have an impact on the cost of the project. While this is usually an objective way of determining the sample size, it is, nonetheless, susceptible to fraud. The sample size is heavily dependent on the standard deviation of the variable and there is no way of knowing the standard deviation until the data have been collected. To resolve this paradox, the computation of the sample size must be performed using an estimate of the standard deviation. This estimate is derived based on secondary data, judgement or a small pilot study. By inflating the standard deviation, it is possible to increase the sample size and thus the project revenue. Using the sample size formula, it can be seen that increasing the standard deviation by 20%, for example, will increase the sample size by 44%. But this is clearly unethical.

Ethical dilemmas can arise even when the standard deviation is estimated honestly. It is possible, indeed common, that the standard deviation in the actual study is different from that estimated initially. When the standard deviation is larger than initially estimated, the confidence interval will also be larger than desired. When such a situation arises, the researcher has the responsibility to disclose this to the client and jointly decide on a course of action. The ethical ramifications of miscommunicating the confidence intervals of survey estimates based on statistical samples are underscored in political polling.

Surveys serve up elections

The dissemination of some survey results has been strongly criticised as manipulative and unethical. In particular, the ethics of releasing political poll results before and during the election have been questioned. Opponents of such surveys claim that voters are misled by these results. First, before the election, voters are influenced by whom the polls predict will win. If they see that the candidate they favour is trailing, they may decide not to vote; they assume that there is no way their candidate can win. The attempt to predict the election results while the election is in progress has come under even harsher criticism. Opponents of this practice feel that this predisposes voters to vote for the projected winner or that it may even discourage voters from voting, even though the polls have not closed, because the media projects that there is already a winner. Furthermore, not only are the effects of these projections questionable, but frequently the accuracy of the projections is questionable as well. Although voters may be told a candidate has a certain percentage of the votes within ± 1 per cent, the confidence interval may be much larger, depending on the sample size. ■

Researchers also have the ethical responsibility to investigate the possibility of non-response bias, and make reasonable effort to adjust for non-response. The methodology adopted and the extent of non-response bias found should be clearly communicated.

Internet and computer applications

The main use of the Internet in sample size calculations is to track down potential sampling frames that could be used to define and classify a population. With different sampling frames collected and 'cleaned' in a database package, the ultimate population size can be determined. If there is a finite size to a population, the Internet can play a vital role in tracking down all elements of that population.

Using database packages to record the identity of survey respondents, the researcher can keep track of non-respondents. The database can help to determine whether there are particular geographical locations or types of non-respondent that are problematic. The rapid identification of non-respondents enables researchers to develop tactics to encourage a response.

Microcomputers and mainframes can determine the sample size for various sampling techniques. For simple applications, appropriate sample size formulas can be entered using spreadsheet programs. The researcher specifies the desired precision level, confidence level and population variance and the program determines the appropriate sample size for the study. By incorporating the cost of each sampling unit, the sample size can be adjusted based upon budget considerations.

Summary

The statistical approaches to determining sample size are based on confidence intervals. These approaches may involve the estimation of the mean or proportion. When estimating the mean, determination of sample size using the confidence interval approach requires the specification of precision level, confidence level and population standard deviation. In the case of proportion, the precision level, confidence level and an estimate of the population proportion must be specified. The sample size determined statistically represents the final or net sample size that must be achieved. To achieve this final sample size, a much greater number of potential respondents have to

be contacted to account for reduction in response due to incidence rates and completion rates.

Non-response error arises when some of the potential respondents included in the sample do not respond. The primary causes of low response rates are refusals and not-at-homes. Refusal rates may be reduced by prior notification, incentives, proper questionnaire design and administration, and follow-up. The percentage of not-at-homes can be substantially reduced by call-backs. Adjustments for non-response can be made by sub-sampling non-respondents, replacement, substitution, subjective estimates, trend analysis, simple weighting and imputation.

The statistical estimation of sample size is even more complicated in international marketing research because the population variance may differ from one country to the next. The preliminary estimation of population variance for the purpose of determining the sample size also has ethical ramifications.

Questions

1 Define:
   (a) the sampling distribution,
   (b) finite population correction,
   (c) confidence intervals.

2 What is the standard error of the mean?

3 What is the procedure for constructing a confidence interval around a mean?

4 Describe the difference between absolute precision and relative precision when estimating a population mean.

5 How do the degree of confidence and the degree of precision differ?

6 Describe the procedure for determining the sample size necessary to estimate a population mean, given the degree of precision and confidence and a known population variance. After the sample is selected, how is the confidence interval generated?

7 Describe the procedure for determining the sample size necessary to estimate a population mean, given the degree of precision and confidence but where the population variance is unknown. After the sample is selected, how is the confidence interval generated?

8 How is the sample size affected when the absolute precision with which a population mean is estimated is doubled?

9 How is the sample size affected when the degree of confidence with which a population mean is estimated is increased from 95% to 99%?

10 Define what is meant by absolute precision and relative precision when estimating a population proportion.

11 Describe the procedure for determining the sample size necessary to estimate a population proportion given the degree of precision and confidence. After the sample is selected, how is the confidence interval generated?

12 How can the researcher ensure that the generated confidence interval will be no larger than the desired interval when estimating a population proportion?

13 When several parameters are being estimated, what is the procedure for determining the sample size?

14 Define incidence rate and completion rate. How do these rates affect the determination of the final sample size?

15 What strategies are available for adjusting for non-response?

Appendix: The normal distribution

In this appendix, we provide a brief overview of the normal distribution and the use of the normal distribution table. The normal distribution is used in calculating the sample size, and it serves as the basis for classical statistical inference. Many continuous phenomena follow the normal distribution or can be approximated by it. The normal distribution can, likewise, be used to approximate many discrete probability distributions.[22]

The normal distribution has some important theoretical properties. It is bell-shaped and symmetrical in appearance. Its measures of central tendency (mean, median, and mode) are all identical. Its associated random variable has an infinite range $(-\infty < x < +\infty)$.

The normal distribution is defined by the population mean μ and population standard deviation σ. Since an infinite number of combinations of μ and σ exist, an infinite number of normal distributions exist and an infinite number of tables would be required. By standardising the data, however, we need only one table, such as Table 2 in the Appendix of Statistical Tables. Any normal random variable X can be converted to a standardised normal random variable z by the formula

$$z = \frac{X - \mu}{\sigma}$$

Note that the random variable z is always normally distributed with a mean of 0 and a standard deviation of 1. The normal probability tables are generally used for two purposes: (1) finding probabilities corresponding to known values of X or z, and (2) finding values of X or z corresponding to known probabilities. Each of these uses is discussed.

Finding probabilities corresponding to known values

Suppose that Figure 15A.1 represents the distribution of the number of engineering contracts received per year by an engineering firm. Because the data span the entire history of the firm, Figure 15A.1 represents the population. Therefore, the probabilities or proportion of area under the curve must add up to 1.0. The Marketing Director wishes to determine the probability that the number of contracts received next year will be between 50 and 55. The answer can be determined by using Table 2 of the Appendix of Statistical Tables.

Table 2 gives the probability or area under the standardised normal curve from the mean (zero) to the standardised value of interest, z. Only positive entries of z are listed in the table. For a symmetrical distribution with zero mean, the area from the
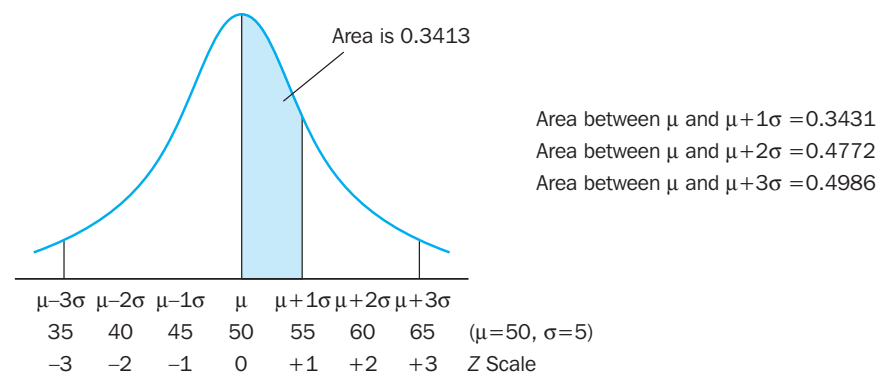
Figure 15A.1
Finding probability corresponding to a known value

Area is 0.3413

Area between μ and $\mu+1\sigma = 0.3431$
Area between μ and $\mu+2\sigma = 0.4772$
Area between μ and $\mu+3\sigma = 0.4986$

$\mu-3\sigma$	$\mu-2\sigma$	$\mu-1\sigma$	μ	$\mu+1\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	
35	40	45	50	55	60	65	$(\mu=50,\ \sigma=5)$
-3	-2	-1	0	$+1$	$+2$	$+3$	Z Scale

mean to $+z$ (i.e. z standard deviations above the mean) is identical to the area from the mean to $-z$ (z standard deviations below the mean).

Note that the difference between 50 and 55 corresponds to a z value of 1.00. Note that, to use Table 2, all z values must be recorded to two decimal places. To read the probability or area under the curve from the mean to $z = +1.00$, scan down the z column of Table 2 until the z value of interest (in tenths) is located. In this case, stop in the row $z = 1.00$. Then read across this row until you intersect the column containing the hundredths place of the z value. Thus, in Table 2, the tabulated probability for $z = 1.00$ corresponds to the intersection of the row $z = 1.0$ with the column $z = 0.00$. This probability is 0.3413. As shown in Figure 15A.1, the probability is 0.3413 that the number of contracts received by the firm next year will be between 50 and 55. It can also be concluded that the probability is 0.6826 (2×0.3413) that the number of contracts received next year will be between 45 and 55.

This result could be generalised to show that for any normal distribution the probability is 0.6826 that a randomly selected item will fall within ±1 standard deviation above or below the mean. Also, it can be verified from Table 2 that there is a 0.9544 probability that any randomly selected normally distributed observation will fall within ±2 standard deviations above or below the mean, and a 0.9973 probability that the observation will fall within ±3 standard deviations above or below the mean.

Finding values corresponding to known properties values

Suppose that the Marketing Director wishes to determine how many contracts must come in so that 5% of the contracts for the year have come in. If 5% of the contracts have come in, 95% of the contracts have yet to come. As shown in Figure 15A.2, this 95% can be broken down into two parts: contracts above the mean (i.e. 50%) and contracts between the mean and the desired z value (i.e. 45%). The desired z value can be determined from Table 2, since the area under the normal curve from the standardised mean, 0, to this z must be 0.4500. From Table 2, we search for the area or probability 0.4500. The closest value is 0.4495 or 0.4505. For 0.4495, we see that the z value corresponding to the particular z row (1.6) and z column (0.04) is 1.64. The z value, however, must be recorded as negative (i.e. $z = -1.64$), since it is below the standardised mean of 0. Similarly, the z value corresponding to the area of 0.4505 is -1.65. Since 0.4500 is midway between 0.4495 and 0.4505, the appropriate z value could be midway between the two z values and estimated as -1.645. The corresponding X value can then be calculated from the standardisation formula, as follows:

$$X = \mu + z\sigma$$
$$= 50 + (-1.645)5$$
$$= 41.775$$
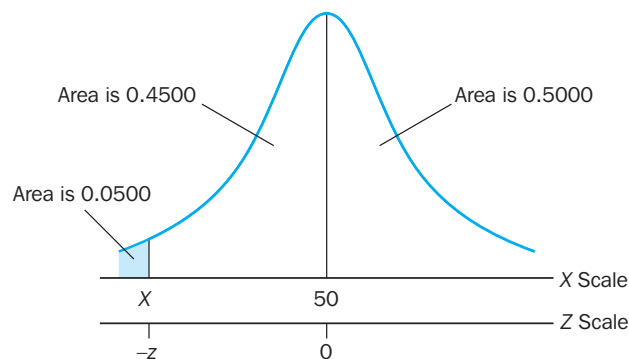
Figure 15A.2
Finding values corresponding to known probabilities
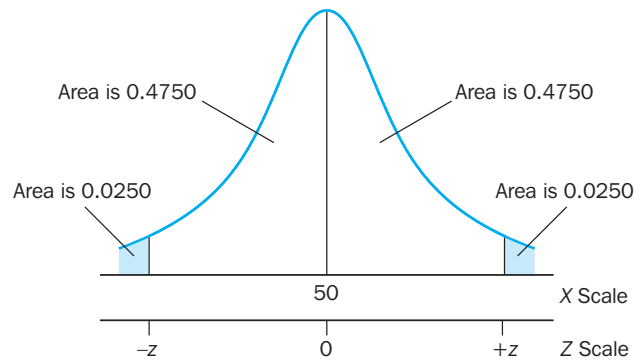
Figure 15 A.3
Finding values corresponding to known probabilities: confidence interval

Area is 0.4750   Area is 0.4750

Area is 0.0250   Area is 0.0250

50   X Scale

−z   0   +z   Z Scale

Suppose that the Marketing Director wanted to determine the interval in which 95% of the contracts for next year are expected to lie. As can be seen from Figure 15A.3, the corresponding z values are ±1.96. This corresponds to X values of 50 ± (1.96)5, or 40.2 and 59.8. This range represents the 95% confidence interval.

Notes

1 A discussion of the sampling distribution may be found in any basic statistics textbook. For example, see Berenson, M.L. and Levine, D.M., *Basic Business Statistics: Concepts and Applications*, 7th edn (Englewood Cliffs, NJ: Prentice Hall, 1999).

2 Other statistical approaches are also available. A discussion of these is beyond the scope of this book, however. The interested reader is referred to Yeh, L. and Van, L.C., 'Bayesian double-sampling plans with normal distributions', *Statistician* 46(2) (1997), 193–207; Blyth, W.G. and Marchant, L.J., 'A self-weighting random sampling technique', *Journal of the Market Research Society* 38(4) (October 1996), 473–9; Nowell, C. and Stanley, L.R., 'Length-biased sampling in mall intercept surveys', *Journal of Marketing Research* 28 (November 1991), 475–9; Gillett, R., 'Confidence interval construction by Stein's method: a practical and economical approach to sample size determination', *Journal of Marketing Research* 26 (May 1989), 237.

3 Chow, S.L., *Statistical Significance* (Thousand Oaks, CA: Sage, 1996).

4 Joseph, L. and Wolfson, D.B., 'Interval-based versus decision-theoretic criteria for the choice of a sample size', *Statistician* 46(2) (1997), 145–9; Frankel, M., 'Sampling theory', in Rossi, P.H., Wright, J.D. and Anderson, A.B. (eds), *Handbook of Survey Research* (New York: Academic Press, 1983), 21–67.

5 For a discussion of estimating sample costs, see Kish, L., *Survey Sampling* (New York: Wiley, 1965) and Sudman, S., *Applied Sampling* (New York: Academic Press, 1976).

6 See, for example, Adcock, C.J., 'Sample size determination – a review', *Statistician* 46(2) (1997), 261–83; Sudman, S., 'Applied sampling', in Rossi, P.H., Wright, J.D. and Anderson, A.B. (eds), *Handbook of Survey Research* (New York: Academic Press, 1983), 145–94.

7 Adjusting for incidence and completion rates is discussed in Dillman, D.A., Singer, E., Clark, J.R. and Treat, J.B., 'Effects of benefits appeals, mandatory appeals, and variations in statements of confidentiality on completion rates for census questionnaires', *Public Opinion Quarterly* 60(3) (Fall 1996), 376–89; Pol, L.G. and Pak, S., 'The use of two stage survey design in collecting data from those who have attended periodic or special events', *Journal of the Market Research Society* 36 (October 1994), 315–26.

8 Lee, K.G., 'Incidence is a key element', *Marketing News* (13 September 1985), 50.

9 Fisher, M.R., 'Estimating the effect of nonresponse bias on angler surveys', *Transactions of the American Fisheries Society* 125(1) (January 1996), 118–26; Martin, C., 'The impact of topic interest on mail survey response behaviour', *Journal of the Market Research Society* 36 (October 1994), 327–38.

10 Hill, A., Roberts, J., Ewings, P. and Gunnell, D., 'Nonresponse bias in a lifestyle survey', *Journal of Public Health Medicine* 19(2) (June 1997), 203–7; McDaniel, S.W., Madden, C.S. and Verille, P., 'Do topic differences affect survey non-response?', *Journal of the Market Research Society* (January 1987), 55–66.

11 For minimising the incidence of non-response and adjusting for its effects, see Chen, H.C., 'Direction, magnitude, and implications of nonresponse bias in mail surveys', *Journal of the Market Research Society* 38(3) (July 1996), 267–76; Brown, M., 'What price response?', *Journal of the Market Research Society* 36 (July 1994), 227–44.

12 Everett, S.A., Price, J.H., Bedell, A.W. and Telljohann, S.K., 'The effect of a monetary incentive in increasing the return rate of a survey of family physicians', *Evaluation and the Health Professions* 20(2) (June 1997), 207–14; Armstrong, J.S. and Lusk, E.J., 'Return postage in mail surveys: a meta-analysis', *Public Opinion Quarterly* (Summer 1987), 233–48; and Yu, J. and Cooper, H., 'A quantitative review of research design effects on response rates to questionnaires', *Journal of Marketing Research* 20 (February 1983), 36–44.

13 Wayman, S., 'The buck stops here when it comes to dollar incentives', *Marketing News* 31(1) (6 January 1997), 9; Biner, P.M. and Kidd, H.J., 'The interactive effects of monetary incentive justification and questionnaire length on mail survey response rates', *Psychology and Marketing* 11(5) (September/October 1994), 483–92.

14 Dillman, D.A., Singer, E., Clark, J.R. and Treat, J.B., 'Effects of benefits appeals, mandatory appeals, and variations in statements of confidentiality on completion rates for census questionnaires', *Public Opinion Quarterly* 60(3) (Fall 1996), 376–89; Gendall, P., Hoek, J. and Esslemont, D., 'The effect of appeal, complexity and tone in a mail survey covering letter', *Journal of the Market Research Society* 37(3) (July 1995),

251–68; Greer, T.V. and Lohtia, R., 'Effects of source and paper color on response rates in mail surveys', *Industrial Marketing Management* 23 (February 1994), 47–54.

15 *Bicycling Magazine's 1997 Semi-annual Study of US Retail Bicycle Stores* (September 1997).

16 Bowen, G.L., 'Estimating the reduction in nonresponse bias from using a mail survey as a backup for nonrespondents to a telephone interview survey', *Research on Social Work Practice* 4(1) (January 1994), 115–28; Kerin, R.A. and Peterson, R.A., 'Scheduling telephone interviews', *Journal of Advertising Research* (May 1983), 44.

17 Rowland, M.L. and Forthofer, R.N., 'Adjusting for non-response bias in a health examination survey', *Public Health Reports* 108(3) (May-June 1993), 380–6.

18 Dey, E.L., 'Working with low survey response rates – the efficacy of weighting adjustments', *Research in Higher Education* 38(2) (April 1997), 215–27.

19 Kessler, R.C., Little, R.J. and Grover, R.M., 'Advances in strategies for minimising and adjusting for survey nonresponse', *Epidemiologic Reviews* 17(1) (1995), 192–204; Ward, J.C., Russick, B. and Rudelius, W., 'A test of reducing call-backs and not-at-home bias in personal interviews by weighting at-home respondents', *Journal of Marketing Research* 2 (February 1985), 66–73.

20 Drane, J.W., Richter, D. and Stoskopf, C., 'Improved imputation of nonresponse to mailback questionnaires', *Statistics in Medicine* 12(3–4) (February 1993), 283–8.

21 Tse, A., 'Estimating the design factor for surveys in Hong Kong', *Marketing Intelligence and Planning* 13(9) (1995), 28–9; *The Economist*, 'Another Chinese take-off' (19 December 1992).

22 This material is drawn from Berenson, M. L. and Levine, D. M., *Basic Business Statistics: Concepts and Applications*, 7th edn (Upper Saddle River, NJ: Prentice Hall, 1999). Adapted by permission of Prentice Hall, Inc. Upper Saddle River, NJ.

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Survey fieldwork

Objectives

After reading this chapter, you should be able to:

1 describe the survey fieldwork process and explain the selecting, training and supervising of fieldworkers, validating fieldwork and evaluating fieldworkers;

2 discuss the training of fieldworkers in making the initial contact, asking the questions, probing, recording the answers and terminating the interview;

3 discuss supervising fieldworkers in terms of quality control and editing, sampling control, control of cheating and central office control;

4 describe evaluating fieldworkers in areas of cost and time, response rates, quality of interviewing and the quality of data;

5 explain the issues related to fieldwork when conducting international marketing research;

6 discuss ethical aspects of survey fieldwork.

No matter how well the research process is designed, the persons working in the field hold the key to quality data.

Overview

Survey fieldwork is a vital process, helping to generate sound marketing research data. During this phase, fieldworkers make contact with potential respondents, administer the questionnaires or observation forms, record the data, and turn in the completed forms for processing. A personal interviewer administering questionnaires door to door, an interviewer intercepting shoppers in the street, a telephone interviewer calling from a central location, a worker mailing questionnaires from an office, an observer counting customers in a particular section of a store, a mystery shopper experiencing the service of a retail outlet and others involved in data collection and supervision of the process are all quantitative fieldworkers.

The marketing researcher faces two major problems when managing fieldwork operations. First of all, fieldwork should be carried out in a consistent manner so that regardless of who administers a questionnaire, the same process is adhered to. This is vital to allow comparisons between all completed questionnaires. Second, fieldworkers to some extent have to approach and motivate potential respondents in a manner that sets the correct purpose for a study and motivates the respondent to spend time answering the questions properly. This cannot be done in a 'robotic' manner; it requires good communication skills and an amount of empathy with respondents, but could be interpreted as a means to bias responses. These two problems may be seen as conflicting, but for the marketing researcher, fieldwork management means resolving these conflicts for each individual data gathering process. This makes survey fieldwork an essential task in the generation of sound research data.

This chapter describes the nature of survey fieldwork and the general survey fieldwork/data collection process. This process involves selecting, training and supervising fieldworkers, validating fieldwork, and evaluating fieldworkers. We briefly discuss survey fieldwork in the context of international marketing research and identify the relevant ethical issues. To begin, we illustrate the rigours of survey fieldwork: imagine trying to conduct interviews in a professional and consistent manner at one of the biggest street parties in Europe.

example

Event of the century[1]

Edinburgh's Hogmanay is branded as the biggest street party in Europe, attracting around 250,000 people into the Scottish capital city, to welcome in the New Year along with national and international coverage. Over the last six years, the celebration, which is actually a programme of events held in the city over five days, has generated around €40m in economic benefit to the city. All the city centre hotels are fully booked and around half of the visitors are from outside Scotland. The major research challenge is to conduct face-to-face interviews with a representative sample of visitors attending the event over the five days but especially with those in the city centre on New Year's Eve itself. The party mood and enhanced interviewer rates means that they are never short of interviewers eager to work on this survey! The questionnaire is designed carefully to produce considerable detail on the patterns of expenditure by visitors to the event but also to establish the extent to which the event attracted visitors from outside the city. ■

The nature of survey fieldwork

Marketing research data are rarely collected by the persons who design the research. Researchers have two major options for collecting their data: they can develop their own organisations or they can contract with a fieldwork agency. In either case, data collection involves the use of some kind of field force. The field force may operate either in the field (personal in-home or in-office, street interview, computer-assisted

personal interviewing, and observation) or from an office (telephone, mail and Internet surveys). The fieldworkers who collect the data typically may have little formal marketing research or marketing training. Their training primarily focuses upon the essential tasks of selecting the correct respondents, motivating them to take part in the research, eliciting the correct answers from them, accurately recording the answers and conveying those answers for analysis. An appreciation of why these tasks fit into the overall context of conducting marketing research is important, but it is not necessary for the survey fieldworker to be trained in the whole array of marketing research skills.

Survey fieldwork and the data collection process

All survey fieldwork involves selecting, training and supervising persons who collect data.[2] The validation of fieldwork and the evaluation of fieldworkers are also parts of the process. Figure 16.1 represents a general framework for the survey fieldwork and data collection process. Even though we describe a general process, it should be recognised that the nature of survey fieldwork varies with the mode of data collection and that the relative emphasis on the different steps will be different for telephone, personal, mail and Internet surveys.

Selecting survey fieldworkers

The first step in the survey fieldwork process is the selection of fieldworkers. The researcher should (1) develop job specifications for the project, taking into account the mode of data collection; (2) decide what characteristics the fieldworkers should have; and (3) recruit appropriate individuals.[3] Interviewers' background characteristics, opinions, perceptions, expectations and attitudes can affect the responses they elicit.[4]

For example, the social acceptability of a fieldworker to the respondent may affect the quality of data obtained, especially in personal interviewing. Researchers generally agree that the more characteristics the interviewer and the respondent have in common, the greater the probability of a successful interview, as illustrated in the following example.

e x a m p l e

Searching for common ground[5]

In a survey dealing with emotional well-being and mental health, older interviewers got better cooperation from older respondents than younger interviewers, and this performance appeared to be independent of years of experience. When the interviewer and the respondent were of the same race the cooperation rate was higher than when there was a mismatch on race. ◼
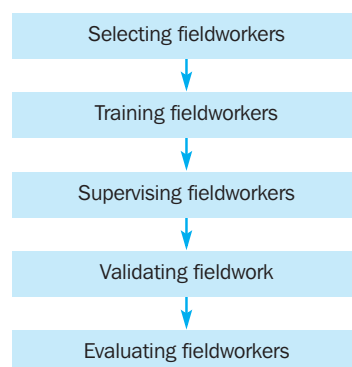
Figure 16.1
Survey fieldwork/data collection process

Selecting fieldworkers → Training fieldworkers → Supervising fieldworkers → Validating fieldwork → Evaluating fieldworkers

Fieldwork can be strenuous and workers must have the stamina to do the job.

Thus, to the extent possible, interviewers should be selected to match respondents' characteristics. The job requirements will also vary with the nature of the problem and the type of data collection method. But there are some general qualifications of survey fieldworkers:

- *Healthy.* Fieldwork can be strenuous, and workers must have the stamina required to do the job.
- *Outgoing.* Interviewers should be able to establish rapport with respondents. They should be able to relate quickly to strangers.
- *Communicative.* Effective speaking, observation and listening skills are a great asset.
- '*Pleasant*' *appearance.* If a fieldworker's physical appearance is unusual (from the respondents' perspective), the data collected may be biased.
- *Educated.* Interviewers must have good reading and writing skills.
- *Experienced.* Experienced interviewers are likely to do a better job in following instructions, obtaining respondent cooperation and conducting the interview, as illustrated in the following example.

example

Your experience counts[6]

Research has found the following effects of interviewer experience on the interviewing process.

- Inexperienced interviewers are more likely to commit coding errors, to mis-record responses, and to fail to probe.
- Inexperienced interviewers have a particularly difficult time filling quotas of respondents.
- Inexperienced interviewers have larger refusal rates. They also accept more 'don't know' responses and refusals to answer individual questions. ■

Training survey fieldworkers

Training survey fieldworkers is critical to the quality of data collected. Training may be conducted in person at a central location or, if the interviewers are geographically dispersed, by mail. Training ensures that all interviewers administer the questionnaire in the same manner so that the data can be collected uniformly. Training should cover making the initial contact, asking the questions, probing, recording the answers and terminating the interview.[7]

Making the initial contact

The initial contact can result in cooperation or the loss of potential respondents.[8] It also sets the potential respondent in a 'frame of mind' to answer subsequent questions. Thus interviewers should be trained to make opening remarks that will convince potential respondents that their participation is important. They should also motivate potential respondents to reflect properly upon the questions posed to them and to answer honestly.

Asking the questions

Even a slight change in the wording, sequence or manner in which a question is asked can distort its meaning and bias the response. Asking questions is an art. Training in asking questions can yield high dividends in eliminating potential sources of bias. Changing the phrasing or order of questions during the interview can make significant differences in the response obtained. The following are guidelines for asking questions in a consistent manner:[9]

1 Be thoroughly familiar with the purpose of the questionnaire.
2 Be thoroughly familiar with the structure of the questionnaire.
3 Ask the questions in the order in which they appear in the questionnaire.
4 Use the exact wording given in the questionnaire.
5 Read each question slowly.
6 Repeat questions that are not understood.
7 Ask every applicable question.
8 Follow instructions, working through any filter questions, and probe carefully.

Probing

Probing
A motivational technique used when asking survey questions to induce the respondents to enlarge on, clarify or explain their answers and to help the respondents focus on the specific content of the interview.

Probing is intended to motive respondents to enlarge on, clarify or explain their answers. Probing also helps respondents focus on the specific content of the interview and provide only relevant information. Probing should not introduce any bias. An example of the effect of interviewer bias comes from a survey in which one of the authors helped in data analysis (but not in the management of the whole research process!). The survey related to bread and cake buying habits with one particular question focusing upon 'large cakes' that respondents had bought over the previous 12 months. In analysing the data a percentage had replied '*Christmas cake*'. When analysed further, all the respondents who said '*Christmas cake*' had been interviewed by the same interviewer. The conclusion from this analysis was that the interviewer in question had used their own probe to make the interview process work. None of the other interviewers had used this probe, which meant there was an inconsistent approach in eliciting answers from respondents. The paradox faced by the survey designers in this example was that the 'rogue' interviewer *may* have used a probe that elicited a true representation of large cake purchasing, the other interviewers consistently failing to draw out a 'true' response.

To help in the process of probing, the following list details some commonly used techniques.[10]

1 *Repeating the question*. Repeating the question in the same words can be effective in eliciting a response.

2 *Repeating the respondent's reply*. Respondents can be stimulated to provide further comments by repeating their replies verbatim. This can be done as the interviewer records the replies.

3 *Using a pause or silent probe*. A silent probe, or an expectant pause or look, can cue the respondent to provide a more complete response. The silence should not become embarrassing, however.

4 *Boosting or reassuring the respondent*. If the respondent hesitates, the interviewer should reassure the respondent with comments such as 'There are no right or wrong answers. We are just trying to get your opinions.' If the respondent needs an explanation of a word or phrase, the interviewer should not offer an interpretation, unless written instructions to do so have been provided. Rather, the responsibility for the interpretation should be returned to the respondent. This can be done with a comment such as 'Just whatever it means to you.'

5 *Eliciting clarification*. The respondent's motivation to cooperate with the interviewer and provide complete answers can be aroused with a question: 'I don't quite understand what you mean by that. Could you please tell me a little more?'

6 *Using objective or neutral questions or comments*. Table 16.1 provides several examples of the common questions or comments used as probes.[11] Corresponding abbreviations are also provided. The interviewer should record the abbreviations on the questionnaire in parentheses next to the question asked.

Table 16.1 Commonly used probes and abbreviations

Standard interviewer's probe	Abbreviation
Any other reason?	(AO?)
Any others?	(Other?)
Anything else?	(AE or Else?)
Could you tell me more about your thinking on that?	(Tell more)
How do you mean?	(How mean?)
Repeat question	(RQ)
What do you mean?	(What mean?)
Which would be closer to the way you feel?	(Which closer?)
Why do you feel this way?	(Why?)
Would you tell me what you have in mind?	(What in mind?)

The above list seems straightforward but there are hidden dangers. For example, probing '*why*' respondents behave in a particular manner or feel about a particular issue takes the interview into the realms of the qualitative interview. Compare the context of the street interview with a short structured questionnaire to the context of the qualitative interview with a questioning approach structured to the respondent and where a greater amount of rapport may be developed. The latter scenario is much more conducive to eliciting '*why*' respondents behave or feel as they do. The question '*why*' is an example of a seemingly simple question that can create many problems of consistency in fieldwork. In the greater majority of circumstances, '*why*' should be treated as a qualitative issue.

Recording the answers

Although recording respondent answers seems simple, several mistakes are common.[12] All interviewers should use the same format and conventions to record the interviews and edit completed interviews. Although the rules for recording answers to structured questions vary with each specific questionnaire, the general rule is to check the box that reflects the respondent's answer. The general rule for recording answers to unstructured questions is to record the responses verbatim. The following guidelines help to record answers to unstructured questions.

1 Record responses during the interview.
2 Use the respondent's own words.
3 Do not summarise or paraphrase the respondent's answers.
4 Include everything that pertains to the question objectives.
5 Include all probes and comments.
6 Repeat the response as it is written down.

Terminating the interview

The interview should not be closed before all the information is obtained. Any spontaneous comments the respondent offers after all the formal questions have been asked should be recorded. The interviewer should answer the respondent's questions about the project. The respondent should be left with a positive feeling about the interview. It is important to thank the respondent and express appreciation.

The Association of Market Survey Organisations (AMSO) publishes a 'Thank You' pamphlet that can be handed out to respondents who have taken part in an interview. As well as thanking respondents for their cooperation, it serves the purpose of educating respondents about the nature and purpose of marketing research, distinguishing marketing research from 'sugging' and 'frugging' as explained in the Ethics section of Chapter 1.

The pamphlet explains why the respondent was chosen, the manner in which the interview was conducted and the purpose of marketing research interviewing. The following example presents extracts from the AMSO pamphlet.

example

Your time has been of great value. Thank you!

Thank you, for taking the time to give this interview; we hope you enjoyed it. Your interviewer is professionally trained by the company for which he/she works, a company which belongs to the Association of Market Survey Organisations. The Association exists to ensure that its members maintain the highest professional standards of market research. Your interviewer carries an identity card which guarantees they are a genuine market researcher.
AND ...
The way in which you were interviewed. In order to make sure that a representative sample was interviewed you may have been asked certain questions about your age, occupation, income and other descriptive details. These questions will be used in the research analysis to check the sample against other statistical information.

The questionnaire which the interviewer used will have been carefully constructed and tested by experienced researchers. The interviewer will have been instructed to read out the questions exactly as they are printed, and is not allowed to change the wording or give a personal opinion. These precautions are taken to ensure that your answers are not influenced in any way by the interviewer. ■

Summary of training issues

To encapsulate the process of interviewer training, the following list summarises the nature and scope of areas in which a marketing research interviewer should be trained.

1 The marketing research process: how a study is developed, implemented and reported.

2 The importance of the interviewer to this process; the need for honesty, objectivity, organisational skills and professionalism.

3 Confidentiality of the respondent and the client.

4 Familiarity with marketing research terminology.

5 The importance of following the exact wording and recording responses verbatim.

6 The purpose and use of probing and clarifying techniques.

7 The reason for and use of classification and respondent information questions.

8 A review of samples of instructions and questionnaires.

9 The importance of the respondent's positive feelings about survey research.

Conversely, marketing researchers should be trained in the 'experience' of gathering data in the field with a practical knowledge of what works in terms of:

■ Motivating potential respondents to take part in a survey
■ Questions that will elicit the required data
■ Probes that can be consistently applied
■ An interview process that does not confuse or cause boredom in the respondent.

The marketing researcher needs to appreciate what the respondent and the interviewer go through in the interview process. Without such an understanding, the questionnaires and interview procedures, which seem fine on paper, can lead to very poor quality data. The following example illustrates the experiences of a 'typical' respondent and researcher. Whilst the feelings expressed cannot be generalised to all interview situations, the lesson from this is that the marketing researcher should aim to understand how their target respondents and interviewers feel about the process. These feelings must form an integral part of any research design that generates sound and accurate data.

| example |

How was it for you? [13]

A respondent and an interviewer describe what their interview experiences were like for them.

The Respondent

'I felt sorry for the interviewer, she was going around all these houses and nobody was in, so I agreed to take part in the survey. The interview did not take that long, only about 10 to 15 minutes, slightly less time than the interviewer said it would. The interviewer was smartly dressed, professional and helpful. She prompted me but did not actually push me. The experience was enjoyable, it was fun, and not a bad way to spend 10 to 15 minutes, although I think that is long enough. I like taking part in a survey if the subject matter is relevant to your life and you feel that your views are being taken into account. I think a lot of women prefer other females (or gay men) to interview them as there is an empathy there, and they might not feel they can be as honest or chatty with men. The age of the interviewer should relate to the subject matter. For example, if you are asking about children, then you should have an interviewer in a mother's age group. I think it is important to actually be in the same position as someone being surveyed. In an interview, you should be honest, do not tell them what you think they want to hear, relax, be friendly and go with the flow. A lot depends on the respondent as well as the interviewer. There has to be a bit of banter between the two of you.'

The Interviewer

'I do not have a typical day. If I am doing quota sampling I will do around 10 interviews a day. If it is pre-selected, then I will do 3 to 4 in-depth interviews. But if it's exit interviewing, I can do as many as 20 in a shift. There are pressures to the job sometimes. Getting your quota is like looking for a needle in a haystack. People are much more suspicious, and fewer will open

the doors these days. I have interviewed through wrought iron gates, letter boxes, front room windows, and with the chain on the door. For your own safety, you must be aware of where you are and what's around you. Technology has not made my job easier. I hate CAPI and I do not use it now. It's slower, heavy to carry around, it sometimes crashes and I feel that interviewing using pen and paper flows better. My job could be made easier by keeping questionnaires short, and using proper screening questions. The essence is to keep it interesting. The worst thing in the world is when you have got a survey that repeats itself and is boring; huge lists are our worst enemy. All I ask of a respondent is that they are honest, they do not have to be articulate or have strong opinions. There are two keys to successful interviewing, smile and be polite at all times, so that it is very hard for people to be rude to you, and be firm and in control of the interview.' ■

Go to the Companion Website and read Professional Perspective 12 'The field-good factor' by Pat Dowding. Pat describes the problems faced by the marketing research industry through the neglect of survey field issues. She goes on to discuss the directions the industry can take to remedy these problems.

Supervising survey fieldworkers

Supervising survey fieldworkers means making sure that they are following the procedures and techniques in which they were trained. Supervision involves quality control and editing, sampling control, control of cheating and central office control.

Quality control and editing

Quality control of fieldworkers requires checking to see whether the field procedures are being properly implemented.[14] If any problems are detected, the supervisor should discuss them with the fieldworkers and provide additional training if necessary. To understand the interviewers' problems related to a specific study, the supervisors should also do some interviewing. Supervisors should collect questionnaires and other observation forms and check them daily. They should examine the questionnaires to make sure all appropriate questions have been completed, that unsatisfactory or incomplete answers have not been accepted, and that the writing is legible.

Supervisors should also keep a record of hours worked and expenses. This will allow a determination of the cost per completed interview, whether the job is moving on schedule, and whether any interviewers are having problems.

Sampling control

Sampling control
An aspect of supervising that ensures that the interviewers strictly follow the sampling plan rather than select sampling units based on convenience or accessibility.

An important aspect of supervision is **sampling control**, which attempts to ensure that the interviewers are strictly following the sampling plan rather than selecting sampling units based on convenience or accessibility.[15] Interviewers tend to avoid homes, offices and people (sampling units) that they perceive as difficult or undesirable. If the sampling unit is not at home, for example, interviewers may be tempted to substitute the next available unit rather than call back. Interviewers sometimes stretch the requirements of quota samples. For example, a 58-year-old person may be placed in the 46 to 55 category and interviewed to fulfil quota requirements.

To control these problems, supervisors should keep daily records of the number of calls made, the number of not-at-homes, the number of refusals, the number of completed interviews for each interviewer, and the total for all interviewers under their control.

Central office control

Supervisors provide quality and cost-control information to the central office so that a total progress report can be maintained. In addition to the controls initiated in the field, other controls may be added at the central office to identify potential problems. Central office control includes tabulation of quota variables, important demographic characteristics and answers to key variables.

Validating survey fieldwork

An interviewer may cheat by falsifying part of an answer to make it acceptable or may fake answers. The most blatant form of cheating occurs when the interviewer falsifies the entire questionnaire, merely filling in fake answers without contacting the respondent. Cheating can be minimised through proper training, rewards, supervision and validation of fieldwork.[16] Validating fieldwork means verifying that the fieldworkers are submitting authentic interviews. One means to achieve this is by asking respondents to give their names and telephone numbers at the end of an interview. To validate the study, the supervisors call 10–25% of the respondents to enquire whether the fieldworkers actually conducted the interviews. The supervisors ask about the length and quality of the interview, reaction to the interviewer and basic demographic data. The demographic information is cross-checked against the information reported by the interviewers on the questionnaires. The major drawback of this approach is that respondents may not trust interviewers with a name and telephone number, perhaps believing that it is to be used to generate a sale, i.e. it can be confused with a 'sugging' or 'frugging' approach.

Evaluating survey fieldworkers

It is important to evaluate survey fieldworkers to provide them with feedback on their performance as well as to identify the better fieldworkers and build a better, high-quality field force. The evaluation criteria should be clearly communicated to the fieldworkers during their training. The evaluation of fieldworkers should be based on the criteria of cost and time, response rates, quality of interviewing and quality of data.[17]

Cost and time

Interviewers can be compared in terms of the total cost (salary and expenses) per completed interview. If the costs differ by city size, comparisons should be made only among fieldworkers working in comparable cities. Fieldworkers should also be evaluated on how they spend their time. Time should be broken down into categories such as actual interviewing, travel and administration.

Response rates

It is important to monitor response rates on a timely basis so that corrective action can be taken if these rates are too low.[18] Supervisors can help interviewers with an inordinate number of refusals by listening to the introductions they use and providing immediate feedback. When all the interviews are over, different fieldworkers' percentage of refusals can be compared to identify the more able interviewers.

Quality of interviewing

To evaluate interviewers on the quality of interviewing, the supervisor must directly observe the interviewing process. The supervisor can do this in person or the field worker can record the interview on tape. The quality of interviewing should be evaluated in terms of (1) the appropriateness of the introduction, (2) the precision with which the fieldworker asks questions, (3) the ability to probe in an unbiased manner, (4) the ability to ask sensitive questions, (5) interpersonal skills displayed during the interview, and (6) the manner in which the interview is terminated.

Quality of data

The completed questionnaires of each interviewer should be evaluated for the quality of data. Some indicators of quality data are that (1) the recorded data are legible; (2) all instructions, including skip patterns, are followed; (3) the answers to unstructured questions are recorded verbatim; (4) the answers to unstructured questions are meaningful and complete enough to be coded; and (5) item non-response occurs infrequently.

International marketing research

The selection, training, supervision and evaluation of survey fieldworkers is critical in international marketing research. Local fieldwork agencies are unavailable in many countries. Therefore, it may be necessary to recruit and train local fieldworkers or import trained foreign workers. The use of local fieldworkers is desirable, because they are familiar with the local language and culture. They can create an appropriate climate for the interview and sensitivity to the concerns of the respondents. Extensive training may be required and close supervision may be necessary. As observed in many countries, local interviewers tend to help the respondent with the answers and select household or sampling units based on personal considerations rather than the sampling plan. Validation of fieldwork is critical. Proper application of fieldwork procedures can greatly reduce these difficulties and result in consistent and useful findings, as the following example illustrates.

example

Americanism unites Europeans [19]

An image study conducted by Research International, a British market research company, showed that despite unification of the European market, European consumers still tend to favour American products. The survey was conducted in Britain, Germany, Italy and the Netherlands. In each country, local interviewers and supervisors were used because it was felt they would be able to identify better with the respondents. The fieldworkers, however, were trained extensively and supervised closely to ensure quality results and to minimise the variability in country-to-country results due to differences in interviewing procedures.

A total of 6,724 personal interviews were conducted. Some of the findings were that Europeans gave US products high marks for being innovative and some countries also regarded them as fashionable and of high quality. Interestingly, France, usually considered anti-American, also emerged as pro-American. Among the 1,034 French consumers surveyed, 40% considered US products fashionable and 38% believed that they were innovative, whereas 15% said US products were of high quality. In addition, when asked what nationality they preferred for a new company in their area, a US company was the first choice. These findings were comparable and consistent across the four countries. A key to the discovery of these findings was the use of local fieldworkers and extensive training and supervision which resulted in high-quality data. ■

Ethics in marketing research

Marketing researchers and survey fieldworkers should make respondents feel comfortable when participating in research activities. This is vital in order to elicit the correct responses for a specific project but also more broadly for the health of the marketing research industry. A respondent who feels that their trust has been abused, who found an interview to be cumbersome and boring, or who fails to see the purpose of a particular study, is less likely to participate in further marketing research efforts. Collectively, the marketing research industry has the responsibility to look after their most precious assets – willing and honest respondents.

Many marketing researchers do not meet respondents face to face, or if they have done it may have occurred many years ago. Not being in the field, marketing researchers can lose an awareness of what it is like to actually collect data in the field. Without this awareness, research designs that on paper seem feasible are difficult to administer in the field in a consistent manner. If there are problems in collecting data in the field, these may not always be attributable to the training and quality of fieldworkers; the blame may lie with the research designer. The marketing researcher, therefore, has an ethical responsibility to the fieldworker and the respondent. Their responsibility lies in an awareness of the process that the fieldworker and respondent go through in the field for each individual piece of research they design. Poor research design can leave fieldworkers facing very disgruntled respondents and can cause great damage.

Good marketing researchers have an awareness of their responsibilities to fieldworkers and respondents. The marketing researcher may take great care in understanding the difficulties of collecting data in the field. They may go to great pains to ensure that the data gathering process works well for the fieldworker and respondent alike. The fieldworker may have been told about the purpose of the study, the purpose of particular questions, the means to select and approach respondents and the means to correctly elicit responses from respondents. However, fieldworkers may behave in an unethical manner. They may cut corners in terms of selecting the correct respondents, posing questions and probes, and recording responses. In such circumstances the fieldworker can cause much damage to an individual study and to the long-term relationship with potential respondents. Thus it becomes a vital part of fieldworker training to demonstrate the ethical responsibilities they have in collecting data.

The use of codes of conduct and guidelines can help fieldworkers to be aware of their responsibilities (see the ESOMAR code on **www.esomar.nl**). It is good to have a specific set of guidelines to help with fieldwork, but applying them is another matter. Marketing researchers who train, develop and reward their survey fieldworkers well create an atmosphere where these guidelines work in practice.

Internet and computer applications

Regardless of which method is used for interviewing (telephone, personal, mail or electronic), the Internet can play a valuable role in all the phases of survey fieldwork: selection, training, supervision, validation and evaluation of fieldworkers. As far as selection is concerned, interviewers can be located, interviewed and hired by using the Internet. This process can be initiated, for example, by posting job vacancies notices for interviewers at the company Website, bulletin boards and other suitable locations. While this would confine the search to only Internet-savvy interviewers, this may well be a qualification to look for in the current marketing research environment.

Similarly, the Internet with its multimedia capabilities can be a good supplementary tool for training the fieldworkers in all aspects of interviewing. Training in this manner can complement personal training programmes and add value to the process. Supervision is enhanced by facilitating communication between the supervisors and the interviewers via email and secured chatrooms. Central office control can be strengthened by posting progress reports, quality and cost control information on a secured location at a Website, so that it is easily available to all the relevant parties.

Validation of fieldwork, especially for personal and telephone interviews, can be easily accomplished for those respondents who have an email address or access to the Internet. These respondents can be sent a short verification survey by email or asked to visit a Website where the survey is posted. Finally, the evaluation criteria can be communicated to the fieldworkers during the training stage by using the Internet, and performance feedback can also be provided to them by using this medium.

Call disposition
Call disposition records the outcome of an interview call.

Microcomputers and mainframes can be used in fieldwork for respondent selection, interviewer planning, supervision and control. Computers can also be used to manage mailing lists. For example, mailing lists can be sorted according to postal codes, geographical regions (which may include drive times from the centre of a city or from a shopping centre) or other pre-specified respondent characteristics. Computers can generate accurate and timely reports for supervision and control purposes. These include quota reports, **call disposition** reports, incidence reports, top-line reports of respondent data, and interviewer productivity reports. Automatic reporting enhances supervision and control and increases the overall quality of data collection. Because less time is spent compiling reports, more time can be spent on data interpretation and on supervision.

Summary

Researchers have two major options in the generation of sound research data: developing their own organisations or contracting with fieldwork agencies. In either case, data collection involves the use of a field force. Fieldworkers should be healthy, outgoing, communicative, pleasant, educated and experienced. They should be trained in important aspects of fieldwork, including making the initial contact, asking the questions, probing, recording the answers and terminating the interview. Supervising fieldworkers involves quality control and editing, sampling control, control of cheating and central office control. Validating fieldwork can be accomplished by calling 10–25% of those who have been identified as interviewees and enquiring whether the interviews took place. Fieldworkers should be evaluated on the basis of cost and time, response rates, quality of interviewing and quality of data collection.

Fieldwork should be carried out in a consistent manner so that, regardless of who administers a questionnaire, the same process is adhered to. This is vital to allow comparisons between collected data. Fieldworkers to some extent have to approach and motivate potential respondents in a manner that sets the correct purpose for a study and motivates the respondent to spend time answering the questions properly. This cannot be done in a 'robotic' manner; it requires good communication skills and an amount of empathy with respondents. This makes the issue of managing fieldwork an essential task in the generation of sound research data.

Selecting, training, supervising and evaluating fieldworkers is even more critical in international marketing research because local fieldwork agencies are not available in many countries. Ethical issues include making the respondents feel comfortable in the data collection process so that their experience is positive.

Questions

1 Why do marketing researchers need to use survey fieldworkers?

2 Describe the survey fieldwork/data collection process.

3 What qualifications should survey fieldworkers possess?

4 What are the guidelines for asking questions?

5 Describe and illustrate the differences between probing in a survey and in a depth interview.

6 Evaluate what may be done to help interviewers probe correctly and consistently.

7 Outline the advantages and disadvantages of the interviewer developing a rapport with respondents.

8 How should the answers to unstructured questions be recorded?

9 How should the survey fieldworker terminate the interview?

10 What aspects are involved in the supervision of survey fieldworkers?

11 How can respondent selection problems be controlled?

12 What is validation of survey fieldwork? How is this done?

13 Describe the criteria that should be used for evaluating survey fieldworkers.

14 Describe the major sources of error related to survey fieldwork.

15 Comment on the following field situations, making recommendations for corrective action.

(a) One of the interviewers has an excessive rate of refusals in in-home personal interviewing.

(b) In a CATI situation, many phone numbers are giving a busy signal during the first dialling attempt.

(c) An interviewer reports that at the end of the interviews many respondents asked if they had answered the questions correctly.

(d) While validating the fieldwork, a respondent reports that she cannot remember being interviewed over the telephone, but the interviewer insists that the interview was conducted.

Notes

1 Costley, T., 'Event of the century', *Research* (June 1999), 44–5.

2 Frey, J.H. and Oishi, S.M., *How to Conduct Interviews by Telephone and In Person* (Thousand Oaks, CA: Sage, 1995); Fowler Jr, E. and Mangione, T.W., 'The role of interviewer training and supervision in reducing effects on survey data', in *Proceedings of the American Statistical Association Meetings, Survey Research Methods Section* (Washington, DC: American Statistical Association, 1983), 124–8.

3 Muller, G.D. and Miller, J., 'Interviewers make the difference', *Marketing Research: A Magazine of Management and Applications* 8(1) (Spring 1996), 8–9; Morton-Williams, J., *Interviewer Approaches* (Brookfield, Ashgate Publishing, 1993); Groves, R.M. and Magilavy, L.J., 'Measuring and explaining interviewer effects in centralized telephone surveys', *Public Opinion Quarterly* 50 (Summer 1986), 251–66.

4 Catina, J.A., Binson, D., Canchola, J., Pollack, L.M. *et al.*, 'Effects of interviewer gender, interviewer choice, and item wording on responses to questions concerning sexual behaviour', *Public Opinion Quarterly* 60(3) (Fall 1996), 345–75; Coulter, P.B., 'Race of interviewer effects on telephone interviews', *Public Opinion Quarterly* 46 (Summer 1982), 278–84; and Singer, E., Frankel, M.R. and Glassman, M.B., 'The effect of interviewer characteristics and expectations on response', *Public Opinion Quarterly* 41 (Spring 1983), 68–83.

5 Davis, D.W., 'Nonrandom measurement error and race of interviewer effects among African Americans', *Public Opinion Quarterly* 61(1) (Spring 1997), 183–207; Barker, R.E., 'A demographic profile of marketing research interviewers', *Journal of the Market Research Society* 29 (July 1987), 279–92.

6 Collins, M. and Butcher, B., 'Interviewer and clustering effects in an attitude survey', *Journal of the Market Research Society* 25 (January 1983), 39–58; Johnson, R.E.Q., 'Pitfalls in research: the interview as an illustrative model', *Psychological Reports* 38 (1976), 3–17.

7 Kiecker, P. and Nelson, J.E., 'Do interviewers follow telephone survey instructions?', *Journal of the Market Research Society* 38(2) (April 1996), 161–76; Guenzel, P.J., Berkmans, T.R. and Cannell, C.F., *General Interviewing Techniques* (Ann Arbor, MI: Institute for Social Research, 1983).

8 Couper, M.P., 'Survey introductions and data quality', *Public Opinion Quarterly* (Summer 1997), 317–38.

9 This section follows closely the material in *Interviewer's Manual*, revised edition (Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan, 1976); and Guenzel, P.J., Berkmans, T.R. and Cannell, C.E., *General Interviewing Techniques* (Ann Arbor, MI: Institute for Social Research, 1983).

10 For an extensive treatment of probing, see *Interviewer's Manual*, revised edition (Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan, 1976), 15–19; and Marchetti, M., 'Probing customer problems', *Sales and Marketing Management* 148(3) (March 1996), 46.

11 *Interviewer's Manual*, revised edition (Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan, 1976), 16. Reprinted by permission of the Institute for Social Research.

12 'Market research industry sets up interviewing quality standards', *Management-Auckland* 44(2) (March 1997), 12; Morton-Williams, J. and Sykes, W., 'The use of interaction coding and follow-up interviews to investigate comprehension of survey questions', *Journal of the Market Research Society* 26 (April 1984), 109–27.

13 Park, C., 'How was it for you?', *Research*, Fieldwork Supplement, (July 2000), 8–9.

14 Hurley, R.F. and Laitamaki, J.M., 'Total quality research: integrating markets and the organization', *California Management Review* 38(1) (Fall 1995), 59–78; Collins, M. and Butcher, B., 'Interviewer and clustering effects in an attitude survey', *Journal of the Market Research Society* 25 (January 1983), 39–58.

15 Czaja, R., Blair, J. and Sebestik, J.P., 'Respondent selection in telephone survey: a comparison of three techniques', *Journal of Marketing Research* (August 1982), 381–5.

16 Greengard, S., '50% of your employees are lying, cheating & stealing', *Workforce* 76(10) (October 1997), 44–53; Tull, D.S. and Richards, L.E., 'What can be done about interviewer bias', in Sheth, J. (ed.), *Research in Marketing* (Greenwich, CT: JAI Press, 1980), 143–62.

17 Pulakos, E.D., Schmitt, N., Whitney, D. and Smith, M., 'Individual differences in interviewer ratings: the impact of standardization, consensus discussion and sampling error on the validity of a structured interview', *Personnel Psychology* 49(1) (Spring 1996), 85–102.

18 Edmonston, J., 'Why response rates are declining', *Advertising Age's Business Marketing* 82(8) (September 1997), 12.

19 Semon, T.T., 'Select local talent when conducting research abroad', *Marketing News* 31(19) (15 September 1997), 28; Wentz, L., 'Poll: Europe favors US products', *Advertising Age* (23 September 1991).

Data preparation

Objectives

After reading this chapter, you should be able to:

1 discuss the nature and scope of data preparation and the data preparation process;

2 explain questionnaire checking and editing and the treatment of unsatisfactory responses by returning to the field, assigning missing values and discarding unsatisfactory responses;

3 describe the guidelines for coding questionnaires, including the coding of structured and unstructured questions;

4 discuss the data cleaning process and the methods used to treat missing responses: substitution of a neutral value, imputed response, casewise deletion and pairwise deletion;

5 state the reasons for and methods of statistically adjusting data: weighting, variable re-specification and scale transformation;

6 describe the procedure for selecting a data analysis strategy and the factors influencing the process;

7 classify statistical techniques and give a detailed classification of univariate techniques as well as a classification of multivariate techniques;

8 understand the intra-cultural, pan-cultural and cross-cultural approaches to data analysis in international marketing research;

9 identify the ethical issues related to data processing, particularly the discarding of unsatisfactory responses, violation of the assumptions underlying the data analysis techniques, and evaluation and interpretation of results.

Perhaps the most neglected series of activities in the marketing research process. Handled with care, data preparation can substantially enhance the quality of statistical results.
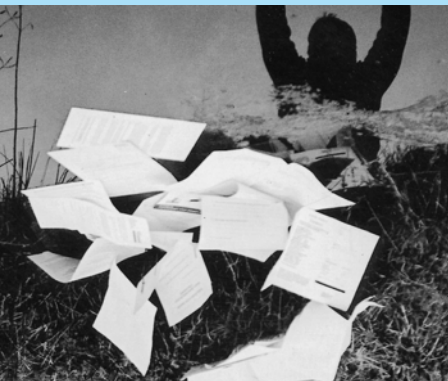
Overview

Decisions related to data preparation and analysis should not take place after data have been collected. Before the raw data contained in the questionnaires can be subjected to statistical analysis, they must be converted into a form suitable for analysis. The suitable form and the means of analysis should be considered as a research design is developed. This ensures that the output of the analyses will satisfy the research objectives set for a particular project.

The care exercised in the data preparation phase has a direct effect upon the quality of statistical results and ultimately the support offered to marketing decision-makers. Paying inadequate attention to data preparation can seriously compromise statistical results, leading to biased findings and incorrect interpretation.

This chapter describes the data collection process, which begins with checking the questionnaires for completeness. Then we discuss the editing of data and provide guidelines for handling illegible, incomplete, inconsistent, ambiguous or otherwise unsatisfactory responses. We also describe coding, transcribing and data cleaning, emphasising the treatment of missing responses and statistical adjustment of data. We discuss the selection of a data analysis strategy and classify statistical techniques. The intra-cultural, pan-cultural and cross-cultural approaches to data analysis in international marketing research are explained. Finally, the ethical issues related to data processing are identified with emphasis on discarding of unsatisfactory responses, violation of the assumptions underlying the data analysis techniques, and evaluation and interpretation of results.

We begin with an illustration of the data preparation process in the GlobalCash Project.

example
GlobalCash Project

Data preparation

In the GlobalCash Project, the data were obtained by postal questionnaires. As the questionnaire was developed and finalised a preliminary plan was drawn up of how the findings could be analysed. The questionnaires were edited by a supervisor as they were being returned from individual European countries by the respective business schools involved in the project. The questionnaires were checked for incomplete, inconsistent and ambiguous responses. Questionnaires with problematic responses were queried with the respective business schools (who kept copies of the original questionnaires). In some circumstances, the business schools were asked to re-contact the respondents to clarify certain issues. Twenty questionnaires were discarded because the proportion of unsatisfactory responses was large. This resulted in a final sample size of 1,075.

A codebook was developed for coding the questionnaires. Coding was extremely difficult given the large number of banks and banking software companies that operate throughout Europe, the number of different names these banks may be known by, and the need for translation in certain open-ended questions. The data were transcribed by being directly keyed into a survey analysis package. The SNAP software used for data entry and analysis has a built-in error-check that identifies out-of-range responses. About 10% of the data were verified for other data entry errors. The data were cleaned by identifying logically inconsistent responses. Most of the rating information was obtained using five-point scales, so responses of 0, 6 and 7 were considered out of range and a code of 9 was assigned to missing responses. If an out-of-range response was keyed in, the SNAP software package did not allow any continuation of data entry; an audible warning was made. In statistically adjusting the data, dummy variables were created for the categorical variables. New variables that were composites of original variables were also created. Finally, a data analysis strategy was developed. ■

The GlobalCash example describes the various phases of the data preparation process. Note that the process is initiated while the fieldwork is still in progress. A systematic description of the data preparation process follows.

The data preparation process

The data preparation process is shown in Figure 17.1. The entire process is guided by the preliminary plan of data analysis that was formulated in the research design phase. The first step is to check for acceptable questionnaires. This is followed by **editing**, **coding** and transcribing the data. The data are cleaned and a treatment for missing responses is prescribed. Often, after the stage of sample validation, statistical adjustment of the data may be necessary to make them representative of the population of interest. The researcher should then select an appropriate data analysis strategy. The final data analysis strategy differs from the preliminary plan of data analysis due to the information and insights gained since the preliminary plan was formulated. Data preparation should begin as soon as the first batch of questionnaires is received from the field, while the fieldwork is still going on. Thus, if any problems are detected, the fieldwork can be modified to incorporate corrective action.

Editing
A review of the questionnaires with the objective of increasing accuracy and precision.

Coding
Assigning a code to represent a specific response to a specific question along with the data record and column position that the code will occupy.

Checking the questionnaire

The initial step in questionnaire checking involves reviewing all questionnaires for completeness and interviewing or completion quality. Often these checks are made while fieldwork is still under way. If the fieldwork was contracted to a data collection agency, the researcher should make an independent check after it is over. A questionnaire returned from the field may be unacceptable for several reasons:

1 Parts of the questionnaire may be incomplete.
2 The pattern of responses may indicate that the respondent did not understand or follow the instructions. For example, filter questions may not have been followed.
3 The responses show little variance. For example, a respondent has ticked only 4s on a series of seven-point rating scales.
4 The returned questionnaire is physically incomplete: one or more pages is missing.
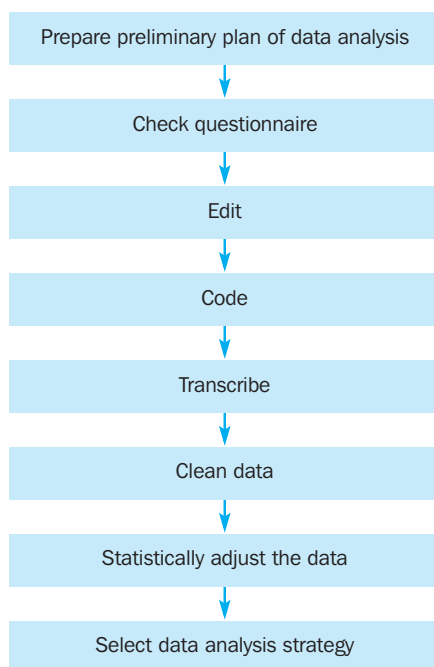
Figure 17.1
Data preparation process

Prepare preliminary plan of data analysis
↓
Check questionnaire
↓
Edit
↓
Code
↓
Transcribe
↓
Clean data
↓
Statistically adjust the data
↓
Select data analysis strategy

5 The questionnaire is received after the pre-established cut-off date.

6 The questionnaire is answered by someone who does not qualify for participation.

If quotas or cell group sizes have been imposed, the acceptable questionnaires should be classified and counted accordingly. Any problems in meeting the sampling requirements should be identified, and corrective action, such as conducting additional interviews in the under-represented cells, should be taken where this is possible, before the data are edited.

Editing

Editing is the review of the questionnaires with the objective of increasing accuracy and precision. It consists of screening questionnaires to identify illegible, incomplete, inconsistent or ambiguous responses. Responses may be illegible if they have been poorly recorded. This is particularly common in questionnaires with a large number of unstructured questions. The data must be legible if they are to be properly coded. Likewise, questionnaires may be incomplete to varying degrees. A few or many questions may be unanswered.

At this stage, the researcher makes a preliminary check for consistency. Certain obvious inconsistencies can be easily detected. For example, a respondent may have answered a whole series of questions relating to their perceptions of a particular bank, yet in other questions may have indicated that they have not used that particular bank or even heard of it.

Responses to unstructured questions may be ambiguous and difficult to interpret clearly. The answer may be abbreviated, or some ambiguous words may have been used. For structured questions, more than one response may be marked for a question designed to elicit a single response. Suppose that a respondent circles 2 and 3 on a five-point rating scale. Does this mean that 2.5 was intended? To complicate matters further, the coding procedure may allow for only a single-digit response.

Treatment of unsatisfactory responses

Unsatisfactory responses are commonly handled by returning to the field to get better data, assigning missing values, and discarding unsatisfactory respondents.

Returning to the field. Questionnaires with unsatisfactory responses may be returned to the field, where the interviewers re-contact the respondents. This approach is particularly attractive for business and industrial marketing surveys, where the sample sizes are small and the respondents are easily identifiable. The data obtained the second time, however, may be different from those obtained during the original survey. These differences may be attributed to changes over time or differences in the mode of questionnaire administration (e.g. telephone versus in-person interview).

Assigning missing values. If returning the questionnaires to the field is not feasible, the editor may assign missing values to unsatisfactory responses. This approach may be desirable if (1) the number of respondents with unsatisfactory responses is small, (2) the proportion of unsatisfactory responses for each of these respondents is small, or (3) the variables with unsatisfactory responses are not the key variables.

Discarding unsatisfactory respondents. In another approach, the respondents with unsatisfactory responses are simply discarded. This approach may have merit when (1) the proportion of unsatisfactory respondents is small (less than 10%); (2) the sample size is large; (3) the unsatisfactory respondents do not differ from sat-

Care must be taken in discarding questionnaires from unsatisfactory respondents.

isfactory respondents in obvious ways (e.g. demographics, product usage characteristics); (4) the proportion of unsatisfactory responses for each of these respondents is large; or (5) responses on key variables are missing. Unsatisfactory respondents may differ from satisfactory respondents in systematic ways, however, and the decision to designate a respondent as unsatisfactory may be subjective. Both these factors bias the results. If the researcher decides to discard unsatisfactory respondents, the procedure adopted to identify these respondents and their number should be reported, as in the following example.

e x a m p l e

Declaring 'discards'[1]

In a cross-cultural survey of marketing managers from English-speaking African countries, questionnaires were posted to 565 firms. A total of 192 completed questionnaires were returned, of which four were discarded because respondents suggested that they were not in charge of overall marketing decisions. The decision to discard the four questionnaires was based on the consideration that the sample size was sufficiently large and the proportion of unsatisfactory respondents was small. ■

Coding

Many questionnaire design and data entry software packages code data automatically. Examples of the options available will be presented in the Internet and computer applications section and on the Companion Website. Learning how to use such packages or even using spreadsheet packages means that the process of coding is now a much simpler task for the marketing researcher. Many of the principles of coding are based on the days of data processing using 'punched cards' or even, much more recently, DOS files. Whilst there may be many data analysts who could present coherent cases for the use of original forms of data entry, the greater majority of researchers enjoy the benefits of a simpler, speedier and less error-prone form of data entry, using proprietary software packages such as SNAP on **www.mercator.co.uk** or Keypoint2 on **www.camsp.com**. The nature and importance of coding for qualitative

data was introduced in Chapters 6 and 9. For quantitative data, which can include coping with open-ended responses or responses to 'Other – Please State…', it is still important to understand the principles of coding, as reference to the process is made by so many in the marketing research industry. The examples of coding presented are based on the original forms of conducting the GlobalCash survey in 1996 and 1998. In the 2000 and 2002 surveys, the process of coding was completely automated using the SNAP package.

Coding means assigning a code, usually a number, to each possible answer to each question. For example, a question on the gender of respondents may be assigned a code of 1 for females and 2 for males. For every individual question in a questionnaire, the researcher decides which codes should be assigned to all its possible answers.

If the question posed has only two possible answers, the codes assigned of 1 or 2 take up one digit space. If the question posed had 300 possible answers such as '*Which European bank do you conduct most business with?*' the possible answers and assigned codes of 1 to 300 would take up three digit spaces. The reason for focusing upon the digit spaces required for any particular question relates to recording the answers from individual questionnaire respondents in 'flat ASCII files'. Such files were typically 80 columns wide. The columns would be set out into 'fields', i.e. assigned columns that relate to specific questions. Thus the task for the researcher after assigning codes to individual question responses was to set out a consecutive series of fields or columns. These fields would represent where the answers to particular questions would be positioned in the ASCII file. In each row of a computer file would be the coded responses from individual questionnaire respondents. Each row is termed a 'record', i.e. all the fields that make up the response from one respondent. All the attitudinal, behavioural, demographic and other classification characteristics of a respondent may be contained in a single record.

Table 17.1 shows an extract from the GlobalCash questionnaire and Table 17.2 illustrates the answers to these questions from a selection of respondents as set out in

Table 17.1 Classification questions from the GlobalCash survey

Question 1 – Do you work in a:

| ☐ Subsidiary | ☐ Country treasury | ☐ Regional treasury | ☐ European treasury | ☐ Group treasury (5) |

Question 2 – What is the main industry in which your group operates?

1 ☐ Oil/gas	9 ☐ Retail/wholesale	(6)
2 ☐ Manufacturing	10 ☐ Services	
3 ☐ Metals	11 ☐ Transport	
4 ☐ Food/beverages/tobacco	12 ☐ Telecommunications	
5 ☐ Chemicals/pharmaceuticals	13 ☐ Non-banking financial institution	
6 ☐ Electronics	14 ☐ Insurance and pensions	
7 ☐ Utilities	15 ☐ Other (please state) …………	
8 ☐ Construction/engineering		

Question 3 – What are your group's approximate annual worldwide sales? (€m) _____ (8)

Question 4 – In how many European countries does your group operate?

| ☐ 1 | ☐ 2 to 5 | ☐ 6 to 10 | ☐ 11 to 15 | ☐ Over 15 (15) |

Question 5 – What position do you hold in your company?

1 ☐ Cash manager	5 ☐ Accountant	(16)
2 ☐ Treasurer	6 ☐ Chief Executive	
3 ☐ Assistant treasurer	7 ☐ Other (please state)	
4 ☐ Finance manager/director		

Table 17.2 Illustrative computer file held on a flat ASCII file

				Fields		
Records	1–4	5	6–7	8–14	15	16 ... n
Record #1	0001	1	12	0001950	3	2
Record #11	0011	3	03	0000760	2	2
Record #21	0021	5	10	0002650	5	1
Record #1011	1011	3	08	0009670	4	9

codes, fields and records. The classification questions set out on Table 17.1 were placed at the end of the questionnaire but, as they represent simple structured questions, they are assumed to be at the start of the questionnaire for this illustration.

Question 1 has five possible answers which are coded 1 to 5, that take up one digit space. Question 2 has 15 possible answers which are coded 1 to 15, that take up two digit spaces. Note that, with question 2, code 15 represents 'Other, Please State'. If the researcher finds that there are other categories emerging that should have been included in the list of answers, these can be given individual codes. As two digit spaces have been allocated to cope with 15 possible answers, up to 99, i.e. 84 additional, categories of 'other' can be catered for. Question 3 is open-ended and requires a number to represent the worldwide sales of the respondent in millions of euros. This question can be pre-coded into bands, e.g. Code 1 = 0 to 50, Code 2 = 51 to 100, Code 3 = 101 to 250, etc. Alternatively, the data can be entered in its raw format, which allows far more precision in later analyses and still retains the ability to put respondents into bands at a later date. By entering the response in a raw data format, eight digit spaces have been allowed, compared with the one or perhaps two had it been pre-coded. Question 4 has the same structure as Question 1, taking up one digit space. Question 5 is similar to Question 2 in that it has an 'Other, Please State' option. Note that at the end of each question is a small number in parentheses. These numbers represent the first field positions of each question as illustrated in Table 17.2.

In Table 17.2, the columns represent the fields and the rows represent the records of each respondent. The field space 1 to 4 is used to record an assigned number to each respondent. Should the record require more than 80 digit spaces, it would wrap around on to a second, third or fourth line, and so on.

Table 17.3 illustrates how the same data may be entered using a spreadsheet. Each row represents an individual respondent, each column representing the fields required to hold the response to an individual question. Note that there is a column that identifies a specific number attached to each record. Many survey analysis packages record a unique ID for each record so that, as the answers to an individual questionnaire are entered, the ID is automatically updated. However, if a unique ID is attached to each questionnaire before it is sent out (for example, in a postal survey), the ID may be entered as a distinct field (see column A).

Table 17.3 Example of computer file held on a spreadsheet program

	Individual fields in columns					
	A	B	C	D	E	F
Records in rows	ID	Question 1	Question 2	Question 3	Question 4	Question 5
1	1	1	12	1950	3	2
11	11	3	3	760	2	2
21	21	5	10	2650	5	1
1011	1011	3	8	9670	4	9

Coding is still required to identify the individual responses to individual questions. Spreadsheets are normally wide enough to allow an individual record to be recorded on one line, and they can be set up so that whoever is entering the data can clearly keep track of which questions relate to which columns. Spreadsheets can be used as a format to analyse data in a wide variety of data analysis packages and so are very versatile. They do, however, have shortcomings. The next paragraph will go on to illustrate these.

In many surveys, multiple choice questions are widely used. An example of a multiple choice question is shown in Table 17.4, a question from the GlobalCash survey which examines respondent's plans for changes over the next two years.

Table 17.4 Multiple choice question from the GlobalCash survey

Question 51 – During the next two years, which of the following changes does your company plan to make? (please tick appropriate boxes)

1	✓	Treasury function to be more centralised
2		Make more use of treasury vehicles
3		Move more towards an in-house bank
4		Out-source certain treasury functions – to banks
5	✓	Put domestic banking out to tender
6	✓	Make greater use of electronic banking
7		Treasury function to be more automated
8		Integrate more services under the treasury function
9	✓	Reduce number of people in the treasury function

In essence, each of the options presented in Question 51 is an individual 'yes' or 'no' question. In the example shown, the respondent has replied 'yes' to the first, fifth, sixth and ninth variables. Using a spreadsheet, this question would be coded as shown in Table 17.5 where the response in Table 17.4 is represented as 'Record 1'. The 'ticks' above have been coded as a '1' to represent 'yes' and '0' as 'no'.

Table 17.5 GlobalCash, Question 51 spreadsheet presentation

	Individual fields in columns								
	Question 51								
	AJ	AK	AL	AM	AN	AO	AP	AQ	AU
Records	Q51(1)	Q51(2)	Q51(3)	Q51(4)	Q51(5)	Q51(6)	Q51(7)	Q51(8)	Q51(9)
1	1	0	0	0	1	1	0	0	1
2	1	1	0	0	0	1	1	0	0
3	0	0	1	1	1	0	0	0	0
4	1	1	0	0	1	0	1	1	1
... n	0	1	1	0	0	1	1	0	0

Entering data for multiple choice questions is a simple task on a spreadsheet, provided that there are not so many of these question types in a survey. In the GlobalCash questionnaire there was one question related to which banks were used in 25 European countries. The multiple choice question of 250 different banks multiplied by 25 countries would have meant a spreadsheet of 6,250 columns, just for the one question! If a respondent had indicated that they used ABN AMRO in Norway, this would mean finding the precise column to enter a '1' and then 6,249 '0s'. This is a lengthy and potentially error-prone task, but fortunately cases like this are rare. This is where proprietary questionnaire design and survey packages really hold many advan-

tages, i.e. they make the data entry task very simple and check for errors. Again, the use of proprietary packages will be outlined at the end of this chapter.

Codebook

Whether, the researcher uses DOS-based systems, spreadsheets or a proprietary package, a summary of the whole questionnaire, showing the position of the fields and the key to all the codes, should be produced. Such a summary is called a codebook. Table 17.6 shows an extract from the GlobalCash codebook. The codebook shown is based upon using a spreadsheet to enter the data. Depending upon which type of data entry is used, the codebook style will change, but in essence the type of information recorded is the same.

Codebook
A book containing coding instructions and the necessary information about the questions and potential answers in a survey.

A **codebook** contains instructions and the necessary information about the questions and potential answers in a survey. A codebook guides the 'coders' in their work and helps the researcher identify and locate the questions properly. Even if the questionnaire has been pre-coded, it is helpful to prepare a formal codebook. As illustrated in Table 17.6, a codebook generally contains the following information: (1) column identifier, (2) question name, (3) question number, and (4) coding instructions.

Table 17.6 **Extract from the GlobalCash survey codebook**

Column identifier	Question name	Question number	Coding instructions
A	Respondent ID		Enter handwritten number from top right-hand corner of the questionnaire.
B	Treasury type	Q. 1	Subsidiary = '1' Country treasury = '2' Regional treasury = '3' European treasury = '4' Group treasury = '5' Missing value = '9'
C	Industry	Q. 2	Enter number as seen alongside ticked box. Missing value = '99'
D	Worldwide sales	Q. 3	Enter actual value in € millions
E	Countries operated in	Q. 4	1 = '1' 2 to 5 = '2' 6 to 10 = '3' 11 to 15 = '4' over 15 = '5' Missing value = '9'
F	Position	Q. 5	Treasurer = '1' Finance director = '2' Accountant = '3' Chief executive = '4' Cash manager = '5' Other = '6' Missing value = '9'

Coding open-ended questions

The coding of structured questions, be they single or multiple choice, is relatively simple because the response options are predetermined. The researcher assigns a code for each response to each question and specifies the appropriate field or column in which it will appear; this is termed 'pre-coding'.[2] The coding of unstructured or open-ended questions is more complex; this is termed 'post-coding'.[3] Respondents'

verbatim responses are recorded on the questionnaire. One option the researcher has is to go through all the completed questionnaires, list the verbatim responses and then develop and assign codes to these responses. Another option that is allowed on some data entry packages is to enter the verbatim responses directly on to the computer, allowing a print-off of the collective responses and codes to be assigned before all of the questionnaires have been entered. The coding process here is similar to the process of assigning codes in the analysis of qualitative data as described in Chapter 9. The verbatim responses to 1,000 questionnaires may generate 1,000 different answers. The words may be different but the essence of the response may mean that 20 issues have been addressed. The researcher decides what those 20 issues are, names the issues and assigns codes from 1 to 20, and then goes through all the 1,000 questionnaires to enter the code alongside the verbatim response.

The following guidelines are suggested for coding unstructured questions and questionnaires in general.[4] Category codes should be mutually exclusive and collectively exhaustive. Categories are mutually exclusive if each response fits into one and only one category code. Categories should not overlap. Categories are collectively exhaustive if every response fits into one of the assigned category codes. This can be achieved by adding an additional category code of 'other' or 'none of the above'. An absolute maximum of 10% of responses should fall into the 'other' category; the researcher should strive to assign all responses into meaningful categories.

Category codes should be assigned for critical issues even if no one has mentioned them. It may be important to know that no one has mentioned a particular response. For example, a bank may be concerned about their new Web page design. In a question 'How did you learn about Credit Card X', the Web should be included as a distinct category, even if no respondents gave this as an answer.

Transcribing

Transcribing data involves keying the coded data from the collected questionnaires into computers. If the data have been collected via the Internet, CATI or CAPI, this step is unnecessary because the data are entered directly into the computer as they are collected. Besides the direct keying of data, they can be transferred by using mark sense forms, optical scanning or computerised sensory analysis (see Figure 17.2). Mark sense forms require responses to be recorded in a pre-designated area coded for
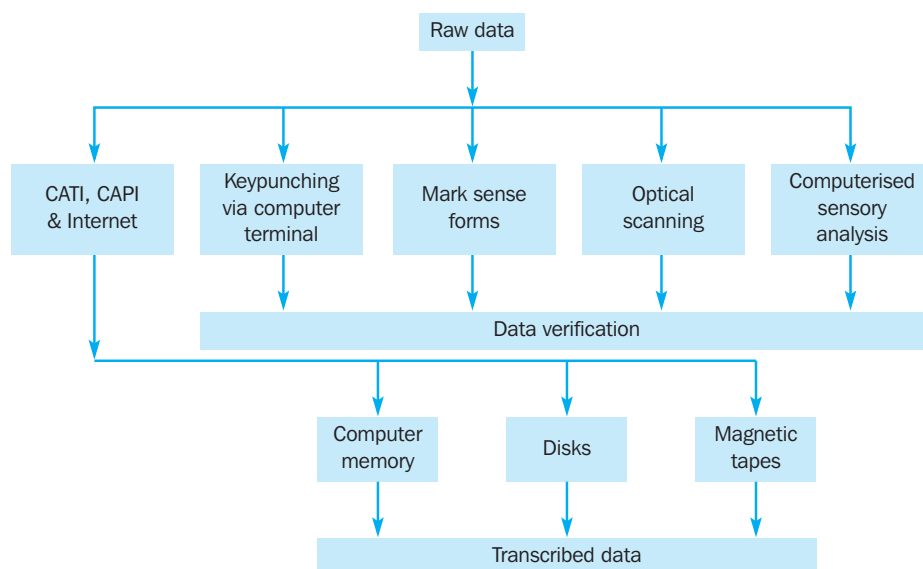
Figure 17.2
Data transcription

that response, and the data can then be read by a machine. Optical scanning involves direct machine reading of the codes and simultaneous transcription. A familiar example of optical scanning is the transcription of universal product code (UPC) data, scanned at supermarket checkout counters. Technological advances have resulted in computerised sensory analysis systems, which automate the data collection process. The questions appear on a computerised gridpad, and responses are recorded directly into the computer using a sensing device.

Except for CATI and CAPI, an original record exists which can be compared with what was either automatically read or keyed. Errors can occur in an automatic read or as data are keyed and it is necessary to verify the dataset, or at least a portion of it, for these errors.

A second operator re-punches the data from the coded questionnaires. The transcribed data from the two operators are compared record by record. Any discrepancy between the two sets of transcribed data is investigated to identify and correct for data keyed in error. Verification of the entire data set will double the time and cost of data transcription. Given the time and cost constraints, and that experienced operators who key data are quite accurate, it is sufficient to verify 10–25% of the data. With automatically read data, the completed data set that has been read can be compared with original records. Again, a percentage may be selected and checks made to see what may have caused differences between the original record and the read data (e.g. respondents entering two ticks when only one was requested).

When CATI, CAPI or the Internet are employed, data are verified as they are collected. In the case of inadmissible responses, the computer will prompt the interviewer or respondent. In the case of admissible responses, the interviewer or the respondent can see the recorded response on the screen and verify it before proceeding.

The selection of a data transcription method is guided by the type of interviewing method used and the availability of equipment. If CATI, CAPI or the Internet are used, the data are entered directly into the computer. Keypunching via a computer terminal is most frequently used for ordinary telephone, in-home and street interviewing and traditional mail interviews. The use of computerised sensory analysis systems in personal interviews is increasing with the growing use of handheld computers, however. Optical scanning can be used in structured and repetitive surveys, and mark sense forms are used in special cases.[5]

Cleaning the data

Data cleaning
Thorough and extensive checks for consistency and treatment of missing responses.

Data cleaning includes consistency checks and treatment of missing responses. Even though preliminary consistency checks have been made during editing, the checks at this stage are more thorough and extensive, because they are made by computer.

Consistency checks

Consistency checks
A part of the data cleaning process that identifies data that are out of range, logically inconsistent, or have extreme values. Data with values not defined by the coding scheme are inadmissible.

Consistency checks identify data that are out of range or logically inconsistent or have extreme values. Out-of-range data values are inadmissible and must be corrected. For example, respondents have been asked to express their degree of agreement with a series of lifestyle statements on a 1 to 5 scale. Assuming that 9 has been designated for missing values, data values of 0, 6, 7 and 8 are out of range. Computer packages can be programmed to identify out-of-range values for each variable and will not progress to another variable within a record until a value in the set range is entered. Other packages can be programmed to print out the respondent code, variable code, variable name, record number, column number and out-of-range value[6]. This makes it easy to check each variable systematically for out-of-range values. The correct responses can be determined by going back to the edited and coded questionnaire.

Responses can be logically inconsistent in various ways. For example, a respondent may indicate that he charges long-distance calls to a calling card from a credit card company, although he does not have such a credit card. Or a respondent reports both unfamiliarity with and frequent usage of the same product. The necessary information (respondent code, variable code, variable name, record number, column number and inconsistent values) can be printed to locate these responses and to take corrective action.

Finally, extreme values should be closely examined. Not all extreme values result from errors, but they may point to problems with the data. For example, in the GlobalCash survey, companies were asked how many banks they use for domestic business. Certain Italian companies recorded 70 or more banks when the mean for all respondents was around six. In these circumstances the extreme values can be identified and the actual figure validated in many cases by re-contacting the respondent or examining secondary data sources. In the GlobalCash survey, contact was made with the respondents who recorded 70 or more banks, and in all cases the figures were correct.

Treatment of missing responses

Missing responses

Values of a variable that are unknown because the respondents concerned provided ambiguous answers to the question or because their answers were not properly recorded.

Missing responses represent values of a variable that are unknown either because respondents provided ambiguous answers or because their answers were not properly recorded. Treatment of missing responses poses problems, particularly if the proportion of missing responses is more than 10%. The following options are available for the treatment of missing responses.[7]

Substitute a neutral value. A neutral value, typically the mean response to the variable, is substituted for the missing responses. Thus, the mean of the variable remains unchanged, and other statistics such as correlations are not affected much. Although this approach has some merit, the logic of substituting a mean value (say 4) for respondents who, if they had answered, might have used either high ratings (6 or 7) or low ratings (1 or 2) is questionable.[8]

Substitute an imputed response. The respondents' pattern of responses to other questions is used to impute or calculate a suitable response to the missing questions. The researcher attempts to infer from the available data the responses the individuals would have given if they had answered the questions. This can be done statistically by determining the relationship of the variable in question to other variables based on the available data. For example, product usage could be related to household size for respondents who have provided data on both variables. Given that respondent's household size, the missing product usage response for a respondent could then be calculated. This approach, however, requires considerable effort and can introduce serious bias. Sophisticated statistical procedures have been developed to calculate imputed values for missing responses.[9]

Casewise deletion

A method for handling missing responses in which cases or respondents with any missing responses are discarded from the analysis.

Pairwise deletion

A method for handling missing responses in which all cases or respondents with any missing resopnses are not automatically discarded; rather, for each calculation, only the cases or respondents with complete responses are considered.

Casewise deletion. In **casewise deletion**, cases or respondents with any missing responses are discarded from the analysis. Because many respondents may have some missing responses, this approach could result in a small sample. Throwing away large amounts of data is undesirable because it is costly and time-consuming to collect data. Furthermore, respondents with missing responses could differ from respondents with complete responses in systematic ways. If so, casewise deletion could seriously bias the results.

Pairwise deletion. In **pairwise deletion**, instead of discarding all cases with any missing responses, the researcher uses only the cases or respondents with complete responses for each calculation. As a result, different calculations in an analysis may be

based on different sample sizes. This procedure may be appropriate when (1) the sample size is large, (2) there are few missing responses, and (3) the variables are not highly related. Yet this procedure can produce unappealing or even infeasible results.

The different procedures for the treatment of missing responses may yield different results, particularly when the responses are not missing at random and the variables are related. Hence, missing responses should be kept to a minimum. The researcher should carefully consider the implications of the various procedures before selecting a particular method for the treatment of non-response.

Statistically adjusting the data

Procedures for statistically adjusting the data consist of weighting, variable re-specification and scale transformation. These adjustments are not always necessary but can enhance the quality of data analysis.

Weighting

Weighting
A statistical procedure that attempts to account for non-response by assigning differential weights to the data depending on the response rates.

In **weighting**, each case or respondent in the database is assigned a weight to reflect its importance relative to other cases or respondents. The value 1.0 represents the unweighted case. The effect of weighting is to increase or decrease the number of cases in the sample that possess certain characteristics. (See Chapter 15, which discussed the use of weighting to adjust for non-response bias.)

Weighting is most widely used to make the sample data more representative of a target population on specific characteristics. For example, it may be used to give greater importance to cases or respondents with higher-quality data. Yet another use of weighting is to adjust the sample so that greater importance is attached to respondents with certain characteristics. If a study is conducted to determine what modifications should be made to an existing product, the researcher might want to attach greater weight to the opinions of heavy users of the product. This could be accomplished by assigning weights of 3.0 to heavy users, 2.0 to medium users, and 1.0 to light users and non-users. Because it destroys the self-weighting nature of the

In weighting, each respondent is assigned a weight.

sample design, weighting should be applied with caution. If used, the weighting procedure should be documented and made a part of the project report.[10]

Determining the weight of community centre users

A mail survey was conducted in the Scottish city of Edinburgh to determine the patronage of a community centre. The resulting sample composition differed in age structure from the area population distribution as compiled from recent census data. Therefore, the sample was weighted to make it representative in terms of age structure. The weights applied were determined by dividing the population percentage by the corresponding sample percentage. The distribution of age structure for the sample and population, as well as the weights applied, are given in the following table.

Age group percentage	Sample percentage	Population	Weight
13 to 18	4.32	6.13	1.42
19 to 24	5.89	7.45	1.26
25 to 34	12.23	13.98	1.14
35 to 44	17.54	17.68	1.01
45 to 54	14.66	15.59	1.06
55 to 64	13.88	13.65	0.98
65 to 74	15.67	13.65	0.87
75 plus	15.81	11.87	0.75
Totals	100.00	100.00	

Age groups under-represented in the sample received higher weights, whereas over-represented age groups received lower weights. Thus, the data for a respondent aged 13 to 18 would be overweighted by multiplying by 1.42, whereas the data for a respondent aged 75 or over would be underweighted by multiplying by 0.75. ■

Variable re-specification

Variable re-specification
The transformation of data to create new variables or the modification of existing variables so that they are more consistent with the objectives of the study.

Variable re-specification involves the transformation of data to create new variables or to modify existing variables. The purpose of re-specification is to create variables that are consistent with the objectives of the study. For example, suppose that the original variable was product usage, with 10 response categories. These might be collapsed into four categories: heavy, medium, light and non-user. Or the researcher may create new variables that are composites of several other variables. For example, in a survey where one question establishes 'gender' and another 'age', a new variable may combine these. The new variable could have four values of 'young women', 'older women', 'young men' and 'older men'. In the GlobalCash study, a composite variable was created based upon how many European countries a company operates in and the size of global sales. Likewise, one may take the ratio of variables. If the amount of purchases at a clothes shop (X_1) and the amount of purchases where a credit card was used (X_2) have been measured, the proportion of purchases charged to a credit card can be a new variable, created by taking the ratio of the two (X_2/X_1). Other re-specifications of variables include square root and log transformations, which are often applied to improve the fit of the model being estimated.

Dummy variables
A re-specification procedure using variables that take on only two values, usually 0 or 1.

An important re-specification procedure involves the use of **dummy variables** for re-specifying categorical variables. Dummy variables are also called binary, dichotomous, instrumental or qualitative variables. They are variables that may take on only two values, such as 0 or 1. The general rule is that to re-specify a categorical variable

with K categories, $K - 1$ dummy variables are needed. The reason for having $K - 1$, rather than K, dummy variables is that only $K - 1$ categories are independent. Given the sample data, information about the Kth category can be derived from information about the other $K - 1$ categories. Consider gender, a variable having two categories. Only one dummy variable is needed. Information on the number or percentage of males in the sample can be readily derived from the number or percentage of females. The following example further illustrates the concept of dummy variables.

e x a m p l e

'Frozen' consumers treated as dummies

In a survey of consumer preferences for frozen foods, the respondents were classified as heavy users, medium users, light users and non-users, and they were originally assigned codes of 4, 3, 2 and 1, respectively. This coding was not meaningful for several statistical analyses. To conduct these analyses, product usage was represented by three dummy variables, X_1, X_2 and X_3, as shown.

Product usage category	Original variable code	Dummy variable code		
		X_1	X_2	X_3
Non-users	1	1	0	0
Light users	2	0	1	0
Medium users	3	0	0	1
Heavy users	4	0	0	0

Note that $X_1 = 1$ for non-users and 0 for all others. Likewise, $X_2 = 1$ for light users and 0 for all others, and $X_3 = 1$ for medium users and 0 for all others. In analysing the data, X_1, X_2, and X_3 are used to represent all user/non-user groups. ■

Scale transformation

Scale transformation
A manipulation of scale values to ensure compatibility with other scales or otherwise to make the data suitable for analysis.

Scale transformation involves a manipulation of scale values to ensure comparability with other scales or otherwise to make the data suitable for analysis. Frequently, different scales are employed for measuring different variables. For example, image variables may be measured on a seven-point semantic differential scale, attitude variables on a continuous rating scale, and lifestyle variables on a five-point Likert scale. Therefore, it would not be meaningful to make comparisons across the measurement scales for any respondent. To compare attitudinal scores with lifestyle or image scores, it would be necessary to transform the various scales. Even if the same scale is employed for all the variables, different respondents may use the scale differently. For example, some respondents consistently use the upper end of a rating scale whereas others consistently use the lower end. These differences can be corrected by appropriately transforming the data.

e x a m p l e

Health care services: transforming consumers[11]

In a study examining preference segmentation of health care services, respondents were asked to rate the importance of 18 factors affecting preferences for hospitals on a three-point scale (very, somewhat, or not important). Before analysing the data, each individual's ratings were transformed. For each individual, preference responses were averaged across all 18 items. Then this mean $\bar{X}$ was subtracted from each item rating X_i, and a constant C was added to the difference. Thus, the transformed data, X_t, were obtained by

$$X_t = X_i - \bar{X} + C$$

Subtraction of the mean value corrected for uneven use of the importance scale. The constant C was added to make all the transformed values positive, since negative importance ratings are not meaningful conceptually. This transformation was desirable because some respondents, especially those with low incomes, had rated almost all the preference items as

very important. Others, high-income respondents in particular, had assigned the very important rating to only a few preference items. Thus, subtraction of the mean value provided a more accurate idea of the relative importance of the factors. ■

Standardisation
The process of correcting data to reduce them to the same scale by subtracting the sample mean and dividing by the standard deviation.

In this example, the scale transformation is corrected only for the mean response. A more common transformation procedure is **standardisation**. To standardise a scale X_i, we first subtract the mean, $\bar{X}$, from each score and then divide by the standard deviation, s_x. Thus, the standardised scale will have a mean of zero and a standard deviation of 1. This is essentially the same as the calculation of z scores (see Chapter 15). Standardisation allows the researcher to compare variables that have been measured using different types of scales.[12] Mathematically, standardised scores, z_i, may be obtained as

$$z_i = \frac{(X_i - \bar{X})}{s_x}$$

Selecting a data analysis strategy

The process of selecting a data analysis strategy is described in Figure 17.3. The selection of a data analysis strategy should be based on the earlier steps of the marketing research process, known characteristics of the data, properties of statistical techniques, and the background and philosophy of the researcher.

Data analysis is not an end in itself. Its purpose is to produce information that will help address the problem at hand. The selection of a data analysis strategy must begin with a consideration of the earlier steps in the process: problem definition (step 1), development of an approach (step 2), and research design (step 3). The preliminary plan of data analysis prepared as part of the research design should be used as a springboard. Changes may be necessary in the light of additional information generated in subsequent stages of the research process.

The next step is to consider the known characteristics of the data. The measurement scales used exert a strong influence on the choice of statistical techniques (see Chapter 12). In addition, the research design may favour certain techniques. For example, analysis of variance (see Chapter 19) is suited for analysing experimental data from causal designs. The insights into the data obtained during data preparation can be valuable for selecting a strategy for analysis.

It is also important to take into account the properties of the statistical techniques, particularly their purpose and underlying assumptions. Some statistical techniques
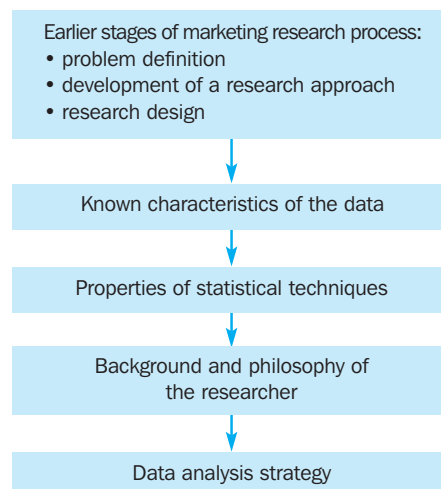
Figure 17.3
Selecting a data analysis strategy

Earlier stages of marketing research process:
• problem definition
• development of a research approach
• research design

↓

Known characteristics of the data

↓

Properties of statistical techniques

↓

Background and philosophy of the researcher

↓

Data analysis strategy

are appropriate for examining differences in variables, others for assessing the magnitudes of the relationships between variables, and still others for making predictions. The techniques also involve different assumptions, and some techniques can withstand violations of the underlying assumptions better than others. A classification of statistical techniques is presented below.

Finally, the researcher's background and philosophy affect the choice of a data analysis strategy. The experienced, statistically trained researcher will employ a range of techniques, including advanced statistical methods. Researchers differ in their willingness to make assumptions about the variables and their underlying populations. Researchers who are conservative about making assumptions will limit their choice of techniques to distribution-free methods. In general, several techniques may be appropriate for analysing the data from a given project. We use the GlobalCash Project for illustration.

<table>
<tr><td>*e x a m p l e*
GlobalCash Project</td><td>

Data analysis strategy

As part of the analysis conducted in the GlobalCash Project, service quality was modelled in terms of the most important quality issues in domestic banking service delivery. The sample was split into halves. The respondents in each half were clustered on the basis of the importance of service quality characteristics. Statistical tests for clusters were conducted, and eight
</td></tr>
</table>
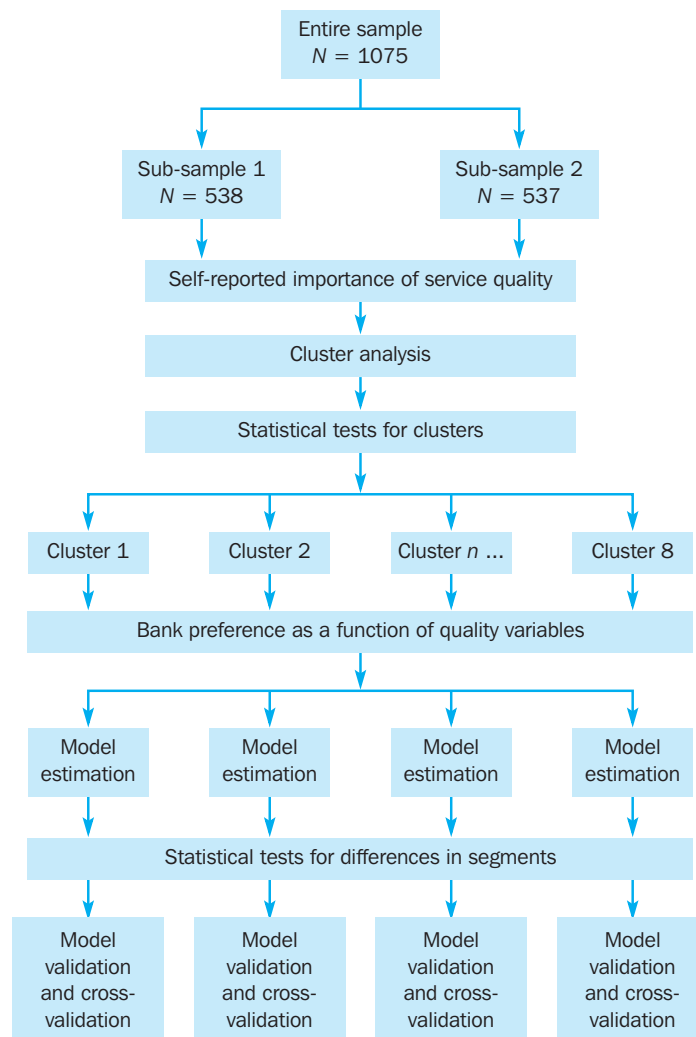
segments were identified. Service quality was modelled in terms of the evaluations of banks on the quality variables. The model was estimated separately for each segment. Differences between segment preference functions were statistically tested. Finally, model verification and cross-validation were conducted for each segment. The data analysis strategy adopted is depicted in the figure opposite.[13] ■

A classification of statistical techniques

Univariate techniques
Statistical techniques appropriate for analysing data when there is a single measurement of each element in the sample, or, if there are several measurements on each element, when each variable is analysed in isolation.

Multivariate techniques
Statistical techniques suitable for analysing data when there are two or more measurements on each element and the variables are analysed simultaneously. Multivariate techniques are concerned with the simultaneous relationships among two or more phenomena.

Metric data
Data that are interval or ratio in nature.

Non-metric data
Data derived from a nominal or ordinal scale.

Statistical techniques can be classified as univariate or multivariate. **Univariate techniques** are appropriate when there is a single measurement of each element in the sample or when there are several measurements of each element but each variable is analysed in isolation. **Multivariate techniques**, on the other hand, are suitable for analysing data when there are two or more measurements of each element and the variables are analysed simultaneously. Multivariate techniques are concerned with the simultaneous relationships among two or more phenomena. Multivariate techniques differ from univariate techniques in that they shift the focus away from the levels (averages) and distributions (variances) of the phenomena, concentrating instead on the degree of relationships (correlations or covariances) among these phenomena.[14] The univariate and multivariate techniques are described in detail in Chapters 18–24, but here we show how the various techniques relate to each other in an overall scheme of classification.

Univariate techniques can be further classified based on whether the data are metric or non-metric (as introduced in Chapter 12). **Metric data** are measured on an interval or ratio scale, whereas **non-metric data** are measured on a nominal or ordinal scale. For metric data, when there is only one sample, the z test and the t test can be used. When there are two or more independent samples, the z test and t test can be used for two samples, and one-way analysis of variance (one-way ANOVA) can be used for more than two samples. In the case of two or more related samples, the paired t test can be used. For non-metric data involving a single sample, frequency distribution, chi-square, Kolmogorov-Smirnov (K-S), runs, and binomial tests can be

Figure 17.4
A classification of univariate techniques

used. For two independent samples with non-metric data, the chi-square, Mann-Whitney, median, K-S and Kruskal-Wallis one-way analysis of variance (K-W ANOVA) can be used. In contrast, when there are two or more related samples, the sign, Wilcoxon, McNemar and chi-square tests should be used.

Multivariate statistical techniques can be classified as dependence techniques or interdependence techniques (see Figure 17.5). **Dependence techniques** are appropriate when one or more variables can be identified as dependent variables and the remaining ones as independent variables. When there is only one dependent variable, cross-tabulation, analysis of variance and covariance, multiple regression, two-group discriminant analysis and conjoint analysis can be used. If there is more than one dependent variable, however, the appropriate techniques are multivariate analysis of variance and covariance, canonical correlation and multiple discriminant analysis. In **interdependence techniques**, the variables are not classified as dependent or independent; rather, the whole set of interdependent relationships is examined. These techniques focus on either variable interdependence or inter-object similarity. The major technique for examining variable interdependence is factor analysis. Analysis of inter-object similarity can be conducted by cluster analysis and multidimensional scaling.[15]

Dependence techniques
Multivariate techniques appropriate when one or more of the variables can be identified as dependent variables and the remaining ones as independent variables.

Interdependence techniques
Multivariate statistical techniques that attempt to group data based on underlying similarity and thus allow for interpretation of the data structures. No distinction is made as to which variables are dependent and which are independent.

International marketing research

Before analysing the data, the researcher should ensure that the units of measurement are comparable across countries or cultural units. For example, the data may have to be adjusted to establish currency equivalents or metric equivalents. Furthermore, standardisation or normalisation of the data may be necessary to make meaningful comparisons and achieve consistent results.

example   **A worldwide scream for ice cream**[16]

Over half the sales of Häagen-Dazs, the US ice cream manufacturer, come from markets outside the United States. How have Häagen-Dazs achieved this situation? Marketing research conducted in several European countries (e.g. Britain, France and Germany) and several Asian countries (e.g. Japan, Singapore and Taiwan) revealed that consumers were

hungry for a high-quality ice cream with a high-quality image and were willing to pay a premium price for it. These consistent findings emerged after the price of ice cream in each country was standardised to have a mean of zero and a standard deviation of unity. Standardisation was desirable because the prices were specified in different local currencies and a common basis was needed for comparison across countries. Also, in each country, the premium price had to be defined in relation to the prices of competing brands. Standardisation accomplished both of these objectives.

Based on these findings, Häagen-Dazs first introduced the brand at a few high-end retailers; it then built company-owned stores in high-traffic areas; and finally it rolled into convenience stores and supermarkets. It maintained the premium quality brand name by starting first with a few high-end retailers. It also supplied free freezers to retailers. Hungry for quality products, consumers in the new markets paid double or triple the price of home brands. In the United States, Häagen-Dazs remains popular, although faced with intense competition and a health-conscious market. This added to the impetus to enter international markets. ■

Data analysis could be conducted at three levels: (1) individual, (2) within country or cultural unit, and (3) across countries or cultural units. Individual level analysis requires that the data from each respondent be analysed separately. For example, one might compute a correlation coefficient or run a regression analysis for each respondent. This means that enough data must be obtained from each individual to allow analysis at the individual level, which is often not feasible. Yet it has been argued that, in international marketing or cross-cultural research, the researcher should possess a sound knowledge of the consumer in each culture. This can best be accomplished by individual level analysis.[17]

Intra-cultural analysis
Within-country analysis of international data.

In within-country or cultural unit analysis, the data are analysed separately for each country or cultural unit. This is also referred to as **intra-cultural analysis**. This level of analysis is quite similar to that conducted in domestic marketing research. The objective is to gain an understanding of the relationships and patterns existing in each country or cultural unit.

Pan-cultural analysis
Across-countries analysis in which the data for all respondents from all the countries are pooled and analysed.

In across-countries analysis, the data from all the countries are analysed simultaneously. Two approaches to this method are possible. The data for all respondents from all the countries can be pooled and analysed. This is referred to as **pan-cultural analysis**. Alternatively, the data can be aggregated for each country, and then these aggregate statistics can be analysed. For example, one could compute means of variables for each country, and then compute correlations on these means. This is referred to as **cross-cultural analysis**. The objective of this level of analysis is to assess the comparability of findings from one country to another. The similarities as well as the differences between countries should be investigated. When examining differences, not only differences in means but also differences in variance and distribution should be assessed.

Cross-cultural analysis
A type of across-countries analysis in which the data could be aggregated for each country and these aggregate statistics analysed.

All the statistical techniques that have been discussed in this book can be applied to within-country or across-country analysis and, subject to the amount of data available, to individual-level analysis as well.[18]

Ethics in marketing research

New ethical issues can arise during the data preparation and analysis step of the marketing research process. While checking, editing, coding, transcribing and cleaning, researchers can get some idea about the quality of the data. Sometimes it is easy to identify respondents who did not take the questionnaire seriously or who otherwise provided data of questionable quality. Consider, for example, a respondent who ticks the 'neither agree nor disagree' response to all the 20 items measuring attitudes towards spectator sports. Decisions as to whether such respondents should be discarded or not included in analyses can raise ethical concerns. A good rule of thumb is to make such decisions during the data preparation phase before conducting any analysis.

In contrast, suppose that the researcher conducted the analysis without first attempting to identify unsatisfactory responses. The analysis, however, does not reveal the expected relationship; the analysis does not show that attitude towards spectator sports influences attendance at spectator sports. The researcher then decides to examine the quality of data obtained. In checking the questionnaires, a few respondents with unsatisfactory data are identified. These respondents are eliminated and the reduced data set is analysed to obtain the expected results. Discarding respondents after analysing the data raises ethical concerns, particularly if the report does not state that the initial analysis was inconclusive. Moreover, the procedure used to identify unsatisfactory respondents and the number of respondents discarded should be clearly disclosed, as in the following example.

example

Elimination of decision-makers unwilling to be ethical[19]

In a study of MBAs' responses to marketing ethics dilemmas, respondents were required to respond to 14 questions regarding ethically ambiguous scenarios by writing a simple sentence regarding what action they would take if they were the manager. The responses were then analysed to determine whether the respondent's answer was indicative of ethical behaviour. However, in the data preparation phase, six respondents out of the 561 total respondents were eliminated from further analysis because their responses indicated that they did not follow the directions which told them to state clearly their choice of action. This is an example of ethical editing of the data. The criterion for unsatisfactory responses is clearly stated, the unsatisfactory respondents are identified before the analysis, and the number of respondents eliminated is disclosed. ■

While analysing the data, the researcher may also have to deal with ethical issues. The assumptions underlying the statistical techniques used to analyse the data must be satisfied to obtain meaningful results. For example, the error terms in bivariate regression must be normally distributed about zero, with a constant variance, and be uncorrelated (Chapter 20). The researcher has the responsibility to test these assumptions and take appropriate corrective actions if necessary. The appropriateness of the statistical techniques used for analysis should be discussed when presenting the results. When this is not done, ethical questions can be raised.

Internet and computer applications

In evaluating software that can help the marketing researcher with data preparation tasks, a distinction should be made between survey design and analysis packages such as SNAP (www.snapsurveys.com), Keypoint2 (www.camsp.com) and SphinxSurvey (www.scolari.co.uk), and statistical packages such as SPSS (www.spss.com), SAS (www.sas.com) and Minitab (www.minitab.com), and office products such as Microsoft Excel (www.microsoft.com). It is advised that these Websites are visited to view the demonstration versions of the packages to gain a feel for their capabilities and applications.

Survey design and analysis packages such as SNAP, Keypoint2 and Sphinx perform a much broader range of tasks to help with the array of data preparation tasks than either the statistical packages or the more generic office products. Primarily, they allow the physical format of a questionnaire to be designed. The designed questionnaire can either be printed out and used for traditional mail surveys, formatted for interviewer-led surveys using CATI or CAPI, or formatted for self-completion Internet surveys. As the questionnaire is designed on screen, the underlying structure of the questionnaire is automatically created. This results in the coding of question replies and the associated field positions automatically being worked out. It also means that any alterations in the design of the questionnaire automatically maintain the structure and integrity of the data being collected, ensuring accuracy in the later analysis stages.

Data can be entered directly by respondents as an Internet survey progresses, or entered by an interviewer on a CATI or CAPI survey. Alternatively, the data can be keyed directly from a paper questionnaire on to the computer. The packages have built in checks for 'out of range' responses which halt the progress of data entry; for example, if a five-point Likert scale is used and a '7' is entered, there is an audible warning and the data entry halts until the error is corrected.

Multiple choice questions typically appear as tick boxes and can be set as either a single response or a multiple response, and it is a simple task to highlight one or more of the boxes. Alternatively, numbers, dates or verbatim responses can be entered. The verbatim data can be transferred to qualitative data analysis packages if needed, or can be post-coded (coded later) and a new variable established for subsequent analysis.

Data preparation effort is considerably reduced when replies are entered directly on to the computer as each question is asked. For paper-based surveys this is not practical and optical scanning can be a cost-effective solution, particularly when there is a large proportion of multiple response questions. The pages of the questionnaire are scanned and the systems search for marks within boxes, and even allow for situations where a reply has been altered or removed. Scanning becomes less cost effective when the proportion of open questions increases, as there is often a high level of manual intervention to clean and code the responses. Scanning systems currently recognise and interpret numbers and hand-printed text, but are not yet capable of accurately recognising hand-written text. Some systems can be 'trained' to recognise an individual style of handwriting, but such a task is inappropriate for normal marketing research surveys, particularly if they are self-completion. For certain types of surveys, scanning is a great time saver. For others, the scanning, coding and cleaning are no faster than manually entering the data on to the computer.

Verification is the process of manually re-entering a proportion of the questionnaires to check the accuracy of data entry, and this facility is regularly used for paper-based surveys. Options are available to set a percentage level (perhaps 5–10%) and the programs then randomly select questionnaires to be re-entered.

The programs also allow for validation checks between questions to be set up, completing a comprehensive error avoidance and checking ability, which in total means that the full array of data cleaning tasks can be performed.

Once the dataset has been created, checked and cleaned, SNAP and Keypoint2 allow statistical adjustment of the data through the creation of weights, scale transformation and the ability to create new variables or modify existing ones. Univariate and basic multivariate statistical analyses can be performed and presented in an array of styles of tables and graphs. Before questionnaires are sent out, the forms of analysis and the resulting tables and graphs can be designed and the instructions for these forms stored. This means that, as questionnaire responses start to build up in any survey format, a full set of tables and graphs as an interim analysis can be performed at any time by the researcher.

In essence, these packages perform the complete array of tasks faced by the marketing researcher, from designing a questionnaire and ensuring a sound dataset is built up to analysing the data and presenting the findings. There are limitations to the extent and array of multivariate statistical analyses that can be performed by these packages. However, they do allow for links to statistical analysis packages. For example, SNAP has the option to transform the files that describe the question and answer structure, and the raw data of a survey, into a fully labelled and coded SPSS SAV file, ready to perform any of the multivariate analysis tasks as detailed in Chapters 18–24. Alternatively, SNAP can both export and import Triple S files, **www.triple-s.org**, a standard used by over 50 survey-related software packages worldwide to represent both the definitions of the survey and the associated data. Go to the Companion Website and read Professional Perspective 20 'Make it snappy' by Tim Macer. Tim reviews the latest version of SNAP, explaining its features and outlining its benefits and limitations.

➤➤➤
See Professional Perspective 20.

Summary

Data preparation begins with a preliminary check of all questionnaires for completeness and interviewing quality. Then more thorough editing takes place. Editing consists of screening questionnaires to identify illegible, incomplete, inconsistent or ambiguous responses. Such responses may be handled by returning questionnaires to the field, assigning missing values or discarding unsatisfactory respondents.

The next step is coding. A numeric or alphanumeric code is assigned to represent a specific response to a specific question along with the column position or field that code will occupy. It is often helpful to prepare a codebook containing the coding instructions and the necessary information about the variables in the dataset. The coded data are transcribed on to disks or magnetic tapes or entered directly into a data analysis package. Mark sense forms, optical scanning or computerised sensory analysis may also be used.

Cleaning the data requires consistency checks and treatment of missing responses. Options available for treating missing responses include substitution of a neutral value such as a mean, substitution of an imputed response, casewise deletion and pairwise deletion. Statistical adjustments such as weighting, variable re-specification and scale transformations often enhance the quality of data analysis. The selection of a data analysis strategy should be based on the earlier steps of the marketing research process, known characteristics of the data, properties of statistical techniques, and the background and philosophy of the researcher. Statistical techniques may be classified as univariate or multivariate.

Before analysing the data in international marketing research, the researcher should ensure that the units of measurement are comparable across countries or cultural units. The data analysis could be conducted at three levels: (1) individual, (2) within-country or cultural unit (intra-cultural analysis), and (3) across countries or cultural units (pan-cultural or cross-cultural analysis). Several ethical issues are related to data processing, particularly the discarding of unsatisfactory responses, violation of the assumptions underlying the data analysis techniques, and evaluation and interpretation of results.

Questions

1 Describe the data preparation process. Why is this process needed?

2 What activities are involved in the preliminary checking of questionnaires that have been returned from the field?

3 What is meant by editing a questionnaire?

4 How are unsatisfactory responses that are discovered in editing treated?

5 What is the difference between pre-coding and post-coding?

6 Describe the guidelines for the coding of unstructured questions.

7 What does transcribing the data involve?

8 What kinds of consistency checks are made in cleaning the data?

9 What options are available for the treatment of missing data?

10 What kinds of statistical adjustments are sometimes made to the data?

11 Describe the weighting process. What are the reasons for weighting?

12 What are dummy variables? Why are such variables created?

13 Explain why scale transformations are made.

14 Which scale transformation procedure is most commonly used? Briefly describe this procedure.

15 What considerations are involved in selecting a data analysis strategy?

Notes

1 Dadzie, K.Q., 'Demarketing strategy in shortage marketing environment', *Journal of the Academy of Marketing Science* (Spring 1989), 157–65. See also Davidson, F., *Principles of Statistical Data Handling* (Thousand Oaks, CA: Sage, 1996).

2 Alreck, P.L. and Settle, R.B., *The Survey Research Handbook* (Homewood, IL: Irwin, 1985), 254–86.

3 For a detailed discussion of coding, see Sidel, P.S., 'Coding', in Ferber, R. (ed.), *Handbook of Marketing Research* (New York: McGraw-Hill, 1974), 2-178–2.199.

4 Luyens, S., 'Coding verbatims by computer', *Marketing Research: A Magazine of Management and Applications* 7(2) (Spring 1995), 20–5.

5 Frendberg, N., 'Scanning questionnaires efficiently', *Marketing Research: A Magazine of Management and Applications* 5(2) (Spring 1993), 38–42.

6 Einspruch, E.L., *An Introductory Guide to SPSS for Windows* (Thousand Oaks, CA: Sage, 1998); Spector, P.E., *SAS Programming for Researchers and Social Scientists* (Thousand Oaks, CA: Sage 1993); Norat, M.A., 'Software reviews', *Economic Journal: The Journal of the Royal Economic Society* 107(442) (May 1997), 857–82.

7 Freedman, V.A. and Wolf, D.A., 'A case study on the use of multiple imputation', *Demography* 32(3) (August 1995), 459–70; Malhotra, N.K., 'Analysing marketing research data with incomplete information on the dependent variable', *Journal of Marketing Research* 24 (February 1987), 74–84.

8 A meaningful and practical value should be imputed. The value imputed should be a legitimate response code. For example, a mean of 3.86 may not be practical if only single-digit response codes have been developed. In such cases, the mean should be rounded to the nearest integer.

9 Malhotra, N.K., 'Analysing marketing research data with incomplete information on the dependent variable', *Journal of Marketing Research* 24 (February 1987), 74–84.

10 Some weighting procedures require adjustments in subsequent data analysis techniques. See Sharot, T., 'Weighting survey results', *Journal of the Market Research Society* 28 (July 1986), 269–84; Frankel, M.R., *Inference from Survey Samples* (Ann Arbor, MI: Institute for Social Research, University of Michigan, 1971).

11 Woodside, A.G., Nielsen, R.L., Walters, F. and Muller, G.D., 'Preference segmentation of health care services: the old-fashioneds, value conscious, affluents, and professional want-it-alls', *Journal of Health Care Marketing* (June 1988), 14–24. See also Jayanti, R., 'Affective responses toward service providers: implications for service encounters', *Health Marketing Quarterly* 14(1) (1996), 49–65.

12 See Swift, B., 'Preparing numerical data', in Sapsford, R. and Jupp, V. (eds), *Data Collection and Analysis* (Thousand Oaks, CA: Sage, 1996); Frank, R.E., 'Use of transformations', *Journal of Marketing Research* (August 1966), 247–53, for specific transformations frequently used in marketing research.

13 For a similar data analysis strategy, see Malhotra, N.K., 'Modelling store choice based on censored preference data', *Journal of Retailing* (Summer 1986), 128–44; Birks, D.F. and Birts, A.N., 'Service quality in domestic banks', in Birks, D.F. (ed.), *Global Cash Management in Europe* (Basingstoke: Macmillan, 1998), 175–205.

14 Bivariate techniques have been included here with multivariate techniques. Although bivariate techniques are concerned with pairwise relationships, multivariate techniques examine more complex simultaneous relationships among phenomena. See Tacq, J., *Multivariate Analysis Techniques in Social Science Research Analysis* (Thousand Oaks, CA: Sage ,1996).

15 Carroll, J.D. and Green, P.E., 'Psychometric methods in marketing research: Part ii: Multidimensional scaling', *Journal of Marketing Research* 34(2) (May 1997), 193–204.

16 Kilburn, D., 'Häagen-Dazs is flavor of month', *Marketing Week* 20(23) (4 September 1997), 30; Maremont, M., 'They're all screaming for Häagen-Dazs', *Business Week* (14 October 1991).

17 Alasuutari, P., *Researching Culture* (Thousand Oaks, CA: Sage, 1995); Tan, C.T., McCullough, J. and Teoh, J., 'An individual analysis approach to cross-cultural research', in Wallendorf, M. and Anderson, P. (eds), *Advances in Consumer Research* 14 (Provo, UT: Association for Consumer Research, 1987), 394–7.

18 See, for example, Spiller, L.D. and Campbell, A.J., 'The use of international direct marketing by small businesses in Canada, Mexico, and the United States: a comparative analysis', *Journal of Direct Marketing* 8 (Winter 1994), 7–16; Nyaw, M.K. and Ng, I., 'A comparative analysis of ethical beliefs: a four country study', *Journal of Business Ethics* 13 (July 1994), 543–56.

19 Newman, D.L. and Brown, R.D., *Applied Ethics for Program Evaluation Analysis* (Thousand Oaks, CA: Sage, 1996); Zinkhan, G.M., Bisesi, M. and Saxton, M.J., 'MBAs' changing attitudes toward marketing dilemmas: 1981–1987', *Journal of Business Ethics* 8 (1989), 963–74.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Frequency distribution, cross-tabulation and hypothesis testing

Objectives

After reading this chapter, you should be able to:

1 describe the significance of preliminary data analysis and the insights that can be obtained from such an analysis;

2 discuss data analysis associated with frequencies, including measures of location, measures of variability and measures of shape;

3 explain data analysis associated with cross-tabulations and the associated statistics: chi-square, phi coefficient, contingency coefficient, Cramer's *V* and lambda coefficient;

4 describe data analysis associated with parametric hypothesis testing for one sample, two independent samples and paired samples;

5 understand data analysis associated with non-parametric hypothesis testing for one sample, two independent samples and paired samples.

Frequency distribution, cross-tabulation and hypothesis testing are the fundamental building blocks of quantitative data analysis. They provide insights into the data, guide subsequent analyses and aid the interpretation of results.

Overview

Once the data have been prepared for analysis (Chapter 17), the researcher should conduct basic analyses. This chapter describes basic data analyses, including frequency distribution, cross-tabulation and hypothesis testing. First, we describe the frequency distribution and explain how it provides both an indication of the number of out-of-range, missing or extreme values as well as insights into the central tendency, variability and shape of the underlying distribution. Next, we introduce hypothesis testing by describing the general procedure. Hypothesis testing procedures are classified as tests of associations or tests of differences. We consider the use of cross-tabulation for understanding the associations between variables taken two or three at a time. Although the nature of the association can be observed from tables, statistics are available for examining the significance and strength of the association. Finally, we present tests for examining hypotheses related to differences based on one or two samples.

Many marketing research projects do not go beyond basic data analysis. These findings are often displayed using tables and graphs, as discussed further in Chapter 25. Although the findings of basic analysis are valuable in their own right, they also provide guidance for conducting multivariate analysis. The insights gained from the basic analysis are also invaluable in interpreting the results obtained from more sophisticated statistical techniques. The following examples provide a 'flavour' of basic data analysis techniques. We illustrate the use of cross-tabulation, chi-square analysis and hypothesis testing.

example
GlobalCash Project

Basic data analyses

In the GlobalCash Project, basic data analysis formed the foundation for conducting subsequent multivariate analysis. Data analysis began by obtaining a frequency distribution and descriptive statistics for each variable or question asked in the survey. In addition to identifying possible problems with the data, this information provided a good idea of the data and insights into how specific variables should be treated in subsequent analyses. For example, should some variables such as the 'global turnover of companies in €million' be treated as categorical, and, if so, how many categories should there be? Several two- and three-variable cross-tabulations were also conducted to identify associations in the data. The effects of variables with two categories on the metric dependent variables of interest were examined by means of *t* tests and other hypothesis testing procedures. ■

example
GlobalCash Project

Bank accounts outside one's home country

Measures of companies' plans over the next two years were conducted in the GlobalCash Project. Results showed differences in the nature of these plans in different countries. The following table focuses upon the results from German respondents. The first question tackled was the intention to change the number of banks a company has business relationships with. The responses to this question were broken down into respondents who plan, or do not plan for more automation in their treasury function. Cross-tabulation and chi-square analysis provided the following:

Intend to change the number of banks you use?	Treasury function to be more automated? (%)	
	Yes	No
No	59	74
Increase	6	7
Decrease	35	19
	$\chi^2 = 6.51$	$p = < 0.05$

These results indicate that in German companies there is little difference in plans to automate when companies plan to increase the number of banks they work with. Where a decrease in the number of banks is planned, a much higher proportion plan to have more automation in their treasury function. ■

Catalogues are risky business[1]

Twelve product categories were examined to compare shopping by catalogue with store shopping. The hypothesis that there is no significant difference in the overall amount of risk perceived when buying products by catalogue compared with buying the same products in a retail store was rejected. The hypothesis was tested by computing 12 paired-observations t tests, one for each product. Mean scores for overall perceived risk for some of the products in both buying situations are presented in the following table, with higher scores indicating greater risk.

Product	Overall perceived risk	
	Catalogue	*Store*
Shoes	58.60	50.80*
Pocket calculator	49.62	42.00*
Hi-fi	48.89	41.98*
Portable television	48.53	40.91*
Digital camera	48.13	39.52*
Athletic socks	35.22	30.22*
Perfume	34.85	29.79*
CDs	32.65	28.74*

* Significant at 0.01 level

As can be seen, a significantly ($p < 0.01$) higher overall amount of perceived risk was attached to products purchased by catalogue as compared with those purchased from a retail store. ■

The first GlobalCash example illustrates the role of basic data analysis used in conjunction with multivariate procedures, whereas the other two examples show how such analysis can be useful in its own right. The cross-tabulation and chi-square analysis in the GlobalCash plans example and the paired t tests in the catalogue shopping example enabled us to draw specific conclusions from the data. Before these types of conclusions are drawn, it is useful to examine the frequency distributions of the relevant variables.

Frequency distribution

Marketing researchers often need to answer questions about a single variable. For example:

- How many users of the brand may be characterised as brand loyal?
- What percentage of the market consists of heavy users, medium users, light users and non-users?
- How many customers are very familiar with a new product offering? How many are familiar, somewhat familiar, or unfamiliar with the brand? What is the mean familiarity rating? Is there much variance in the extent to which customers are familiar with the new product?
- What is the income distribution of brand users? Is this distribution skewed towards low income brackets?

Frequency distribution
A mathematical distribution whose objective is to obtain a count of the number of responses associated with different values of one variable and to express these counts in percentage terms.

The answers to these kinds of questions can be determined by examining **frequency distributions**. In a frequency distribution, one variable is considered at a time.

The objective is to obtain a count of the number of responses associated with different values of the variable. The relative occurrence, or frequency, of different values of the variable is expressed in percentages. A frequency distribution for a variable produces a table of frequency counts, percentages and cumulative percentages for all the values associated with that variable.

Table 18.1 **Rating of 'good electronic banking system' as a criterion in selecting a bank**

Value label	Value	Frequency (N)	Percentage	Valid percentage	Cumulative percentage
Most important criterion	1	203	18.8	18.9	18.9
Second most important	2	166	15.4	15.5	34.4
Third most important	3	138	12.8	12.8	47.2
Fourth most important	4	104	9.6	9.7	56.9
Fifth most important	5	77	7.1	7.2	64.1
Not rated	6	386	35.8	35.9	100.0
Missing	9	5	0.5	missing	
TOTAL		1079	100.0	100.0	

Table 18.1 gives the frequency distribution of the replies to a question on selecting banks in the GlobalCash Project. Respondents were asked to rank order the top five criteria they use in selecting a bank. One of those criteria was a 'good electronic banking system' and the replies to that criterion are shown in Table 18.1. In the table, the first column contains the labels assigned to the different categories of the variable and the second column indicates the codes assigned to each value. Note that a code of 9 has been assigned to missing values. The third column gives the number of respondents ticking each value. For example, 203 respondents ticked value 1, indicating that they felt that a good electronic banking system was their most important criterion when selecting a bank. The fourth column displays the percentage of respondents ticking each value.

The fifth column shows percentages calculated by excluding the cases with missing values. If there are no missing values, columns 4 and 5 are identical. The last column represents cumulative percentages after adjusting for missing values. As can be seen, of the 1,079 respondents who participated in the survey, 18.8% entered a figure of '1' indicating the criterion that was the 'most important'. If the five respondents with missing values are excluded, this changes to 18.9%. Examining the cumulative percentage column, it is clear that 47.2% of respondents would rate 'a good electronic banking system' as being in their top three criteria in selecting a bank. In other words, 47.2% of the respondents with valid responses indicated a value of 3 or less.

A frequency distribution helps determine the extent of item non-response. It also indicates the extent of illegitimate responses. Values of 0, 7 and 8 would be illegitimate responses, or errors. The cases with these values could be identified and corrective action could be taken. The presence of outliers or cases with extreme values can also be detected. For example, in the case of a frequency distribution of household size, a few isolated families with household sizes of nine or more might be considered outliers. A frequency distribution also indicates the shape of the empirical distribution of the variable. The frequency data may be used to construct a histogram, or a vertical bar chart in which the values of the variable are portrayed along the x axis and the absolute or relative frequencies of the values are placed along the y axis.
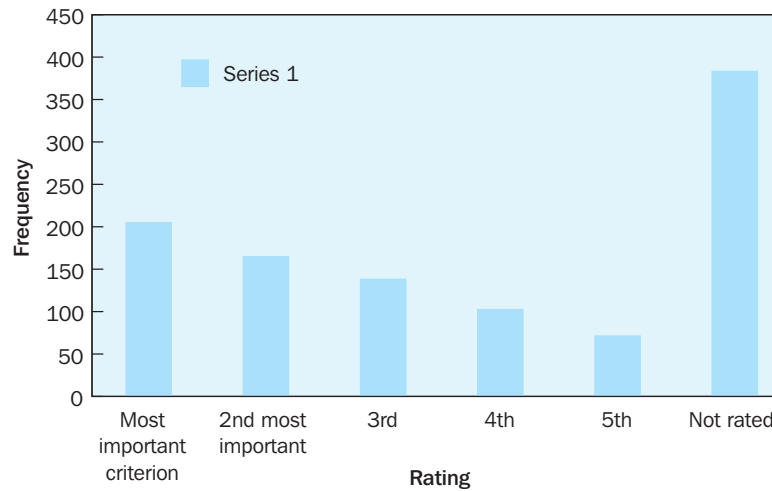
Figure 18.1 is a histogram of the frequency data in Table 18.1. From the histogram, one could examine whether the observed distribution is consistent with an expected or assumed distribution.

Statistics associated with frequency distribution

A frequency distribution is a convenient way of looking at different values of a variable. A frequency table is easy to read and provides basic information, but sometimes this information may be too detailed and the researcher must summarise it by the use of descriptive statistics.[2] The most commonly used statistics associated with frequencies are measures of location (mean, mode and median), measures of variability (range, interquartile range, variance, standard deviation and coefficient of variation), and measures of shape (skewness and kurtosis).[3]

Measures of location

Measure of location
A statistic that describes a location within a data set. Measures of central tendency describe the centre of the distribution.

The **measures of location** that we discuss are measures of central tendency because they tend to describe the centre of the distribution. If the entire sample is changed by adding a fixed constant to each observation, then the mean, mode and median change by the same fixed amount.

Mean
The average; that value obtained by summing all elements in a set and dividing by the number of elements.

Mean. The **mean**, or average value, is the most commonly used measure of central tendency. The measure is used to estimate the mean when the data have been collected using an interval or ratio scale. The data should display some central tendency, with most of the responses distributed around the mean.

The mean, $\bar{X}$, is given by

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

where

X_i = observed values of the variable X

n = number of observations (sample size).

Generally, the mean is a robust measure and does not change markedly as data values are added or deleted. For the frequencies given in Table 18.1, the mean value is calculated as follows:

$$\bar{X} = \frac{(203 \times 1) + (166 \times 2) + (138 \times 3) + (104 \times 4) + (77 \times 5) + (386 \times 6)}{1079}$$

$$= \frac{203 + 332 + 414 + 416 + 385 + 2316}{1079}$$

$$= \frac{4066}{1079}$$

$$= 3.768$$

Mode. The **mode** is the value that occurs most frequently. It represents the highest peak of the distribution. The mode is a good measure of location when the variable is inherently categorical or has otherwise been grouped into categories. The mode in Table 18.1 is 6.

Median. The **median** of a sample is the middle value when the data are arranged in ascending or descending order. If the number of data points is even, the median is usually estimated as the midpoint between the two middle values by adding the two middle values and dividing their sum by 2. The median is the 50th percentile. The median is an appropriate measure of central tendency for ordinal data. In Table 18.1, the middle value is 4, so the median is 4.

As can be seen from Table 18.1, the three measures of central tendency for this distribution are different (mean = 3.768, mode = 6, median = 4). This is not surprising, since each measure defines central tendency in a different way. So which measure should be used? If the variable is measured on a nominal scale, the mode should be used. If the variable is measured on an ordinal scale, as in Figure 18.1 and Table 18.1, the median is appropriate.

If the variable is measured on an interval or ratio scale, the mode is a poor measure of central tendency. In general, for interval or ratio data, the median is a better measure of central tendency, although it too ignores available information about the variable. The actual values of the variable above and below the median are ignored. The mean is the most appropriate measure of central tendency for interval or ratio data. The mean makes use of all the information available since all of the values are used in computing it. The mean, however, is sensitive to extremely small or extremely large values (outliers). When there are outliers in the data, the mean is not a good measure of central tendency, and it is useful to consider both the mean and the median.

Measures of variability

The **measures of variability**, which are calculated on interval or ratio data, include the range, interquartile range, variance or standard deviation, and coefficient of variation.

Range. The **range** measures the spread of the data. It is simply the difference between the largest and smallest values in the sample:

$$\text{range} = X_{\text{largest}} - X_{\text{smallest}}$$

As such, the range is directly affected by outliers. If all the values in the data are multiplied by a constant, the range is multiplied by the same constant. The range in Table 18.1 is $6 - 1 = 5$.

Mode
A measure of central tendency given as the value that occurs with the most frequency in a sample distribution.

Median
A measure of central tendency given as the value above which half of the values fall and below which half of the values fall.

Measure of variability
A statistic that indicates the distribution's dispersion.

Range
The difference between the smallest and largest values of a distribution.

Interquartile range
The range of a distribution encompassing the middle 50% of the observations.

Interquartile range. The **interquartile range** is the difference between the 75th and 25th percentiles. For a set of data points arranged in order of magnitude, the pth percentile is the value that has $p\%$ of the data points below it and $(100 - p)\%$ above it. If all the data points are multiplied by a constant, the interquartile range is multiplied by the same constant. The interquartile range in Table 18.1 is $6 - 2 = 4$.

Variance
The mean squared deviation of all the values of the mean.

Variance. The difference between the mean and an observed value is called the deviation from the mean. The **variance** is the mean squared deviation from the mean. The variance can never be negative. When the data points are clustered around the mean, the variance is small. When the data points are scattered, the variance is large. If all the data values are multiplied by a constant, the variance is multiplied by the square of the constant.

Standard deviation
The square root of the variance.

Standard deviation. The **standard deviation** is the square root of the variance. Thus, the standard deviation is expressed in the same units as the data, rather than in squared units. The standard deviation of a sample, s_x, is calculated as:

$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

We divide by $n - 1$ instead of n because the sample is drawn from a population and we are trying to determine how much the responses vary from the mean of the entire population. The population mean is unknown, however; therefore, the sample mean is used instead. The use of the sample mean makes the sample seem less variable than it really is. By dividing by $n - 1$ instead of by n, we compensate for the smaller variability observed in the sample. For the data given in Table 18.1, the variance is calculated as follows:

$$s^2_x = \frac{\{203 \times (1 - 3.768)^2 + 166 \times (2 - 3.768)^2 + 138 \times (3 - 3.768)^2}{1073}$$
$$+ 104 \times (4 - 3.768)^2 + 77 \times (5 - 3.768)^2 + 368 \times (6 - 3.768)^2\}$$

$$= \frac{1555.35 + 518.87 + 81.40 + 5.60 + 116.87 + 1922.98}{1073}$$

$$= \frac{4201.07}{1073}$$

$$= 3.91$$

The standard deviation, therefore, is calculated as

$$s_x = \sqrt{3.91}$$

$$= 1.98$$

Coefficient of variation
A useful expression in sampling theory for the standard deviation as a percentage of the mean.

Coefficient of variation. The **coefficient of variation** is the ratio of the standard deviation to the mean expressed as a percentage, and it is a unitless measure of relative variability. The coefficient of variation, CV, is expressed as

$$CV = \frac{s_x}{\bar{X}}$$

The coefficient of variation is meaningful only if the variable is measured on a ratio scale. It remains unchanged if all the data values are multiplied by a constant. Since the data in Table 18.1 are not measured on a ratio scale, it is not meaningful to calculate the coefficient of variation.

Measures of shape

In addition to measures of variability, measures of shape are also useful in understanding the nature of the distribution. The shape of a distribution is assessed by examining skewness and kurtosis.

Skewness. Distributions can be either symmetric or skewed. In a symmetric distribution, the values on either side of the centre of the distribution are the same, and the mean, mode and median are equal. The positive and corresponding negative deviations from the mean are also equal. In a skewed distribution, the positive and negative deviations from the mean are unequal. Skewness is the tendency of the deviations from the mean to be larger in one direction than in the other. It can be thought of as the tendency for one tail of the distribution to be heavier than the other (see Figure 18.2). The skewness value for the data of Table 18.1 is –0.352, indicating a negative skew.

Kurtosis. Kurtosis is a measure of the relative peakedness or flatness of the curve defined by the frequency distribution. The kurtosis of a normal distribution is zero. If the kurtosis is positive, then the distribution is more peaked than a normal distribution. A negative value means that the distribution is flatter than a normal distribution. The value of this statistic for Table 18.1 is –0.0113, indicating that the distribution is flatter than a normal distribution.

Skewness
A characteristic of a distribution that assesses its symmetry about the mean.

Kurtosis
A measure of the relative peakedness of the curve defined by the frequency distribution.

A general procedure for hypothesis testing

Basic analysis invariably involves some hypothesis testing. Examples of hypotheses generated in marketing research abound:

- A cinema is being patronised by more than 10% of the households in a city.
- The heavy and light users of a brand differ in terms of psychographic characteristics.
- One hotel has a more 'luxurious' image than its close competitor.
- Familiarity with a restaurant results in greater preference for that restaurant.

Chapter 15 covered the concepts of the sampling distribution, standard error of the mean or the proportion, and the confidence interval.[4] All these concepts are relevant to hypothesis testing and should be reviewed. We now describe a general procedure for hypothesis testing that can be applied to test hypotheses about a wide range of parameters.
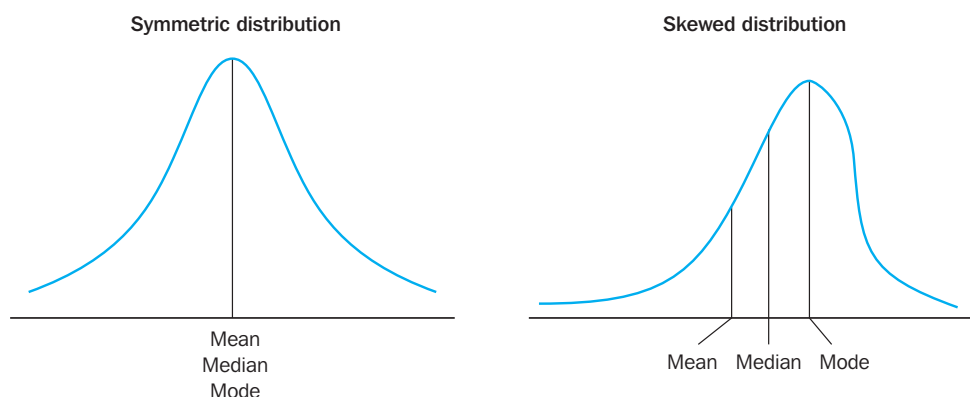
Figure 18.2
Skewness of a distribution

The following steps are involved in hypothesis testing (Figure 18.3).

1 Formulate the null hypothesis H_0 and the alternative hypothesis H_1.
2 Select an appropriate statistical technique and the corresponding test statistic.
3 Choose the level of significance, α.
4 Determine the sample size and collect the data. Calculate the value of the test statistic.
5 Determine the probability associated with the test statistic under the null hypothesis, using the sampling distribution of the test statistic. Alternatively, determine the critical values associated with the test statistic that divide the rejection and non-rejection region.
6 Compare the probability associated with the test statistic with the level of significance specified. Alternatively, determine whether the test statistic has fallen into the rejection or the non-rejection region.
7 Make the statistical decision to reject or not reject the null hypothesis.
8 Express the statistical decision in terms of the marketing research problem.

Step 1: Formulate the hypothesis

Null hypothesis
A statement in which no difference or effect is expected. If the null hypothesis is not rejected, no changes will be made.

Alternative hypothesis
A statement that some difference or effect is expected. Accepting the alternative hypothesis will lead to changes in opinions or actions.

The first step is to formulate the null and alternative hypotheses. A **null hypothesis** is a statement of the status quo, one of no difference or no effect. If the null hypothesis is not rejected, no changes will be made. An **alternative hypothesis** is one in which some difference or effect is expected. Accepting the alternative hypothesis will lead to changes in opinions or actions. Thus, the alternative hypothesis is the opposite of the null hypothesis.

The null hypothesis is always the hypothesis that is tested. The null hypothesis refers to a specified value of the population parameter (e.g. μ, σ, π), not a sample statistic (e.g. $\overline{X}$). A null hypothesis may be rejected, but it can never be accepted based on a single test. A statistical test can have one of two outcomes: that the null hypothesis is rejected and the alternative hypothesis accepted, or that the null hypothesis is not
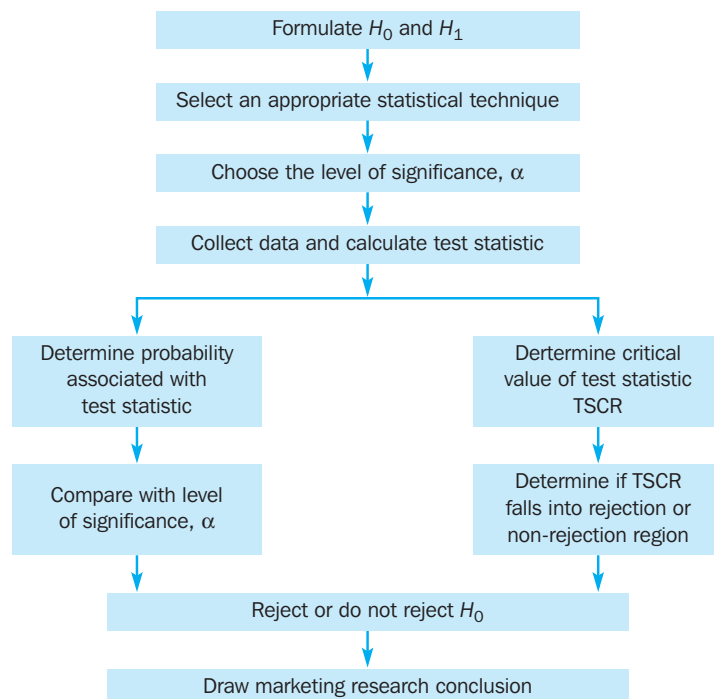
Figure 18.3
A general procedure for hypothesis testing

453

rejected based on the evidence. It would be incorrect, however, to conclude that since the null hypothesis is not rejected, it can be accepted as valid. In classical hypothesis testing, there is no way to determine whether the null hypothesis is true.

In marketing research, the null hypothesis is formulated in such a way that its rejection leads to the acceptance of the desired conclusion. The alternative hypothesis represents the conclusion for which evidence is sought. For example, a garage is considering introducing a collection and delivery system when customers' cars need repairs. Given the investment in personnel to make this plan work well, it will only be introduced if it is preferred by more than 40% of the customers. The appropriate way to formulate the hypotheses is

$$H_0: \pi \leq 0.40$$

$$H_1: \pi > 0.40$$

If the null hypothesis H_0 is rejected, then the alternative hypothesis H_1 will be accepted and the new collection and delivery service introduced. On the other hand, if H_0 is not rejected, then the new collection and delivery service should not be introduced unless additional evidence is obtained. The test of the null hypothesis is a **one-tailed test** because the alternative hypothesis is expressed directionally: the proportion of customers who express a preference is greater than 0.40.

On the other hand, suppose that the researcher wanted to determine whether the new collection and delivery service is different (superior or inferior) from the existing form of getting a car to the garage for repairs, which is preferred by 40 per cent of the customers. Then a **two-tailed test** would be required, and the hypotheses would be expressed as

$$H_0: \pi = 0.40$$

$$H_1: \pi \neq 0.40$$

In commercial marketing research, the one-tailed test is used more often than a two-tailed test. Typically, there is some preferred direction for the conclusion for which evidence is sought. For example, the higher the profits, sales and product quality, the better. The one-tailed test is more powerful than the two-tailed test. The power of a statistical test is discussed further in step 3.

Step 2: Select an appropriate statistical technique

To test the null hypothesis, it is necessary to select an appropriate statistical technique. The researcher should take into consideration how the **test statistic** is computed and the sampling distribution that the sample statistic (e.g. the mean) follows. The test statistic measures how close the sample has come to the null hypothesis. The test statistic often follows a well-known distribution, such as the normal, t, or chi-square distribution. Guidelines for selecting an appropriate test or statistical technique are discussed later in this chapter. In our example, the z statistic, which follows the standard normal distribution, would be appropriate. This statistic would be computed as follows:

$$z = \frac{p - \pi}{\sigma_p}$$

where

$$\sigma_p = \sqrt{\frac{\pi(1 - \pi)}{n}}$$

One-tailed test
A test of the null hypothesis where the alternative hypothesis is expressed directionally.

Two-tailed test
A test of the null hypothesis where the alternative hypothesis is not expressed directionally.

Test statistic
A measure of how close the sample has come to the null hypothesis. It often follows a well-known distribution, such as the normal, t, or chi-square distribution.

Step 3: Choose the level of significance

Whenever we draw inferences about a population, there is a risk that an incorrect conclusion will be reached. Two types of error can occur.

Type I error occurs when the sample results lead to the rejection of the null hypothesis when it is in fact true. In our example, a type I error would occur if we concluded, based on sample data, that the proportion of customers preferring the new collection and delivery service was greater than 0.40, when in fact it was less than or equal to 0.40. The probability of type I error (α) is also called the **level of significance**. The type I error is controlled by establishing the tolerable level of risk of rejecting a true null hypothesis. The selection of a particular risk level should depend on the cost of making a type I error.

Type II error occurs when, based on the sample results, the null hypothesis is not rejected when it is in fact false. In our example, the type II error would occur if we concluded, based on sample data, that the proportion of customers preferring the new collection and delivery service was less than or equal to 0.40 when in fact it was greater than 0.40. The probability of type II error is denoted by β. Unlike α, which is specified by the researcher, the magnitude of β depends on the actual value of the population parameter (proportion). The probability of type I error (α) and the probability of type II error (β) are shown in Figure 18.4.

The complement $(1 - \beta)$ of the probability of a type II error is called the **power of a statistical test**. The power of a test is the probability $(1 - \beta)$ of rejecting the null hypothesis when it is false and should be rejected. Although β is unknown, it is related to α. An extremely low value of α (e.g. 0.001) will result in intolerably high β errors. So it is necessary to balance the two types of errors. As a compromise, α is often set at 0.05; sometimes it is 0.01; other values of α are rare. The level of α along with the sample size will determine the level of β for a particular research design. The risk of both α and β can be controlled by increasing the sample size. For a given level of α, increasing the sample size will decrease β, thereby increasing the power of the test.

Step 4: Collect the data and calculate the test statistic

Sample size is determined after taking into account the desired α and β errors and other qualitative considerations, such as budget constraints. Then the required data are collected and the value of the test statistic is computed. Suppose, in our example, that 500 customers were surveyed and 220 expressed a preference for the new

Type I error
An error that occurs when the sample results lead to the rejection of a null hypothesis that is in fact true. Also called alpha error (a).

Level of significance
The probability of making a type I error.

Type II error
An error that occurs when the sample results lead to acceptance of a null hypothesis that is in fact false. Also called beta error (β)

Power of a statistical test
The probability of rejecting the null hypothesis when it is in fact false and should be rejected.

Figure 18.4
Type I error (α) and Type II error (β)

collection and delivery service. Thus the value of the sample proportion is $\hat{p} = 220/500 = 0.44$. The value of $\sigma \hat{p}$ can be determined as follows:

$$\sigma_{\hat{p}} = \sqrt{\frac{\pi(1-\pi)}{n}}$$

$$= \sqrt{\frac{0.40 \times 0.60}{500}}$$

$$= 0.0219$$

The test statistic z can be calculated as follows:

$$z = \frac{\hat{p} - \pi}{\sigma_{\hat{p}}}$$

$$= \frac{0.44 - 0.40}{0.0219}$$

$$= 1.83$$

Step 5: Determine the probability or the critical value

Using standard normal tables (Table 2 of the Appendix of Statistical Tables), the probability of obtaining a z value of 1.83 can be calculated (see Figure 18.5). The shaded area between $-\infty$ and 1.83 is 0.9664. Therefore, the area to the right of $z = 1.83$ is $1.0000 - 0.9664 = 0.0336$.

Alternatively, the critical value of z, which will give an area to the right side of the critical value of 0.05, is between 1.64 and 1.65 and equals 1.645. Note that, in determining the critical value of the test statistic, the area to the right of the critical value is either α or $\alpha/2$. It is α for a one-tailed test and $\alpha/2$ for a two-tailed test.

Steps 6 and 7: Compare the probability or critical values and make the decision

The probability associated with the calculated or observed value of the test statistic is 0.0336. This is the probability of getting a p value of 0.44 when $\hat{p} = 0.40$. This is less than the level of significance of 0.05. Hence, the null hypothesis is rejected. Alternatively, the calculated value of the test statistic $z = 1.83$ lies in the rejection region, beyond the value of 1.645. Again, the same conclusion to reject the null hypothesis is reached. Note that the two ways of testing the null hypothesis are equivalent but mathematically opposite in the direction of comparison. If the probability associated with the calculated or observed value of the test statistic (TS_{CAL}) is less than the level of significance (α), the null hypothesis is rejected. If the calculated value of the test statistic is *greater than* the critical value of the test statistic (TS_{CR}), however,
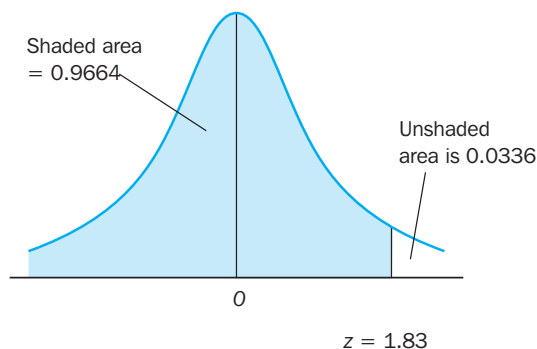
Figure 18.5
Probability of *z* with a one-tailed test

the null hypothesis is rejected. The reason for this sign shift is that the larger the value of TS_{CAL}, the smaller the probability of obtaining a more extreme value of the test statistic under the null hypothesis. This sign shift can be easily seen:

$$\text{if probability of } TS_{CAL} < \text{significance level } (\alpha), \text{ then reject } H_0$$

but

$$\text{if } TS_{CAL} > TS_{CR}, \text{ then reject } H_0$$

Step 8: Draw the marketing research conclusion

The conclusion reached by hypothesis testing must be expressed in terms of the marketing research problem. In our example, we conclude that there is evidence that the proportion of customers preferring the new service plan is significantly greater than 0.40. Hence, the recommendation would be to introduce the new collection and delivery service.

As can be seen from Figure 18.6, hypothesis testing can be related to either an examination of associations or an examination of differences. In tests of associations the null hypothesis is that there is no association between the variables (H_0: ... is NOT related to ...). In tests of differences the null hypothesis is that there is no difference (H_0: ... is NOT different than ...). Tests of differences could relate to distributions, means, proportions, or medians or rankings. First, we discuss hypotheses related to associations in the context of cross-tabulations.

Cross-tabulations

Although answers to questions related to a single variable are interesting, they often raise additional questions about how to link that variable to other variables. To introduce the frequency distribution, we posed several representative marketing research questions. For each of these, a researcher might pose additional questions to relate these variables to other variables. For example:

Cross-tabulation
A statistical technique that describes two or more variables simultaneously and results in tables that reflect the joint distribution of two or more variables that have a limited number of categories or distinct values.

- How many brand-loyal users are males?
- Is product use (measured in terms of heavy users, medium users, light users and non-users) related to interest in outdoor leisure activities (high, medium and low)?
- Is familiarity with a new product related to age and income levels?
- Is product ownership related to income (high, medium and low)?

The answers to such questions can be determined by examining **cross-tabulations**. A frequency distribution describes one variable at a time, but a cross-tabulation
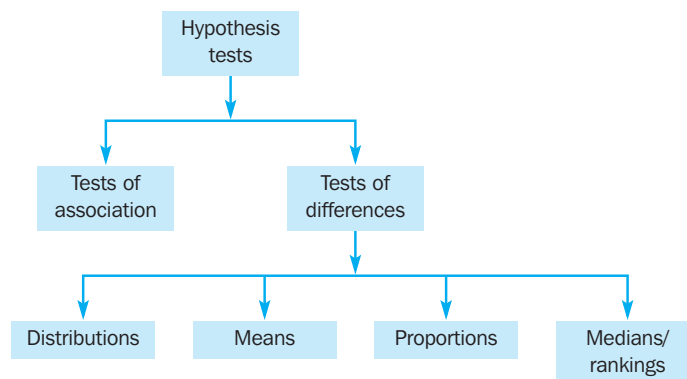
Figure 18.6
A broad classification of hypothesis testing procedures

describes two or more variables simultaneously. Cross-tabulation results in tables that reflect the joint distribution of two or more variables with a limited number of categories or distinct values. The categories of one variable are cross-classified with the categories of one or more other variables. Thus, the frequency distribution of one variable is subdivided according to the values or categories of the other variables.

Using the GlobalCash Project as an example, suppose that interest was expressed in determining whether the number of European countries that a company operates in was associated with the plans to change the number of banks they do business with. The cross-tabulation is shown in Table 18.2. A cross-tabulation includes a cell for every combination of the categories of the two variables. The number in each cell shows how many respondents gave that combination of responses. In Table 18.2, 105 operated in only one European country and did not plan to change the number of banks they do business with.

Table 18.2 Number of countries in Europe that a company operates in and plans to change the number of banks that a company does business with

Do you intend to change the number of banks you use?	In how many countries does your company operate?					
	1	2 to 5	6 to 10	11 to 15	Over 15	Row total
No	105	130	115	92	172	614
Increase	12	22	7	9	14	64
Decrease	34	57	64	44	95	294
Column total	151	209	186	145	281	972

The totals in this table indicate that, of the 972 respondents with valid responses on both the variables, 614 had no plans to change, 64 would increase and 294 would decrease the number of banks they do business with. Based on how many European countries a company operates in, 151 operate in one country, 209 in two to five countries, 186 in six to 10 countries, 145 in 11 to 15 countries and 281 in more than 15 countries. Note that this information could have been obtained from a separate frequency distribution for each variable. In general, the margins of a cross-tabulation show the same information as the frequency tables for each of the variables.

Contingency table
A cross-tabulation table. It contains a cell for every combination of categories of the two variables.

Cross-tabulation tables are also called **contingency tables**. The data are considered to be qualitative or categorical data, because each variable is assumed to have only a nominal scale.[5] Cross-tabulation is widely used in commercial marketing research because (1) cross-tabulation analysis and results can be easily interpreted and understood by managers who are not statistically oriented; (2) the clarity of interpretation provides a stronger link between research results and managerial action; (3) a series of cross-tabulations may provide greater insights into a complex phenomenon than a single multivariate analysis; (4) cross-tabulation may alleviate the problem of sparse cells, which could be serious in discrete multivariate analysis; and (5) cross-tabulation analysis is simple to conduct and appealing to less-sophisticated researchers.[6] We will discuss cross-tabulation for two and three variables.

Two variables

Cross-tabulation with two variables is also known as bivariate cross-tabulation. Consider again the cross-classification of the number of countries in Europe that a company operates in and plans to change the number of banks that a company does business with, given in Table 18.2. Is operating in a high number of European countries related to plans to reduce bank numbers? It appears to be from Table 18.2. We

see that disproportionately more of the respondents who operate in over 15 European countries plan to decrease the number of banks that they do business with compared with those that operate in 15 or fewer countries. Computation of percentages can provide more insights.

Because two variables have been cross-classified, percentages could be computed either column-wise, based on column totals (Table 18.3), or row-wise, based on row totals (Table 18.4). Which table is more useful?

Table 18.3 Plans to change the number of banks that a company does business with by number of countries in Europe that a company operates in

Do you intend to change the number of banks you use?	In how many countries does your company operate?				
	1	2 to 5	6 to 10	11 to 15	over 15
No	70%	62%	62%	63%	61%
Increase	7%	11%	4%	7%	5%
Decrease	23%	27%	34%	30%	34%
Column total	100%	100%	100%	100%	100%

Table 18.4 Number of countries in Europe that a company operates in by plans to change the number of banks that a company does business with

Do you intend to change the number of banks you use?	In how many countries does your company operate?					
	1	2 to 5	6 to 10	11 to 15	Over 15	Row total
No	17%	21%	19%	15%	28%	100%
Increase	19%	34%	11%	14%	22%	100%
Decrease	12%	19%	22%	15%	32%	100%

The answer depends on which variable will be considered as the independent variable.[7] The general rule is to compute the percentages in the direction of the independent variable, across the dependent variable. In our analysis, number of countries may be considered as the independent variable and planned changes as the dependent variable, and the correct way of calculating percentages is shown in Table 18.3. Note that while 70% of those operating in one country do not plan to make any changes, 61% of those who operate in over 15 countries do not plan to change their number of banks. This seems plausible given the costs and complexity of operating many bank accounts in many countries. Companies faced with such an array of accounts may be seeking to make further cuts and savings, especially with the introduction of European Monetary Union.

Note that computing percentages in the direction of the dependent variable across the independent variable, as shown in Table 18.4, is not meaningful in this case. Table 18.4 implies that plans to change the number of bank relationships influence companies' decisions to operate in certain numbers of European countries. This latter finding seems implausible. It is possible, however, that the association between 'change plans' and 'numbers of countries' is mediated by a third variable, such as the country where an ultimate parent company in a group operates from, e.g. although Hitachi has operations in Britain, Germany and Italy, their ultimate parentage is Japanese. It is possible that companies whose group parentage is in areas of the globe whose economic conditions are more or less favourable than Europe are affected in the extent of planned changes to their bank relationships. This kind of possibility points to the need to examine the effect of a third variable.
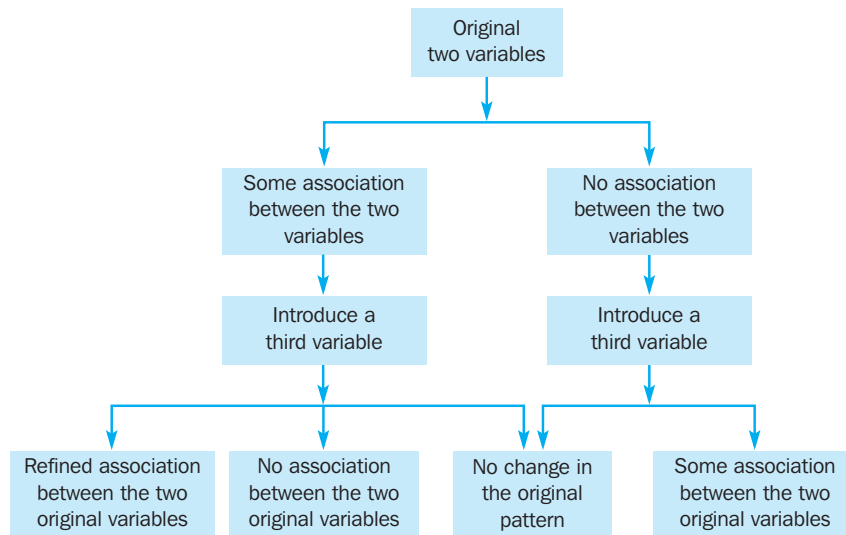
Figure 18.7
The introduction of a third variable in cross-tabulation

Three variables

Often the introduction of a third variable clarifies the initial association (or lack of it) observed between two variables. As shown in Figure 18.7, the introduction of a third variable can result in four possibilities:

1 It can refine the association observed between the two original variables.
2 It can indicate no association between the two original variables, although an association was initially observed. In other words, the third variable indicates that the initial association between the two variables was spurious.
3 It can reveal some association between the two original variables, although no association was initially observed. In this case, the third variable reveals a suppressed association between the first two variables.
4 It can indicate no change in the initial pattern.[8]

These cases are explained with examples based on a sample of 1,000 respondents. Although these examples are contrived to illustrate specific cases, such cases are not uncommon in commercial marketing research.

Refine an initial relationship. An examination of the relationship between the purchase of 'designer' clothing and marital status resulted in the data reported in Table 18.5. The respondents were classified into either high or low categories based on their purchase of 'designer' clothing. Marital status was also measured in terms of two categories: currently married or unmarried. As can be seen from Table 18.5, 52% of unmarried respondents fell in the high-purchase category as opposed to 31% of the married respondents. Before concluding that unmarried respondents purchase more 'designer' clothing than those who are married, a third variable, the buyer's gender, was introduced into the analysis.

Table 18.5 **Purchase of 'designer' clothing by marital status**

	Marital status	
Purchase of 'designer' clothing	*Married*	*Unmarried*
High	31%	52%
Low	69%	48%
Column	100%	100%
Number of respondents	700	300

The buyer's gender was selected as the third variable based on past research. The relationship between purchase of 'designer' clothing and marital status was re-examined in light of the third variable, as shown in Table 18.6. In the case of females, 60% of the unmarried respondents fall in the high-purchase category compared with 25% of those who are married. On the other hand, the percentages are much closer for males, with 40% of the unmarried respondents and 35% of the married respondents falling in the high-purchase category. Hence, the introduction of gender (third variable) has refined the relationship between marital status and purchase of 'designer' clothing (original variables). Unmarried respondents are more likely to fall into the high-purchase category than married ones, and this effect is much more pronounced for females than for males.

Table 18.6 Purchase of 'designer' clothing by marital status and gender

	Gender			
	Male Marital status		Female Marital status	
Purchase of 'designer' clothing	Married	Unmarried	Married	Unmarried
High	35%	40%	25%	60%
Low	65%	60%	75%	40%
Column	100%	100%	100%	100%
Number of respondents	400	120	300	180

Initial relationship was spurious. A researcher working for an advertising agency promoting a car brand costing more than €60,000 was attempting to explain the ownership of expensive cars (see Table 18.7). The table shows that 32 of those with university degrees own an expensive (more than €60,000) car compared with 21% of those without university degrees. The researcher was tempted to conclude that education influenced ownership of expensive cars. Realising that income may also be a factor, the researcher decided to re-examine the relationship between education and ownership of expensive cars in the light of income level. This resulted in Table 18.8. Note that the percentages of those with and without university degrees who own expensive cars are the same for each income group. When the data for the high-income and low-income groups are examined separately, the association between education and ownership of expensive cars disappears, indicating that the initial relationship observed between these two variables was spurious.

Table 18.7 Ownership of expensive cars by education level

	Education	
Own expensive car	Degree	No degree
Yes	32%	21%
No	68%	79%
Column	100%	100%
Number of respondents	250	750

Reveal suppressed association. A researcher suspected that desire to travel abroad may be influenced by age. A cross-tabulation of the two variables produced the results in Table 18.9, indicating no association. When gender was introduced as the third variable, Table 18.10 was obtained. Among men, 60% of those under 45 indicated a

Table 18.8 Ownership of expensive cars by education and income levels

	Income			
	Low income Education		High income Education	
Own expensive car	Degree	No degree	Degree	No degree
Yes	20%	20%	40%	40%
No	80%	80%	60%	60%
Column totals	100%	100%	100%	100%
Number of respondents	100	700	150	50

desire to travel abroad compared with 40% of those 45 or older. The pattern was reversed for women, where 35% of those under 45 indicated a desire to travel abroad as opposed to 65% of those 45 or older. Since the association between desire to travel abroad and age runs in the opposite direction for males and females, the relationship between these two variables is masked when the data are aggregated across gender as in Table 18.9. But when the effect of gender is controlled, as in Table 18.10, the suppressed association between preference and age is revealed for the separate categories of males and females.

Table 18.9 Desire to travel abroad by age

	Age	
Desire to travel abroad	Under 45	45 or older
Yes	50%	50%
No	50%	50%
Column totals	100%	100%
Number of respondents	500	500

Table 18.10 Desire to travel abroad by age and gender

	Gender			
	Male Age		Female Age	
Desire to travel abroad	Under 45	45 or older	Under 45	45 or older
Yes	60%	40%	35%	65%
No	40%	60%	65%	35%
Column totals	100%	100%	100%	100%
Number of respondents	300	300	200	200

No change in initial relationship. In some cases, the introduction of the third variable does not change the initial relationship observed, regardless of whether the original variables were associated. This suggests that the third variable does not influence the relationship between the first two. Consider the cross-tabulation of family size and the tendency to eat in fast-food restaurants frequently, as shown in Table 18.11. The respondents' families were classified into small- and large-size categories based on a median split of the distribution, with 500 respondents in each category. No association is observed. The respondents were further classified into high- or low-

Table 18.11 Eating frequently in fast-food restaurants by family size

Eat frequently in fast-food restaurants	Family size	
	Small	Large
Yes	65%	65%
No	35%	35%
Column totals	100%	100%
Number of respondents	500	500

Table 18.12 Eating frequently in fast-food restaurants by family size and income

Eat frequently in fast-food restaurants	Income			
	Low income Family size		High income Family size	
	Small	Large	Small	Large
Yes	65%	65%	65%	65%
No	35%	35%	35%	35%
Column total	100%	100%	100%	100%
Number of respondents	250	250	250	250

income groups based on a median split. When income was introduced as a third variable in the analysis, Table 18.12 was obtained. Again, no association was observed.

General comments on cross-tabulation

Even though more than three variables can be cross-tabulated, the interpretation is quite complex. Also, because the number of cells increases multiplicatively, maintaining an adequate number of respondents or cases in each cell can be problematic. As a general rule, there should be at least five expected observations in each cell for the computed statistics to be reliable. Thus, cross-tabulation is an inefficient way of examining relationships when there are more than a few variables. Note that cross-tabulation examines association between variables, not causation. To examine causation, the causal research design framework should be adopted (see Chapter 11).

Statistics associated with cross-tabulation

We now discuss the statistics commonly used for assessing the statistical significance and strength of association of cross-tabulated variables. The statistical significance of the observed association is commonly measured by the chi-square statistic. The strength of association, or degree of association, is important from a practical or substantive perspective. Generally, the strength of association is of interest only if the association is statistically significant. The strength of the association can be measured by the phi correlation coefficient, the contingency coefficient, Cramer's *V*, and the lambda coefficient. These statistics are described in detail.

Chi-square

Chi-square statistic
The statistic used to test the statistical significance of the observed association in a cross-tabulation. It assists us in determining whether a systematic association exists between the two variables.

The **chi-square statistic** (χ^2) is used to test the statistical significance of the observed association in a cross-tabulation. It assists us in determining whether a systematic association exists between the two variables. The null hypothesis, H_0, is that there is no

association between the variables. The test is conducted by computing the cell frequencies that would be expected if no association were present between the variables, given the existing row and column totals. These expected cell frequencies, denoted f_e, are then compared with the actual observed frequencies, f_o, found in the cross-tabulation to calculate the chi-square statistic. The greater the discrepancies between the expected and observed frequencies, the larger the value of the statistic. Assume that a cross-tabulation has r rows and c columns and a random sample of n observations. Then the expected frequency for each cell can be calculated by using a simple formula:

$$f_e = \frac{n_r n_c}{n}$$

where n_r = total number in the row
n_c = total number in the column
n = total sample size.

For the data in Table 18.2, the expected frequencies for the cells, going from left to right and from top to bottom, are

$$\frac{614 \times 151}{972} = 95.4 \qquad \frac{614 \times 209}{972} = 132 \qquad \frac{614 \times 186}{972} = 117.5$$

$$\frac{614 \times 145}{972} = 91.6 \qquad \frac{614 \times 281}{972} = 177.5 \qquad \frac{64 \times 151}{972} = 9.9$$

$$\frac{64 \times 209}{972} = 13.8 \qquad \frac{64 \times 186}{972} = 12.2 \qquad \frac{64 \times 145}{972} = 9.5$$

$$\frac{64 \times 281}{972} = 18.5 \qquad \frac{294 \times 151}{972} = 45.7 \qquad \frac{294 \times 209}{972} = 63.2$$

$$\frac{294 \times 186}{972} = 56.3 \qquad \frac{294 \times 145}{972} = 43.9 \qquad \frac{294 \times 281}{972} = 85$$

Then the value of χ^2 is calculated as follows:

$$\chi^2 = \sum_{all\ cells} \frac{(f_0 - f_e)^2}{f_e}$$

For the data in Table 18.2, the value of χ^2 is calculated as

$$\chi^2 = \frac{(105 - 95.4)^2}{95.4} + \frac{(130 - 132)^2}{132} + \frac{(115 - 117.5)^2}{117.5} +$$

$$\frac{(92 - 91.6)^2}{91.6} + \frac{(172 - 177.5)^2}{177.5} + \frac{(12 - 9.9)^2}{9.9} +$$

$$\frac{(22 - 13.8)^2}{13.8} + \frac{(7 - 12.2)^2}{12.2} + \frac{(9 - 9.5)^2}{9.5} +$$

$$\frac{(14 - 18.5)^2}{18.5} + \frac{(34 - 45.7)^2}{45.7} + \frac{(57 - 63.2)^2}{63.2} +$$

$$\frac{(64 - 56.3)^2}{56.3} + \frac{(44 - 43.9)^2}{43.9} + \frac{(95 - 85)^2}{85} = 15.8$$

To determine whether a systematic association exists, the probability of obtaining a value of chi-square as large as or larger than the one calculated from the cross-tabulation is estimated. An important characteristic of the chi-square statistic is the number of degrees of freedom (*df*) associated with it. In general, the number of degrees of freedom is equal to the number of observations less the number of constraints needed to calculate a statistical term. In the case of a chi-square statistic associated with a cross-tabulation, the number of degrees of freedom is equal to the product of number of rows (*r*) less one and the number of columns (*c*) less one. That is, $df = (r - 1) \times (c - 1)$.[9] The null hypothesis (H_0) of no association between the two variables will be rejected only when the calculated value of the test statistic is greater than the critical value of the chi-square distribution with the appropriate degrees of freedom, as shown in Figure 18.8.

Chi-square distribution
A skewed distribution whose shape depends solely on the number of degrees of freedom. As the number of degrees of freedom increases, the chi-square distribution becomes more symmetrical.

The **chi-square distribution** is a skewed distribution whose shape depends solely on the number of degrees of freedom.[10] As the number of degrees of freedom increases, the chi-square distribution becomes more symmetrical. Table 3 in the Statistical Appendix contains upper-tail areas of the chi-square distribution for different degrees of freedom. In this table, the value at the top of each column indicates the area in the upper portion (the right side, as shown in Figure 18.8) of the chi-square distribution. To illustrate, for 8 degrees of freedom, the value for an upper-tail area of 0.05 is 15.507. This indicates that for 2 degrees of freedom the probability of exceeding a chi-square value of 15.507 is 0.05. In other words, at the 0.05 level of significance with 8 degrees of freedom, the critical value of the chi-square statistic is 15.507.

For the cross-tabulation given in Table 18.2, there are $(3 - 1) \times ((5 - 1) = 8$ degrees of freedom. The calculated chi-square statistic had a value of 15.8. Since this exceeds the critical value of 15.507, the null hypothesis of no association can be rejected, indicating that the association is statistically significant at the 0.05 level.

The chi-square statistic can also be used in goodness-of-fit tests to determine whether certain models fit the observed data. These tests are conducted by calculating the significance of sample deviations from assumed theoretical (expected) distributions and can be performed on cross-tabulations as well as on frequencies (one-way tabulations). The calculation of the chi-square statistic and the determination of its significance is the same as illustrated above.

The chi-square statistic should be estimated only on counts of data. When the data are in percentage form, they should first be converted to absolute counts or numbers. In addition, an underlying assumption of the chi-square test is that the observations are drawn independently. As a general rule, chi-square analysis should not be conducted when the expected or theoretical frequency in any of the cells is less than five. If the number of observations in any cell is less than 10, or if the table has two rows and two columns (a 2×2 table), a correction factor should be applied.[11] In the case of a 2×2 table, the chi-square is related to the phi coefficient.
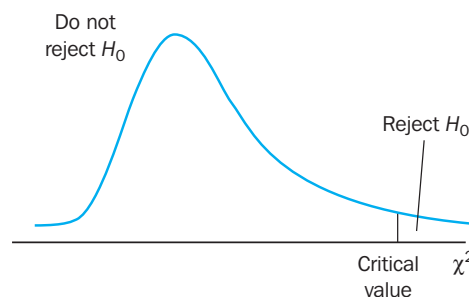
Figure 18.8
Chi-square test of association

Phi coefficient

Phi coefficient (ϕ)
A measure of the strength of association in the special case of a table with two rows and two columns (a 2 × 2 table).

The **phi coefficient** (ϕ) is used as a measure of the strength of association in the special case of a table with two rows and two columns (a 2 × 2 table). The phi coefficient is proportional to the square root of the chi-square statistic. For a sample of size n, this statistic is calculated as

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

It takes the value of 0 when there is no association, which would be indicated by a chi-square value of 0 as well. When the variables are perfectly associated, phi assumes the value of 1 and all the observations fall just on the main or minor diagonal. (In some computer programs, phi assumes a value of –1 rather than +1 when there is perfect negative association.) In the more general case involving a table of any size, the strength of association can be assessed by using the contingency coefficient.

Contingency coefficient

Contingency coefficient
A measure of the strength of association in a table of any size.

Although the phi coefficient is specific to a 2 × 2 table, the **contingency coefficient** (C) can be used to assess the strength of association in a table of any size. This index is also related to chi-square, as follows:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

The contingency coefficient varies between 0 and 1. The 0 value occurs in the case of no association (i.e. the variables are statistically independent), but the maximum value of 1 is never achieved. Rather, the maximum value of the contingency coefficient depends on the size of the table (number of rows and number of columns). For this reason, it should be used only to compare tables of the same size. The value of the contingency coefficient for Table 18.2 is

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

$$= \sqrt{\frac{15.8}{987.8}}$$

$$= 0.1264$$

This value of C indicates that the association is not very strong.

Cramer's V

Cramer's V
A measure of the strength of association used in tables larger than 2 × 2.

Cramer's V is a modified version of the phi correlation coefficient, ϕ, and is used in tables larger than 2 × 2. When phi is calculated for a table larger than 2 × 2, it has no upper limit. Cramer's V is obtained by adjusting phi for either the number of rows or the number of columns in the table based on which of the two is smaller. The adjustment is such that V will range from 0 to 1. A large value of V merely indicates a high degree of association. It does not indicate how the variables are associated. For a table with r rows and c columns, the relationship between Cramer's V and the phi correlation coefficient is expressed as

$$V = \sqrt{\frac{\phi^2}{\min\,(r-1),\,(c-1)}}$$

or

$$V = \sqrt{\frac{\chi^2/n}{\min{(r-1),(c-1)}}}$$

The value of Cramer's V for Table 18.2 is

$$V = \sqrt{\frac{15.8/972}{8}}$$

$$= 0.045$$

Thus, the association is not strong.

Lambda coefficient

The lambda coefficient assumes that the variables are measured on a nominal scale. **Asymmetric lambda** measures the percentage improvement in predicting the value of the dependent variable, given the value of the independent variable. The lambda coefficient also varies between 0 and 1. A value of 0 means no improvement in prediction. A value of 1 indicates that the prediction can be made without error. This happens when each independent variable category is associated with a single category of the dependent variable.

Asymmetric lambda is computed for each of the variables (treating it as the dependent variable). The two asymmetric lambdas are likely to be different, since the marginal distributions are not usually the same. A **symmetric lambda**, a kind of average of the two asymmetric values, is also computed. The symmetric lambda does not make an assumption about which variable is dependent. It measures the overall improvement when prediction is done in both directions.[12]

Other statistics

Note that in the calculation of the chi-square statistic the variables are treated as being measured only on a nominal scale. Other statistics such as **tau *b***, **tau *c***, and **gamma** are available to measure association between two ordinal-level variables. All these statistics use information about the ordering of categories of variables by considering every possible pair of cases in the table. Each pair is examined to determine whether its relative ordering on the first variable is the same as its relative ordering on the second variable (concordant), the ordering is reversed (discordant), or the pair is tied. The manner in which the ties are treated is the basic difference between these statistics. Both tau *b* and tau *c* adjust for ties. Tau *b* is the most appropriate with square tables in which the number of rows and the number of columns are equal. Its value varies between +1 and −1. For a rectangular table in which the number of rows is different from the number of columns, tau *c* should be used. Gamma does not make an adjustment for either ties or table size. Gamma also varies between +1 and −1 and generally has a higher numerical value than tau *b* or tau *c*. Other statistics for measuring the strength of association, namely product moment correlation and non-metric correlation, are discussed in Chapter 20.

Cross-tabulation in practice

While conducting cross-tabulation analysis in practice, it is useful to proceed through the following steps.

■ Test the null hypothesis that there is no association between the variables using the chi-square statistic. If you fail to reject the null hypothesis, then there is no relationship.

Asymmetric lambda
A measure of the percentage improvement in predicting the value of the dependent variable given the value of the independent variable in contingency table analysis. Lambda also varies between 0 and 1.

Symmetric lambda
The symmetric lambda does not make an assumption about which variable is dependent. It measures the overall improvement when prediction is done in both directions.

tau *b*
A test statistic that measures the association between two ordinal-level variables. It makes an adjustment for ties and is the most appropriate when the table of variables is square.

tau *c*
A test statistic that measures the association between two ordinal-level variables. It makes an adjustment for ties and is most appropriate when the table of variables is not square but a rectangle.

Gamma
A test statistic that measures the association between two ordinal-level variables. It does not make an adjustment for ties.

■ If H_0 is rejected, then determine the strength of the association using an appropriate statistic (phi coefficient, contingency coefficient, Cramer's V, lambda coefficient, or other statistics), as discussed earlier.

■ If H_0 is rejected, interpret the pattern of the relationship by computing the percentages in the direction of the independent variable, across the dependent variable.

■ If the variables are treated as ordinal rather than nominal, use tau b, tau c, or gamma as the test statistic. If H_0 is rejected, then determine the strength of the association using the magnitude, and the direction of the relationship using the sign of the test statistic.

Hypothesis testing related to differences

The previous section considered hypothesis testing related to associations. We now focus on hypothesis testing related to differences. A classification of hypothesis testing procedures for examining differences is presented in Figure 18.9. Note that this figure is consistent with the classification of univariate techniques presented in Figure 17.4.

Hypothesis testing procedures can be broadly classified as parametric or non-parametric, based on the measurement scale of the variables involved. **Parametric tests** assume that the variables of interest are measured on at least an interval scale. The most popular parametric test is the t test conducted for examining hypotheses about means. The t test could be conducted on the mean of one sample or two samples of observations. In the case of two samples, the samples could be independent or paired.

Non-parametric tests assume that the variables are measured on a nominal or ordinal scale. These tests can be further classified based on whether one or two or more samples are involved.

Non-parametric tests based on observations drawn from one sample include the chi-square test, the Kolmogorov-Smirnov test, the runs test and the binomial test. In the case of two independent samples, the chi-square test, the Mann-Whitney U test,

Parametric tests
Hypothesis testing procedures that assume that the variables of interest are measured on at least an interval scale.

Non-parametric tests
Hypothesis testing procedures that assume that the variables are measured on a nominal or ordinal scale.
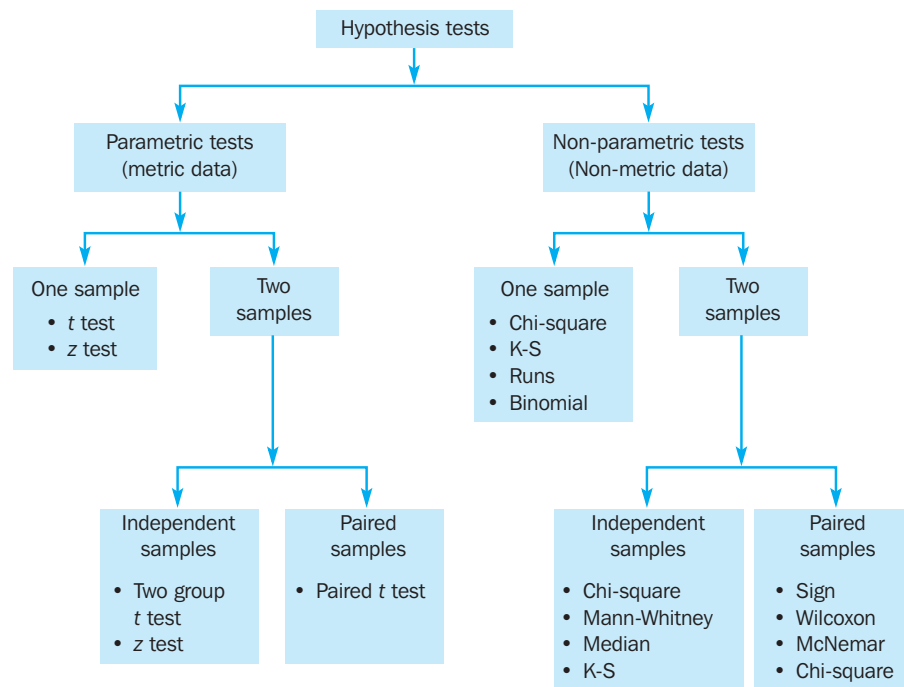
Figure 18.9
Hypothesis testing procedures

the median test and the Kolmogorov-Smirnov two-sample test are used. These tests are non-parametric counterparts of the two-group t test. For paired samples, non-parametric tests include the sign test, the Wilcoxon matched-pairs signed-ranks test, the McNemar test and the chi-square test. These tests are the counterparts of the paired t test. As explained in Chapter 17, the number of samples is determined based on how the data are treated for the purpose of analysis, not based on how the data were collected. The samples are *independent* if they are drawn randomly from different populations. For the purpose of analysis, data pertaining to different groups of respondents, for example males and females, are generally treated as independent samples. On the other hand, the samples are *paired* when the data for the two samples relate to the same group of respondents.

Parametric as well as non-parametric tests are also available for evaluating hypotheses relating to more than two samples. These tests are considered in later chapters.

Parametric tests

t test
A univariate hypothesis test using the t distribution, which is used when the standard deviation is unknown and the sample size is small.

t statistic
A statistic that assumes that the variable has a symmetric bell-shaped distribution, that the mean is known (or assumed to be known), and that the population variance is estimated from the sample.

t distribution
A symmetrical bell-shaped distribution that is useful for sample testing ($n<30$). It is similar to the normal distribution in appearance.

Parametric tests provide inferences for making statements about the means of parent populations. A **t test** is commonly used for this purpose. This test is based on the Student's t statistic. The **t statistic** assumes that the variable is normally distributed and the mean is known (or assumed to be known) and the population variance is estimated from the sample. Assume that the random variable X is normally distributed, with mean μ and unknown population variance σ^2, which is estimated by the sample variance s^2. Recall that the standard deviation of the sample mean, $\overline{X}$, is estimated as $s_{\overline{x}} = s/\sqrt{n}$. Then $t = (\overline{X} - \mu)/s_{\overline{x}}$ is t distributed with $n-1$ degrees of freedom.

The **t distribution** is similar to the normal distribution in appearance. Both distributions are bell-shaped and symmetric. Compared with the normal distribution, however, the t distribution has more area in the tails and less in the centre. This is because the population variance σ^2 is unknown and is estimated by the sample variance s^2. Given the uncertainty in the value of s^2, the observed values of t are more variable than those of z. Thus, we must go out a larger number of standard deviations from zero to encompass a certain percentage of values from the t distribution than is the case with the normal distribution. Yet, as the number of degrees of freedom increases, the t distribution approaches the normal distribution. In fact, for large samples of 120 or more, the t distribution and the normal distribution are virtually indistinguishable. Table 4 in the Statistical Appendix shows selected percentiles of the t distribution. Although normality is assumed, the t test is quite robust to departures from normality.

The procedure for hypothesis testing, for the special case when the t statistic is used, is as follows.

1 Formulate the null (H_0) and the alternative (H_1) hypotheses.
2 Select the appropriate formula for the t statistic.
3 Select a significance level, α, for testing H_0. Typically, the 0.05 level is selected.[13]
4 Take one or two samples and compute the mean and standard deviation for each sample.
5 Calculate the t statistic assuming that H_0 is true.
6 Calculate the degrees of freedom and estimate the probability of getting a more extreme value of the statistic from Table 4 in the Statistical Appendix. (Alternatively, calculate the critical value of the t statistic.)
7 If the probability computed in step 6 is smaller than the significance level selected in step 3, reject H_0. If the probability is larger, do not reject H_0. (Alternatively, if the

value of the calculated t statistic in step 5 is larger than the critical value determined in step 6, reject H_0. If the calculated value is smaller than the critical value, do not reject H_0.) Failure to reject H_0 does not necessarily imply that H_0 is true. It only means that the true state is not significantly different from that assumed by H_0.[14]

8 Express the conclusion reached by the t test in terms of the marketing research problem.

We illustrate the general procedure for conducting t tests in the following sections, beginning with the one-sample case.

One sample

In marketing research, the researcher is often interested in making statements about a single variable against a known or given standard. Examples of such statements are that the market share for a new product will exceed 15%, at least 65% of customers will like a new package design, and 80% of retailers will prefer a new pricing policy. These statements can be translated to null hypotheses that can be tested using a one-sample test, such as the t test or the z test. In the case of a t test for a single mean, the researcher is interested in testing whether the population mean conforms to a given hypothesis (H_0). Suppose that a new machine attachment would be introduced if it receives a mean of at least 7 on a 10-point scale (where 0 = dreadful addition to machine, and 10 = exemplary addition to machine). A sample of 20 engineers is shown the attachment and asked to evaluate it. The results indicate a mean rating of 7.9 with a standard deviation of 1.6. A significance level of $\alpha = 0.05$ is selected. Should the part be introduced?

$$H_0: \mu \leq 7.0$$
$$H_1: \mu > 7.0$$

$$t = \frac{\overline{X} - \mu}{s_{\overline{x}}}$$

$$s_{\overline{x}} = \frac{s}{\sqrt{n}}$$

$$s_{\overline{x}} = \frac{1.6}{\sqrt{20}} = \frac{1.6}{4.472} = 0.358$$

$$t = \frac{7.9 - 7.0}{0.358} = \frac{0.9}{0.358} = 2.514$$

The degrees of freedom for the t statistic to test the hypothesis about one mean are $n - 1$. In this case, $n - 1 = 20 - 1$, or 19. From Table 4 in the Statistical Appendix, the probability of getting a more extreme value than 2.514 is less than 0.05. (Alternatively, the critical t value for 19 degrees of freedom and a significance level of 0.05 is 1.7291, which is less than the calculated value.) Hence, the null hypothesis is rejected, favouring the introduction of the part.

Note that if the population standard deviation was assumed to be known as 1.5, rather than estimated from the sample, a **z test** would be appropriate. In this case, the value of the z statistic would be

z test
A univariate hypothesis test using the standard normal distribution.

$$z = \frac{\overline{X} - \mu}{\sigma_{\overline{x}}}$$

where

$$\sigma_{\overline{x}} = \frac{1.5}{\sqrt{20}} = \frac{1.5}{4.472} = 0.335$$

and

$$z = \frac{7.9 - 7.0}{0.335} = \frac{0.9}{0.335} = 2.687$$

From Table 2 in the Statistical Appendix, the probability of getting a more extreme value of z than 2.687 is less than 0.05. (Alternatively, the critical z value for a one-tailed test and a significance level of 0.05 is 1.645, which is less than the calculated value.) Therefore, the null hypothesis is rejected, reaching the same conclusion arrived at earlier by the t test.

The procedure for testing a null hypothesis with respect to a proportion was illustrated earlier in this chapter when we introduced hypothesis testing.

Two independent samples

Several hypotheses in marketing relate to parameters from two different populations: for example, the users and non-users of a brand differ in terms of their perceptions of the brand, the high-income consumers spend more on leisure activities than low-income consumers, or the proportion of brand-loyal users in segment I is more than the proportion in segment II. Samples drawn randomly from different populations are termed **independent samples**. As in the case for one sample, the hypotheses could relate to means or proportions.

Independent samples
Two samples that are not experimentally related. The measurement of one sample has no effect on the values of the second sample.

Means. In the case of means for two independent samples, the hypotheses take the following form:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

The two populations are sampled and the means and variances are computed based on samples of sizes n_1 and n_2. If both populations are found to have the same variance, a pooled variance estimate is computed from the two sample variances as follows:

$$s^2 = \frac{\sum_{i=1}^{n_1}(X_{i_1} - \overline{X}_1)^2 + \sum_{i=1}^{n_2}(X_{i_2} - \overline{X}_2)^2}{n_1 + n_2 - 2}$$

The standard deviation of the test statistic can be estimated as

$$s_{\overline{X}_1 - \overline{X}_2} = \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

The appropriate value of t can be calculated as

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{s_{\overline{X}_1 - \overline{X}_2}}$$

The degrees of freedom in this case are $(n_1 + n_2 - 2)$.

If the two populations have unequal variances, an exact t cannot be computed for the difference in sample means. Instead, an approximation to t is computed. The number of degrees of freedom in this case is usually not an integer, but a reasonably accurate probability can be obtained by rounding to the nearest integer.[15]

An **F test** of sample variance may be performed if it is not known whether the two populations have equal variance. In this case the hypotheses are:

F test
A statistical test of the equality of the variances of two populations.

$$H_0: \sigma_1^2 = \mu_2^2$$
$$H_1: \sigma_1^2 / \mu_2^2$$

F statistic
The ratio of two sample variances.

The **F statistic** is computed from the sample variances as follows:

$$F_{(n_1-1),(n_2-1)} = \frac{s_1^2}{s_2^2}$$

where n_1 = size of sample 1

 n_2 = size of sample 2

 $n_1 - 1$ = degrees of freedom for sample 1

 $n_2 - 1$ = degrees of freedom for sample 2

 s_1^2 = sample variance for sample 1

 s_2^2 = sample variance for sample 2.

F distribution
A frequency distribution that depends upon two sets of degrees of freedom: the degrees of freedom in the numerator and the degrees of freedom in the denominator.

As can be seen, the critical value of the **F distribution** depends on two sets of degrees of freedom: those in the numerator and those in the denominator. The critical values of F for various degrees of freedom for the numerator and denominator are given in Table 5 of the Statistical Appendix. If the probability of F is greater than the significance level α, H_0 is not rejected and t based on the pooled variance estimate can be used. On the other hand, if the probability of F is less than or equal to α, H_0 is rejected and t based on a separate variance estimate is used. We illustrate this using the GlobalCash example.

example

GlobalCash Project

Electronic banking security

Suppose that decision-makers wanted to understand whether respondents from Germany and the Netherlands who use Citibank for pan-European transactions prefer 'security' to other functions that need to be improved in their electronic banking.

A two independent samples t test was conducted, and the results are presented in Table 18.13. Note that the F test of sample variances has a probability that exceeds 0.05. Accordingly, H_0 cannot be rejected, and the t test based on the pooled variance estimate should be used. The t value is –1.99, and with 265 degrees of freedom this gives a probability of 0.048, which is less than the significance level of 0.05. Therefore, the null hypothesis of equal means is rejected. Since the mean importance of 'user friendliness' for German companies is 3.9778 and for Dutch companies is 4.3712, Dutch companies (who use Citibank for pan-European transactions) attach significantly greater importance to 'security' when seeking improvements to electronic banking functions than German companies. ∎

Table 18.13 **Two independent samples *t* test**

Summary statistics			
	Number of cases	*Mean*	*Standard deviation*
German companies	135	3.9778	1.604
Dutch companies	132	4.3712	1.627
***F* test for equality of variances**			
F value		*Two-tail probability*	
1.03		0.871	
***t* test**			
Pooled variance estimate			*Separate variance estimate*		
t value	*Degrees of freedom*	*Two-tail probability*	*t value*	*Degrees of freedom*	*Two-tail probability*
–1.99	265	0.048	–1.99	264.65	0.048

We also show the t test using a separate variance estimate, since most computer programs automatically conduct the t test both ways. As an application of the t test, consider the following example.

Shops seek to suit elderly to a 't'[16]

A study based on a sample of 789 respondents who were 65 or older attempted to determine the effect of lack of mobility on shop patronage. A major research question related to the differences in the physical requirements of dependent and self-reliant elderly persons. That is, did the two groups require different things to get to the shop or after they arrived at the shop? A more detailed analysis of the physical requirements conducted by the t tests of two independent samples (shown in the table) indicated that dependent elderly persons are more likely to look for shops that offer home delivery and phone orders and for shops to which they have accessible transportation. They are also more likely to look for a variety of shops located close together. ■

Differences in physical requirements between dependent and self-reliant elderly

Physical requirement items	Mean*		
	Self-reliant	Dependent	t test probability
Delivery to home	1.787	2.000	0.023
Phone in order	2.030	2.335	0.003
Transportation to store	2.188	3.098	0.000
Convenient parking	4.001	4.095	0.305
Location close to home	3.177	3.325	0.137
Variety of shops close together	3.456	3.681	0.023

* Measured on a five-point scale from not important (1) to very important (5)

In this example, we tested the difference between means. A similar test is available for testing the difference between proportions for two independent samples.

Proportions. A case involving proportions for two independent samples is illustrated in Table 18.14, which gives the number of companies (employing more than 5,000) that have their own 'in-house' banks in Germany and France.

Table 18.14 Comparing the proportions of German and French companies that have an 'in-house' bank

Sample	Use an 'in-house' bank		
	Have an 'in-house' bank	Do not have an 'in-house' bank	Row totals
Germany	160	40	200
France	120	80	200
Column totals	280	120	

Is the proportion the same in the German and French samples? The null and alternative hypotheses are:

$$H_0: \pi_1 = \pi_2$$
$$H_1: \pi_1 \neq \pi_2$$

473

A z test is used as in testing the proportion for one sample. In this case, however, the test statistic is given by

$$z = \frac{P_1 - P_2}{s_{\bar{P}_1 - \bar{P}_2}}$$

In the test statistic, the numerator is the difference between the proportions in the two samples, P_1 and P_2. The denominator is the standard error of the difference in the two proportions and is given by

$$s_{\bar{P}_1 - \bar{P}_2} = \sqrt{P(1 - P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

where

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

A significance level of $\alpha = 0.05$ is selected. Given the data in Table 18.14, the test statistic can be calculated as

$$P_1 - P_2 = 0.8 - 0.6 = 0.2$$

$$P = \frac{200 \times 0.8 + 200 \times 0.6}{200 + 200} = 0.7$$

$$s_{\bar{P}_1 - \bar{P}_2} = \sqrt{0.7 \times 0.3\left(\frac{1}{200} + \frac{1}{200}\right)} = 0.04583$$

$$z = \frac{0.2}{0.04583} = 4.36$$

Given a two-tail test, the area to the right of the critical value is $\alpha/2$, or 0.025. Hence, the critical value of the test statistic is 1.96. Since the calculated value exceeds the critical value, the null hypothesis is rejected. Thus, the proportion of companies with 'in-house' banks (80% for German companies and 60% for French companies) is significantly different for the two samples.

Paired samples

In many marketing research applications, the observations for the two groups are not selected from independent samples. Rather, the observations relate to **paired samples** in that the two sets of observations relate to the same respondents. A sample of respondents may rate competing brands, may indicate the relative importance of two attributes of a product, or may evaluate a brand at two different times. The differences in these cases are examined by a **paired samples t test**. To compute t for paired samples, the paired difference variable, denoted by D, is formed and its mean and variance calculated. Then the t statistic is computed. The degrees of freedom are $n - 1$, where n is the number of pairs. The relevant formulas are

Paired samples
In hypothesis testing, the observations are paired so that the two sets of observations relate to the same respondents.

Paired samples t test
A test for differences in the means of paired samples.

$$H_0: \mu_D = 0$$
$$H_1: \mu_D \neq 0$$

$$t_{n-1} = \frac{\bar{D} - \mu_D}{\dfrac{s_D}{\sqrt{n}}}$$

where

$$\bar{D} = \frac{\sum\limits_{i=1}^{n} D_i}{n}$$

$$s_D = \sqrt{\frac{\sum\limits_{i=1}^{n} (D_i - \bar{D})^2}{n-1}}$$

The following example illustrates an application of a paired samples t test.

Seconds count[17]

A survey of 83 media directors of advertising agencies was conducted to determine the relative effectiveness of 15-second versus 30-second commercial advertisements. By use of a five-point rating scale (1 being excellent and 5 being poor), 15- and 30-second commercials were rated by each respondent for brand awareness, main idea recall, persuasion, and ability to tell an emotional story. The table indicates that 30-second commercials were rated more favourably on all the dimensions. Paired t tests indicated that these differences were significant, and the 15-second commercials were evaluated as less effective. ■

Mean rating of 15- and 30-second commercials on four communication variables

Brand awareness		Main idea recall		Persuasion		Ability to tell an emotional story	
15	30	15	30	15	30	15	30
2.5	1.9	2.7	2.0	3.7	2.1	4.3	1.9

The difference in proportions for paired samples can be tested by using the McNemar test or the chi-square test, as explained in the following section on non-parametric tests.

Non-parametric tests

Non-parametric tests are used when the variables are non-metric. Like parametric tests, non-parametric tests are available for testing variables from one sample, two independent samples, or two related samples.

One sample

Sometimes the researcher wants to test whether the observations for a particular variable could reasonably have come from a particular distribution, such as the normal, uniform or Poisson distribution. Knowledge of the distribution is necessary for finding probabilities corresponding to known values of the variable, or variable values corresponding to known probabilities (see Appendix 15A). The **Kolmogorov-Smirnov (K-S) one-sample test** is one such goodness-of-fit test. The K-S compares the cumulative distribution function for a variable with a specified distribution. A_i denotes the cumulative relative frequency for each category of the theoretical (assumed) distribution, and O_i denotes the comparable value of the sample frequency. The K-S test is based on the maximum value of the absolute difference between A_i and O_i. The test statistic is

$$K = \max |A_i - O_i|$$

Kolmogorov-Smirnov (K-S) one-sample test
A one-sample non-parametric goodness-of-fit test that compares the cumulative distribution function for a variable with a specified distribution.

The decision to reject the null hypothesis is based on the value of K. The larger K is, the more confidence we have that H_0 is false. Note that this is a one-tailed test, since the value of K is always positive, and we reject H_0 for large values of K. For $\alpha = 0.05$, the critical value of K for large samples (over 35) is given by $1.36/\sqrt{n}$.[18] Alternatively, K can be transformed into a normally distributed z statistic and its associated probability determined.

e x a m p l e

GlobalCash Project

Security deviates from the normal

Suppose that one wanted to test whether the distribution of the importance attached to 'security' in electronic banking functions was normal. A K-S one-sample test is conducted, yielding the data shown in Table 18.15.

The largest absolute difference between the observed and normal distribution was $K = 0.1975$. The critical value for K is $1.36/\sqrt{271} = 0.083$. Since the calculated value of K is larger than the critical value, the null hypothesis is rejected. Alternatively, Table 18.15 indicates that the probability of observing a K value of 0.1975, as determined by the normalised z statistic, is less than 0.001. Since this is less than the significance level of 0.05, the null hypothesis is rejected, leading to the same conclusion. Hence, the distribution of the importance attached to 'security' deviates significantly from the normal distribution. ■

Table 18.15 K-S one-sample test for normality

	Test distribution, normal
Mean	4.19
Standard deviation	1.62
Cases	271

	Most extreme differences			
Absolute	*Positive*	*Negative*	*K-S z*	*Two-tailed p*
0.19754	0.13150	–0.19754	3.253	0.000

As mentioned earlier, the chi-square test can also be performed on a single variable from one sample. In this context, the chi-square serves as a goodness-of-fit test. It tests whether a significant difference exists between the observed number of cases in each category and the expected number.

Other one-sample non-parametric tests include the **runs test** and the **binomial test**. The runs test is a test of randomness for the dichotomous variables. This test is conducted by determining whether the order or sequence in which observations are obtained is random. The binomial test is also a goodness-of-fit test for dichotomous variables. It tests the goodness of fit of the observed number of observations in each category to the number expected under a specified binomial distribution. For more information on these tests, refer to standard statistical literature.[19]

Two independent samples

When the difference in the location of two populations is to be compared based on observations from two independent samples and the variable is measured on an ordinal scale, the **Mann-Whitney U test** can be used.[20] This test corresponds to the two independent sample t test, for interval scale variables, when the variances of the two populations are assumed equal.

In the Mann-Whitney U test, the two samples are combined and the cases are ranked in order of increasing size. The test statistic, U, is computed as the number of

Runs test
A test of randomness for a dichotomous variable.

Binominal test
A goodness-of-fit statistical test for dichotomous variables. It tests the goodness of fit of the observed number of observations in each category to the number expected under a specified binominal distribution.

Mann-Whitney U test
A statistical test for a variable measured on an ordinal scale, comparing the differences in the location of two populations based on observations from two independent samples.

times a score from sample or group 1 precedes a score from group 2. If the samples are from the same population, the distribution of scores from the two groups in the rank list should be random. An extreme value of U would indicate a non-random pattern pointing to the inequality of the two groups. For samples of less than 30, the exact significance level for U is computed. For larger samples, U is transformed into a normally distributed z statistic. This z can be corrected for ties within ranks.

The Dutch seek more security

Since the distribution of importance attached to 'security' was determined to be non-normal, it is appropriate to examine again whether German companies attach different importance to Dutch companies in their views of electronic banking security. This time, though, the Mann-Whitney U test is used. The results are given in Table 18.16. Again, a significant difference is found between the two groups, corroborating the results of the two independent samples t test reported earlier. Since the ranks are assigned from the smallest observation to the largest, the higher mean rank (144.39) of Dutch respondents indicates that they attach greater importance to electronic banking security than German respondents (mean rank = 123.84). ■

Table 18.16 Mann-Whitney *U* Wilcoxon rank sum *W* test: importance of electronic banking security

Mean rank	Cases	
123.84	135	Germany = 1.00
144.39	132	Netherlands = 2.00
	267	Total

U	W	z	Corrected for ties, two-tailed p
7538.00	19060.00	−2.2219	0.0263

Note: U = Mann-Whitney test statistics, W = Wilcoxon W statistic, z = U transformed into a normally distributed z statistic.

Researchers often wish to test for a significant difference in proportions obtained from two independent samples. In this case, as an alternative to the parametric z test considered earlier, one could also use the cross-tabulation procedure to conduct a chi-square test.[21] In this case, we will have a 2×2 table. One variable will be used to denote the sample and will assume a value of 1 for sample 1 and a value of 2 for sample 2. The other variable will be the binary variable of interest.

Two other independent-samples non-parametric tests are the median test and Kolmogorov-Smirnov test. The **two-sample median test** determines whether the two groups are drawn from populations with the same median. It is not as powerful as the Mann-Whitney U test because it merely uses the location of each observation relative to the median, and not the rank, of each observation. The **Kolmogorov-Smirnov (K-S) two-sample test** examines whether the two distributions are the same. It takes into account any differences between the two distributions, including the median, dispersion and skewness, as illustrated by the following example.

Two-sample median test
Non-parametric test statistic that determines whether two groups are drawn from populations with the same median. This is not as powerful as the Mann-Whitney U.

Kolmogorov-Smirnov (K-S) two-sample test
Non-parametric test statistic that determines whether two distributions are the same. It takes into account any differences in the two distributions, including median, dispersion and skewness.

Directors change direction[22]

How do marketing research directors and users in *Fortune 500* manufacturing firms perceive the role of marketing research in initiating changes in marketing strategy formulation? It was found that the marketing research directors were more strongly in favour of initiating changes in strategy and less in favour of holding back than were users of marketing research. Using the Kolmogorov-Smirnov test, these differences of role definition were statistically significant at the 0.05 level, as shown overleaf. ■

The role of marketing research in strategy formulation

Stage/item			Responses (%)					Kolmogorov-Smirnov test
Strategy formulation	*Sample*	*n*	*Absolutely must*	*Preferably should*	*May or may not*	*Preferably should not*	*Absolutely must not*	*Significance*
'Initiate change in the marketing strategy of the firm whenever possible'	D	77	7	26	43	19	5	
	U	68	2	15	32	35	16	0.05

D = directors, U = users.

In this example, the marketing research directors and users comprised two independent samples. The samples, however, are not always independent. In the case of paired samples, a different set of tests should be used.

Paired samples

Wilcoxon matched-pairs signed-ranks test
A non-parametric test that analyses the differences between the paired observations, taking into account the magnitude of the differences.

An important non-parametric test for examining differences in the location of two populations based on paired observations is the **Wilcoxon matched-pairs signed-ranks test**. This test analyses the differences between the paired observations, taking into account the magnitude of the differences. It computes the differences between the pairs of variables and ranks the absolute differences. The next step is to sum the positive and negative ranks. The test statistic, z, is computed from the positive and negative rank sums. Under the null hypothesis of no difference, z is a standard normal variate with mean 0 and variance 1 for large samples. This test corresponds to the paired t test considered earlier.[23]

example
GlobalCash Project

User friendliness more important than security

Respondents were asked to rank the importance of electronic banking criteria such as 'security' and 'user friendliness'. As these were measured on ordinal rather than interval scales, we used the Wilcoxon test. The results are shown in Table 18.17 where a significant difference is found in the variables. There are 177 negative differences (importance attached to security is

There's always someone who doesn't quite understand that a pair means two.

less than that attached to user friendliness). The mean rank of these negative differences is 99.88. On the other hand, there are only 14 positive differences (the importance attached to security exceeds that of user friendliness). The mean rank of these differences is 46.89. There are 80 ties, or observations with the same value for both variables. These numbers indicate that user friendliness is more important than security. Furthermore, the probability associated with the t statistic is less than 0.05, indicating that the difference is indeed significant. ■

Table 18.17 Wilcoxon matched pairs signed-ranks test

Security with user friendliness		
Security–user friendliness	Cases	Mean rank
– Ranks	177	99.88
+ Ranks	14	46.89
Ties	80	
Total	271	
$z = -11.1262$		Two-tailed $p = 0.000$

Sign test
A non-parametric test for examining differences in the location of two populations, based on paired populations, that compares only the signs of the differences between pairs of variables without taking into account the magnitude of the differences.

Another paired sample non-parametric test is the **sign test**.[24] This test is not as powerful as the Wilcoxon matched-pairs signed-ranks test because it only compares the signs of the differences between pairs of variables without taking into account the magnitude of the differences. In the special case of a binary variable where the researcher wishes to test differences in proportions, the McNemar test can be used. Alternatively, the chi-square test can also be used for binary variables.

The various parametric and non-parametric tests are summarised in Table 18.18. The tests in Table 18.18 can be easily related to those in Figure 18.9. Table 18.18 classifies the tests in more detail, as parametric tests (based on metric data) are classified separately for means and proportions. Likewise, non-parametric tests (based on non-metric data) are classified separately for distributions and rankings/medians.

Table 18.18 A summary of hypothesis testing

Sample	Application	Level of scaling	Test/comments
One sample			
One sample	Distributions	Non-metric	K-S and chi-square for goodness of fit Runs test for randomness Binomial test for goodness of fit for dichotomous variables
One sample	Means	Metric	t test, if variance is unknown z test, if variance is known
One sample	Proportions	Metric	z test
Two independent samples			
Two independent samples	Distributions	Non-metric	K-S two-sample test for examining equivalence of two distributions
Two independent samples	Means	Metric	Two-group t test F test for equality of variances
Two independent samples	Proportions	Metric Non-metric	z test Chi-square test
Two independent samples	Rankings/medians	Non-metric	Mann-Whitney U test more powerful than median test

Table 18.18 continued

Paired samples	Applications	Level of scaling	Test/comments
Paired samples	Means	Metric	Paired t test
Paired samples	Proportions	Non-metric	McNemar test for binary variables Chi-square test
Paired samples	Rankings/medians	Non-metric	Wilcoxon matched pairs ranked-signs test more powerful than sign test

Internet and computer applications

SPSS (www.spss.com)

The main program in SPSS is FREQUENCIES. It produces a table of frequency counts, percentages and cumulative percentages for the values of each variable. It gives all of the associated statistics except for the coefficient of variation. If the data are interval scaled and only the summary statistics are desired, the DESCRIPTIVES procedure can be used. All of the statistics computed by DESCRIPTIVES are available in FREQUENCIES. However, DESCRIPTIVES is more efficient because it does not sort values into a frequency table. An additional program, MEANS, computes means and standard deviations for a dependent variable over subgroups of cases defined by independent variables.

CROSSTABS displays cross-classification tables and provides cell counts, row and column percentages, the chi-square test for significance, and all the measures of the strength of the association that have been discussed. The major program for conducting t tests is T-TEST. This program can be used to conduct t tests on independent as well as paired samples. All the non-parametric tests that we have discussed can be conducted by using the NPAR TESTS program.

SAS (www.sas.com)

The main program is UNIVARIATE. In addition to providing a frequency table, this program provides all the associated statistics. Another procedure available is FREQ. For one-way frequency distribution, FREQ does not provide any associated statistics. If only summary statistics are desired, procedures such as MEANS, SUMMARY and TABULATE can be used.

FREQ displays cross-classification tables and provides cell counts, row and column percentages, the chi-square test for significance, and all the measures of the strength of the association that have been discussed. TABULATE can be used for obtaining cell counts and row and column percentages, although it does not provide any of the associated statistics. The program T-TEST can be used. The non-parametric tests may be conducted by using NPAR1WAY. This program will conduct the two independent samples test (Mann-Whitney, median and K-S) as well as the Wilcoxon test for paired samples.

Minitab (www.minitab.com)

The main function is Stats>Descriptive Statistics. The output values include the mean, median, mode, standard deviation, minimum, maximum and quartiles. A histogram in a bar chart or graph can be produced from the Graph>Histogram selection. Cross-tabulations (crosstabs) and chi-square are under the Stats>Tables function. Each of these features must be selected separately under the Tables function. Parametric tests available in Minitab in descriptive stat function are z test for

means, *t* test for means, and two-sample *t* test. The non-parametric tests can be accessed under the Stat>Time Series function. The output includes the one-sample sign, one-sample Wilcoxon, Mann-Whitney, Kruskal-Wallis, median test, Friedman, runs test and pairwise differences.

Excel (www.microsoft.com/catalog/)

In Excel the output produces the mean, standard error, median, mode, standard deviation, variance, kurtosis, skewness, range, minimum, maximum, sum, count and confidence level. Frequencies can be selected under the histogram function. A histogram can be produced in bar format. The Data>Pivot Table performs crosstabs in Excel. To do additional analysis or customise data, select a different summary function such as max, min, average or standard deviation. In addition, a custom calculation can be selected to calculate based on other cells in the data plane. The chi-square test can be accessed under the Insert>Function>Statistical>ChiTest function. The available parametric tests include the *t* test: paired samples for means; *t* test: two independent samples assuming equal variances; *t* test: two independent samples assuming unequal variances; *z* test: two samples for means; and *F* test for variances of two samples.

Snap (www.snapsurveys.com)

In Snap, Descriptive Statistics produces the count, mean, mode, quartiles, median, sum, minimum, maximum, range, standard error, standard deviation, variance, skewness and kurtosis. In either Tables or Charts, the following can be produced displaying absolute values, percentages or means:

- A single question or several questions together
- Comparing one question against another, or several questions against one or more other questions (cross-tabulation)
- Grids of a body of scaled items, e.g. a body of Likert scales
- Holecount tables, i.e. a summary of all or any chosen responses to the questionnaire
- Summary results in counts or percentages in the format of the original questionnaire.

Chi-square and *t* tests can be calculated using the Charts option.

Summary

Basic data analysis provides valuable insights and guides the rest of the data analysis as well as the interpretation of the results. A frequency distribution should be obtained for each variable in the data. This analysis produces a table of frequency counts, percentages and cumulative percentages for all the values associated with that variable. It indicates the extent of out-of-range, missing or extreme values. The mean, mode and median of a frequency distribution are measures of central tendency. The variability of the distribution is described by the range, the variance or standard deviation, coefficient of variation and interquartile range. Skewness and kurtosis provide an idea of the shape of the distribution.

Cross-tabulations are tables that reflect the joint distribution of two or more variables. In cross-tabulation, the percentages can be computed either by column, based on column totals, or by row, based on row totals. The general rule is to compute the per-

centages in the direction of the independent variable, across the dependent variable. Often the introduction of a third variable can provide additional insights. The chi-square statistic provides a test of the statistical significance of the observed association in a cross-tabulation. The phi coefficient, contingency coefficient, Cramer's *V* and lambda coefficient provide measures of the strength of association between the variables.

Parametric and non-parametric tests are available for testing hypotheses related to differences. In the parametric case, the *t* test is used to examine hypotheses related to the population mean. Different forms of the *t* test are suitable for testing hypotheses based on one sample, two independent samples or paired samples. In the non-parametric case, popular one-sample tests include the chi-square, Kolmogorov-Smirnov and binomial tests. For two independent non-parametric samples, the chi-square, Mann-Whitney *U*, median and Kolmogorov-Smirnov tests can be used. For paired samples, the sign, Wilcoxon matched-pairs signed-ranks, McNemar and chi-square tests are useful for examining hypotheses related to measures of location.

Questions

1 Describe the procedure for computing frequencies.

2 What measures of location are commonly computed?

3 What measures of variability are commonly computed?

4 How is the relative flatness or peakedness of a distribution measured?

5 What is a skewed distribution? What does it mean?

6 What is the major difference between cross-tabulation and frequency distribution?

7 What is the general rule for computing percentages in cross-tabulation?

8 Define a spurious correlation.

9 What is meant by a suppressed association? How is it revealed?

10 Discuss the reasons for the frequent use of cross-tabulations. What are some of the limitations?

11 Present a classification of hypothesis testing procedures.

12 Describe the general procedure for conducting a *t* test.

13 What is the major difference between parametric and non-parametric tests?

14 Which non-parametric tests are the counterparts of the two independent samples *t* test for parametric data?

15 Which non-parametric tests are the counterparts of the paired samples *t* test for parametric data?

Notes

1 Dowling, M., 'To go direct – or not', *Catalog Age* 14(9) (1 September 1997), 5; Festervand, T.A., Snyder, D.R. and Tsalikis, J.D., 'Influence of catalog vs. store shopping and prior satisfaction on perceived risk', *Journal of the Academy of Marketing Science* (Winter 1986), 28–36.

2 For an application of frequencies and descriptive statistics, see Bitner, M.J., Booms, B.H. and Mohr, L.A., 'Critical service encounters: the employee's viewpoint', *Journal of Marketing* 58 (October 1994), 95–106.

3 See any introductory statistics book for a more detailed description of these statistics; for example, see Berenson, M.L., Levine, D.M. and Krehbiel, T.C., *Basic Business Statistics: Concepts and Applications*, 8th edn (Englewood Cliffs, NJ: Prentice Hall, 2002).

4 For our purposes, no distinction will be made between formal hypothesis testing and statistical inference by means of confidence intervals.

5 Semon, T.T., 'Let's hear it for disaggregate analysis', *Marketing News* 30(7) (25 March 1996), 10. Excellent discussions of ways to analyse cross-tabulations can be found in Hellevik, O., *Introduction to Causal Analysis: Exploring Survey Data by Crosstabulation* (Beverly Hills, CA: Sage, 1984).

6 Feick, L.F., 'Analyzing marketing research data with association models', *Journal of Marketing Research* 21 (November 1984), 376–86. For an application, see Kamakura, W.A. and Wedel, M., 'Statistical data fusion for cross tabulation', *Journal of Marketing Research* 34(4) (November 1997), 485–98.

7 Lenell, W. and Boissoneau, R., 'Using causal-comparative and correlational designs in conducting market research', *Journal of Professional Services Marketing* 13(2) (1996), 59–69. See also the classic book by Zeisel, H., *Say It with Figures*, 5th edn (New York: Harper & Row, 1968).

8 Sirkin, R.M., *Statistics for the Social Sciences* (Thousand Oaks, CA: Sage, 1994).

9 Pett, M.A., *Nonparametric Statistics for Health Care Research* (Thousand Oaks, CA: Sage, 1997). For a more extensive treatment see Lancaster, H.O., *The Chi Squared Distribution* (New York: Wiley, 1969).

10 Berenson, M.L., Levine, D.M. and Krehbiel, T.C., *Basic Business Statistics: Concepts and Applications*, 8th edn (Englewood Cliffs, NJ: Prentice Hall, 2002).

11 Some statisticians, however, disagree. They feel that a correction should not be applied. See, for example, Overall, J.E., 'Power of chi-square tests for 2×2 contingency tables with small expected frequencies', *Psychological Bulletin* (January 1980), 132–5.

12 Significance tests and confidence intervals are also available for either lambda asymmetric or lambda symmetric. See Goodman, L.A. and Kruskal, W.H., 'Measures of association for cross classification: appropriate sampling theory', *Journal of the American Statistical Association* 88 (June 1963), 310–64.

13 Verma, R. and Goodale, J.C., 'Statistical power in operations management research', *Journal of Operations Management* 13(2) (August 1995), 139–52; Cowles, M. and Davis, C., 'On the origins of the .05 level of statistical significance', *American Psychologist* (May 1982), 553–8. See also Kotabe, M., Duhan, D.E., Smith Jr, D.K. and Wilson, R.D., 'The perceived veracity of PIMS strategy principles in Japan: an empirical inquiry', *Journal of Marketing* 55 (January 1991), 26–41.

14 Technically, a null hypothesis cannot be accepted. It can be either rejected or not rejected. This distinction, however, is inconsequential in marketing research.

15 The condition when the variances cannot be assumed to be equal is known as the Behrens-Fisher problem. There is some controversy over the best procedure in this case.

16 Balazs, A.L., 'Positioning the retail shopping center for aging customers', *Stores* 77(4) (April 1995), RR10–RR11; Lumpkin, J.R. and Hunt, J.B., 'Mobility as an influence on retail patronage behavior of the elderly: testing conventional wisdom', *Journal of the Academy of Marketing Science* (Winter 1989), 1–12.

17 Parpis, E., 'Playing for the ring', *Adweek* (Eastern Edition) 39(3) (19 January 1998), 29–31; Dunst, L., 'Is it possible to get creative in 15 seconds?' *Advertising Age* 64(50) (29 November 1993), 18; Rosenblatt, J.A. and Mainprize, J., 'The history and future of 15-second commercials: an empirical investigation of the perception of ad agency media directors', in Lazer, W., Shaw, E. and Wee, C.H. (eds), *World Marketing Congress,* International Conference Series, Vol. 4 (Boca Raton, FL: Academy of Marketing Science, 1989), 169–77.

18 Kanji, G.K., *100 Statistical Tests* (Thousand Oaks, CA: Sage, 1993); Harnett, D.L., *Statistical Methods*, 3rd edn (Reading, MA: Addison-Wesley, 1982).

19 Pett, M.A., *Nonparametric Statistics for Health Care Research* (Thousand Oaks, CA: Sage, 1997).

20 There is some controversy over whether non-parametric statistical techniques should be used to make inferences about population parameters. See also Cheung, Y.K. and Klotz, J.H., 'The Mann-Whitney-Wilcoxon distribution using linked lists', *Statistica Sinica* 7(3) (July 1997), 805–13.

21 The *t* test in this case is equivalent to a chi-square test for independence in a 2×2 contingency table. The relationship is

$$\chi^2.95(1) = t^2.05(n_1 + n_2 - 2)$$

For large samples, the *t* distribution approaches the normal distribution and so the *t* test and the *z* test are equivalent.

22 Krum, J.R., Rau, P.A. and Keiser, S.K., 'The marketing research process: role perceptions of researchers and users', *Journal of Advertising Research* (December–January 1988), 9–21. See also Miller, C., 'Gallup Brothers analyze the research industry', *Marketing News* 31(1) (6 January 1997), 2.

23 For an example of Wilcoxon matched-pairs signed ranks test, see Kalwani, M.U. and Narayandas, N., 'Long-term manufacturer–supplier relationships: do they pay off for supplier firms?', *Journal of Marketing* 59 (January 1995), 1–16.

24 Pett, M.A., *Nonparametric Statistics for Health Care Research* (Thousand Oaks, CA: Sage, 1997); Field, J.G., 'The world's simplest test of significance', *Journal of the Market Research Society* (July 1971), 170–2.

25 Einspruch, E.L., *An Introductory Guide to SPSS for Windows* (Thousand Oaks, CA: Sage, 1997); Spector, P.E., *SAS Programming for Researchers and Social Scientists* (Thousand Oaks, CA: Sage, 1993); Norat, M.A., 'Software reviews', *Economic Journal: The Journal of the Royal Economic Society* 107 (May 1997), 857–82; Wass, J., 'How statistical software can be assessed', *Scientific Computing and Automation* (October 1966).

Analysis of variance and covariance

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Objectives

After reading this chapter, you should be able to:

1 discuss the scope of the analysis of variance (ANOVA) technique and its relationship to *t* test, and regression;

2 describe one-way analysis of variance, including decomposition of the total variation, measurement of effects significance testing, and interpretation of results;

3 describe *n*-way analysis of variance and the testing of the significance of the overall effect, the interaction effect and the main effect of each factor;

4 describe analysis of covariance and show how it accounts for the influence of uncontrolled independent variables;

5 explain key factors pertaining to the interpretation of results with emphasis on interactions, relative importance of factors and multiple comparisons;

6 discuss specialised ANOVA techniques applicable to marketing, such as repeated measures ANOVA, non-metric analysis of variance, and multivariate analysis of variance (MANOVA).

Analysis of variance is a straightforward way to examine the differences between groups of responses that are measured on interval or ratio scales.

Overview

In Chapter 18, we examined tests of differences between two means or two medians. In this chapter, we discuss procedures for examining differences between more than two means or medians. These procedures are called analysis of variance and analysis of covariance. These procedures have traditionally been used for analysing experimental data, but they are also used for analysing survey or observational data.

We describe analysis of variance and covariance procedures and discuss their relationship to other techniques. Then we describe one-way analysis of variance, the simplest of these procedures, followed by *n*-way analysis of variance and analysis of covariance. Special attention is given to issues in interpretation of results as they relate to interactions, relative importance of factors, and multiple comparisons. Some specialised topics such as repeated measures analysis of variance, non-metric analysis of variance, and multivariate analysis of variance are briefly discussed. We begin with an example illustrating the application of analysis of variance.

e x a m p l e

Antacids are treatment for ANOVA[1]

An investigation was conducted to determine the role of 'verbal content' and 'relative newness of a brand' in the effectiveness of a comparative advertising format, for over-the-counter antacids. The measure of attitude towards the sponsoring brand was the dependent variable. Three factors – advertising format, relative newness and verbal content – were the independent variables, each manipulated at two levels. *Advertising format* was either non-comparative (1st) or comparative (2nd). In the comparative format, well-known brands (Rolaids and Tums) were used for comparison. *Relative newness* was manipulated by changing the brand's sponsor. Alka-Seltzer (1st) was the sponsor in the well-established brand treatment, whereas Acid-Off (2nd) was the sponsor in the new brand condition. The name 'Acid-Off' was chosen based on a pre-test. Verbal content was manipulated to reflect factual (1st) or evaluative content (2nd) in an ad. The subjects were recruited at a shopping centre and randomly assigned to the treatment by an interviewer who was blind to the purpose of the study. A total of 207 responses was collected, 200 of which were usable. Twenty-five respondents were assigned to each of the eight ($2 \times 2 \times 2$) treatments.

A three-way analysis of variance was performed, with attitude as the dependent variable. The overall results were significant. The three-way interaction was also significant. The only two-way interaction that was significant was between ad format and relative newness. A major conclusion from these results was that a comparative format that emphasised factual information was best suited for launching a new brand. ∎

In this example, *t* tests were not appropriate because the effect of each factor was not independent of the effect of other factors (in other words, interactions were significant). Analysis of variance provided a meaningful conclusion in this study.

Relationship among techniques

Analysis of variance and analysis of covariance are used for examining the differences in the mean values of the dependent variable associated with the effect of the controlled independent variables, after taking into account the influence of the uncontrolled independent variables. Essentially, **analysis of variance (ANOVA)** is used as a test of means for two or more populations. The null hypothesis, typically, is that all means are equal. For example, suppose that the researcher was interested in examining whether heavy users, medium users, light users and non-users of yoghurt differed in their preference for Muller yoghurt, measured on a nine-point Likert scale.

Analysis of variance (ANOVA)
A statistical technique for examining the differences among means for two or more populations.

485

The null hypothesis that the four groups were not different in preference for Muller could be tested using analysis of variance.

In its simplest form, analysis of variance must have a dependent variable (preference for Muller yoghurt) that is metric (measured using an interval or ratio scale). There must also be one or more independent variables (product use: heavy, medium, light and non-users). The independent variables must be all categorical (non-metric). Categorical independent variables are also called **factors**. A particular combination of factor levels, or categories, is called a **treatment**. **One-way analysis of variance** (ANOVA) involves only one categorical variable, or a single factor. The differences in preference of heavy users, medium users, light users and non-users would be examined by one-way ANOVA. In one-way analysis of variance, a treatment is the same as a factor level (medium users constitute a treatment). If two or more factors are involved, the analysis is termed **n-way analysis of variance**. If, in addition to product use, the researcher also wanted to examine the preference for Muller yoghurt of customers who are loyal and those who are not, an n-way analysis of variance would be conducted.

If the set of independent variables consists of both categorical and metric variables, the technique is called **analysis of covariance** (ANCOVA). For example, analysis of covariance would be required if the researcher wanted to examine the preference of product use groups and loyalty groups, taking into account the respondents' attitudes towards nutrition and the importance they attached to dairy products. The latter two variables would be measured on nine-point Likert scales. In this case, the categorical independent variables (product use and brand loyalty) are still referred to as factors, whereas the metric-independent variables (attitude towards nutrition and importance attached to dairy products) are referred to as **covariates**.

The relationship of analysis of variance to t tests and other techniques, such as regression (see Chapter 20), is shown in Figure 19.1. These techniques all involve a metric-dependent variable. ANOVA and ANCOVA can include more than one independent variable (product use, brand loyalty, attitude, importance, etc.).

Factors
Categorical independent variables in ANOVA. The independent variables must all be categorical (non-metric) to use ANOVA.

Treatment
In ANOVA, a particular combination of factor levels or categories.

One-way analysis of variance
An ANOVA technique in which there is only one factor.

n-way analysis of variance
An ANOVA model where two or more factors are involved.

Analysis of covariance (ANCOVA)
An advanced analysis of variance procedure in which the effects of one or more metric-scaled extraneous variables are removed from the dependent variable before conducting the ANOVA.

Covariate
A metric-independent variable used in ANCOVA.

Figure 19.1
Relationship between t test, analysis of variance, analysis of covariance and regression

Furthermore, at least one of the independent variables must be categorical, and the categorical variables may have more than two categories (in our example, product use has four categories). A *t* test, on the other hand, involves a single, binary independent variable. For example, the difference in the preferences of loyal and non-loyal respondents could be tested by conducting a *t* test. Regression analysis, like ANOVA and ANCOVA, can also involve more than one independent variable. All the independent variables, however, are generally interval scaled, although binary or categorical variables can be accommodated using dummy variables. For example, the relationship between preference for Muller yoghurt, attitude towards nutrition, and importance attached to dairy products could be examined via regression analysis.

One-way analysis of variance

Marketing researchers are often interested in examining the differences in the mean values of the dependent variable for several categories of a single independent variable or factor. For example:

■ Do various market segments differ in terms of their volume of product consumption?
■ Do brand evaluations of groups exposed to different commercials vary?
■ Do retailers, wholesalers and agents differ in their attitudes towards the firm's distribution policies?
■ How do consumers' intentions to buy the brand vary with different price levels?
■ What is the effect of the types of business customer a company has, upon the number of banks it holds accounts with?

The answer to these and similar questions can be determined by conducting one-way analysis of variance. Before describing the procedure, we define the important statistics associated with one-way analysis of variance.[2]

Retailers, wholesalers and agents seem to differ in their attitudes towards distribution policies – but is the difference significant?

eta² (η^2). The strength of the effects of X (independent variable or factor) on Y (dependent variable) is measured by eta² (η^2). The value of η^2 varies between 0 and 1.

F statistic. The null hypothesis that the category means are equal in the population is tested by an *F statistic* based on the ratio of mean square related to X and mean square related to error.

Mean square. This is the sum of squares divided by the appropriate degrees of freedom.

$SS_{between}$. Also denoted as SS_x, this is the variation in Y related to the variation in the means of the categories of X. This represents variation between the categories of X or the portion of the sum of squares in Y related to X.

SS_{within}. Also denoted as SS_{error}, this is the variation in Y due to the variation within each of the categories of X. This variation is not accounted for by X.

SS_y. This is the total variation in Y.

Conducting one-way analysis of variance

The procedure for conducting one-way analysis of variance is described in Figure 19.2. It involves identifying the dependent and independent variables, decomposing the total variation, measuring the effects, testing significance and interpreting the results. We consider these steps in detail and illustrate them with some applications.

Identifying the dependent and independent variables

The dependent variable is denoted by Y and the independent variable by X, and X is a categorical variable having c categories. There are n observations on Y for each category of X, as shown in Table 19.1. As can be seen, the sample size in each category of X is n, and the total sample size $N = n \times c$. Although the sample sizes in the categories of X (the group sizes) are assumed to be equal for the sake of simplicity, this is not a requirement.

Decomposition of the total variation

In one-way ANOVA, separation of the variation observed in the dependent variable into the variation due to the independent variables plus the variation due to error.

Decomposing the total variation

In examining the differences among means, one-way analysis of variance involves the **decomposition of the total variation** observed in the dependent variable. This variation is measured by the sums of squares corrected for the mean (SS). Analysis of
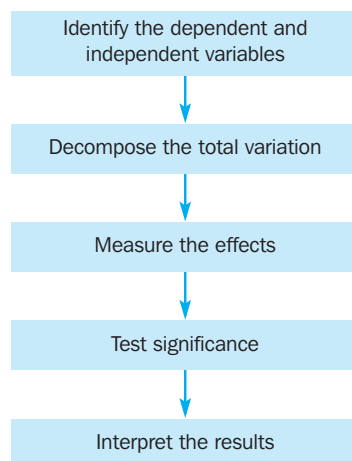
Figure 19.2
Conducting one-way ANOVA

Identify the dependent and independent variables

Decompose the total variation

Measure the effects

Test significance

Interpret the results

Table 19.1 Decomposition of the total variation: one-way ANOVA

variance is so named because it examines the variability or variation in the sample (dependent variable) and, based on the variability, determines whether there is reason to believe that the population means differ.

The total variation in Y, denoted by SS_y, can be decomposed into two components:

$$SS_y = SS_{between} + SS_{within}$$

where the subscripts *between* and *within* refer to the categories of X.[3] $SS_{between}$ is the variation in Y related to the variation in the means of the categories of X. It represents variation between the categories of X. In other words, $SS_{between}$ is the portion of the sum of squares in Y related to the independent variable or factor X. For this reason, $SS_{between}$ is also denoted as SS_x. SS_{within} is the variation in Y related to the variation within each category of X. SS_{within} is not accounted for by X. Therefore, it is referred to as SS_{error}. The total variation in Y may be decomposed as

$$SS_y = SS_x + SS_{error}$$

where $\quad SS_y = \sum_{i=1}^{N}(Y_i - \overline{Y})^2$

$$SS_x = \sum_{j=1}^{c}n(\overline{Y}_j - \overline{Y})^2$$

$$SS_{error} = \sum_{j}^{c}\sum_{i}^{n}(\overline{Y}_{ij} - \overline{Y}_j)^2$$

and Y_i = individual observation

$\quad \overline{Y}_j$ = mean for category j

$\quad \overline{Y}$ = mean over the whole sample or grand mean

$\quad \overline{Y}_{ij}$ = ith observation in the jth category.

The logic of decomposing the total variation in Y, SS_y into $SS_{between}$ and SS_{within} to examine differences in group means can be intuitively understood. Recall from Chapter 18 that, if the variation of the variable in the population was known or estimated, one could estimate how much the sample mean should vary because of random variation alone. In analysis of variance, there are several different groups (e.g. heavy, medium, and light users and non-users). If the null hypothesis is true and all

the groups have the same mean in the population, one can estimate how much the sample means should vary because of sampling (random) variations alone. If the observed variation in the sample means is more than what would be expected by sampling variation, it is reasonable to conclude that this extra variability is related to differences in group means in the population.

In analysis of variance, we estimate two measures of variation: within groups (SS_{within}) and between groups ($SS_{between}$). Within-group variation is a measure of how much the observations, Y values, within a group vary. This is used to estimate the variance within a group in the population. It is assumed that all the groups have the same variation in the population. But because it is not known that all the groups have the same mean, we cannot calculate the variance of all the observations together. The variance for each of the groups must be calculated individually, and these are combined into an 'average' or 'overall' variance. Likewise, another estimate of the variance of the Y values may be obtained by examining the variation between the means. (This process is the reverse of determining the variation in the means, given the population variances.) If the population mean is the same in all the groups, then the variation in the sample means and the sizes of the sample groups can be used to estimate the variance of Y. The reasonableness of this estimate of the Y variance depends on whether the null hypothesis is true. If the null hypothesis is true and the population means are equal, the variance estimate based on between-group variation is correct. On the other hand, if the groups have different means in the population, the variance estimate based on between-group variation will be too large. Thus, by comparing the Y variance estimates based on between-group and within-group variation, we can test the null hypothesis.[3] Decomposition of the total variation in this manner also enables us to measure the effects of X on Y.

Measuring the effects

The effects of X on Y are measured by SS_x. Since SS_x is related to the variation in the means of the categories of X, the relative magnitude of SS_x increases as the differences among the means of Y in the categories of X increase. The relative magnitude of SS_x also increases as the variations in Y within the categories of X decrease. The strength of the effects of X on Y is measured as follows:

$$\eta^2 = \frac{SSx}{SS_y} = \frac{SS_y - SS_{error}}{SSy}$$

The value of η^2 varies between 0 and 1. It assumes a value of 0 when all the category means are equal, indicating that X has no effect on Y. The value of η^2 will be 1 when there is no variability within each category of X but there is some variability between categories. Thus, η^2 is a measure of the variation in Y that is explained by the independent variable X. Not only can we measure the effects of X on Y, but we can also test for their significance.

Testing the significance

In one-way analysis of variance, the interest lies in testing the null hypothesis that the category means are equal in the population.[4] In other words,

$$H_0: \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_c$$

Under the null hypothesis, SS_x and SS_{error} come from the same source of variation. In such a case, the estimate of the population variance of Y can be based on either between-category variation or within-category variation. In other words, the estimate of the population variance of Y,

$$S_y^2 = \frac{SS_x}{c-1}$$

$$= \text{mean square due to } X$$

$$= MS_x$$

or

$$S_y^2 = \frac{SS_{error}}{N-c}$$

$$= \text{mean square due to error}$$

$$= MS_{error}$$

The null hypothesis may be tested by the F statistic based on the ratio between these two estimates:

$$F = \frac{SS_x/(c-1)}{SS_{error}/(N-c)} = \frac{MS_x}{MS_{error}}$$

This statistic follows the F distribution, with $(c-1)$ and $(N-c)$ degrees of freedom (df). A table of the F distribution is given as Table 5 in the Statistical Appendix at the end of the book. As mentioned in Chapter 18, the F distribution is a probability distribution of the ratios of sample variances. It is characterised by degrees of freedom for the numerator and degrees of freedom for the denominator.[5]

Interpreting results

If the null hypothesis of equal category means is not rejected, then the independent variable does not have a significant effect on the dependent variable. On the other hand, if the null hypothesis is rejected, then the effect of the independent variable is significant. In other words, the mean value of the dependent variable will be different for different categories of the independent variable. A comparison of the category mean values will indicate the nature of the effect of the independent variable. Other salient issues in the interpretation of results, such as examination of differences among specific means, are discussed later.

Illustrative applications of one-way analysis of variance

We illustrate the concepts discussed in this section using the data presented in Table 19.2. These data were generated by an experiment in which a bank wanted to examine the effect of direct mail offers and in-branch promotions upon the level of sales of personal loans. In-branch promotion was varied at three levels: high (1), medium (2) and low (3). Direct mail efforts were manipulated at two levels. Either a travel alarm clock was offered to customers who took out a loan (denoted by 1) or it was not (denoted by 2 in Table 19.2). In-branch promotion and direct mail offer were crossed, resulting in a 3×2 design with six cells. Thirty bank branches were randomly selected, and five branches were randomly assigned to each treatment condition. The experiment ran for two months. The sales level of loans were measured, normalised to account for extraneous factors (e.g., branch size, competitive banks within walking distance) and converted to a 1 to 10 scale (10 representing the highest level of sales). In addition, a qualitative assessment was made of the relative affluence of the clientele of each branch, again using a 1 to 10 scale (10 representing the most affluent client base).

To illustrate the concepts of ANOVA, we begin with an example showing calculations done by hand and then by computer. Suppose that only one factor, namely in-branch promotion, was manipulated, that is, let us ignore the direct mail efforts for the purpose of this illustration. The bank is attempting to determine the effect of in-

Table 19.2 Direct mail offer, in-branch promotion, sales of personal loans and clientele rating

Branch number	Direct mail offer	In-branch promotion	Sales	Clientele rating
1	1	1	10	9
2	1	1	9	10
3	1	1	10	8
4	1	1	8	4
5	1	1	9	6
6	1	2	8	8
7	1	2	8	4
8	1	2	7	10
9	1	2	9	6
10	1	2	6	9
11	1	3	5	8
12	1	3	7	9
13	1	3	6	6
14	1	3	4	10
15	1	3	5	4
16	2	1	8	10
17	2	1	9	6
18	2	1	7	8
19	2	1	7	4
20	2	1	6	9
21	2	2	4	6
22	2	2	5	8
23	2	2	5	10
24	2	2	6	4
25	2	2	4	9
26	2	3	2	4
27	2	3	3	6
28	2	3	2	10
29	2	3	1	9
30	2	3	2	8

branch promotion (X) on the sales of personal loans (Y). For the purpose of illustrating hand calculations, the data of Table 19.2 are transformed in Table 19.3 to show the branch (Y_{ij}) for each level of promotion.

The null hypothesis is that the category means are equal:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

To test the null hypothesis, the various sums of squares are computed as follows:

$$
\begin{aligned}
SS_y &= (10 - 6.067)^2 + (9 - 6.067)^2 + (10 - 6.067)^2 + (8 - 6.067)^2 + (9 - 6.067)^2 + (8 - 6.067)^2 \\
&\quad + (9 - 6.067)^2 + (7 - 6.067)^2 + (7 - 6.067)^2 + (6 - 6.067)^2 + (8 - 6.067)^2 + (8 - 6.067)^2 \\
&\quad + (7 - 6.067)^2 + (9 - 6.067)^2 + (6 - 6.067)^2 + (4 - 6.067)^2 + (5 - 6.067)^2 + (5 - 6.067)^2 \\
&\quad + (6 - 6.067)^2 + (4 - 6.067)^2 + (5 - 6.067)^2 + (7 - 6.067)^2 + (6 - 6.067)^2 + (4 - 6.067)^2 \\
&\quad + (5 - 6.067)^2 + (2 - 6.067)^2 + (3 - 6.067)^2 + (2 - 6.067)^2 + (1 - 6.067)^2 + (2 - 6.067)^2 \\
&= 185.867
\end{aligned}
$$

$$
\begin{aligned}
SS_x &= 10(8.3 - 6.067)^2 + 10(6.2 - 6.067)^2 + 10(3.7 - 6.067)^2 \\
&= 106.067
\end{aligned}
$$

Table 19.3 Effect of in-branch promotion on sales of new bank loans

Branch number	Normalised sales level of in-branch promotion		
	High	Medium	Low
1	10	8	5
2	9	8	7
3	10	7	6
4	8	9	4
5	9	6	5
6	8	4	2
7	9	5	3
8	7	5	2
9	7	6	1
10	6	4	2
Column totals	83	62	37
Category means: $\bar{Y}_j$	$\dfrac{83}{10}$ = 8.3	$\dfrac{62}{10}$ = 6.2	$\dfrac{37}{10}$ = 3.7

Grand means: $\bar{Y} = \dfrac{83 + 62 + 37}{30} = 6.067$

$$\begin{aligned} SS_{error} &= (10 - 8.3)^2 + (9 - 8.3)^2 + (10 - 8.3)^2 + (8 - 8.3)^2 + (9 - 8.3)^2 + (8 - 8.3)^2 + (9 - 8.3)^2 \\ &+ (7 - 8.3)^2 + (7 - 8.3)^2 + (6 - 8.3)^2 + (8 - 6.2)^2 + (8 - 6.2)^2 + (7 - 6.2)^2 + (9 - 6.2)^2 \\ &+ (6 - 6.2)^2 + (4 - 6.2)^2 + (5 - 6.2)^2 + (5 - 6.2)^2 + (6 - 6.2)^2 + (4 - 6.2)^2 + (5 - 3.7)^2 \\ &+ (7 - 3.7)^2 + (6 - 3.7)^2 + (4 - 3.7)^2 + (5 - 3.7)^2 + (2 - 3.7)^2 + (3 - 3.7)^2 + (2 - 3.7)^2 \\ &+ (1 - 3.7)^2 + (2 - 3.7)^2 \\ &= 79.8 \end{aligned}$$

It can be verified that

$$SS_y = SS_x + SS_{error}$$

as follows:

$$185.867 = 106.067 + 79.80$$

The strength of the effects of X on Y are measured as follows:

$$\eta^2 = \frac{SS_x}{SS_y}$$

$$= \frac{106.067}{185.897}$$

$$= 0.571$$

In other words, 57.1% of the variation in sales (Y) is accounted for by in-branch promotion (X), indicating a modest effect. The null hypothesis may now be tested.

$$F = \frac{SS_x/(c-1)}{SS_{error}/(N-c)} = \frac{MS_x}{MS_{error}}$$

$$= \frac{106.067/(3-1)}{79.8/(30-3)}$$

$$= 17.944$$

From Table 5 in the Statistical Appendix we see that, for 2 and 27 degrees of freedom, the critical value of F is 3.35 for $\alpha = 0.05$. Because the calculated value of F is greater than the critical value, we reject the null hypothesis. We conclude that the population means for the three levels of in-branch promotion are indeed different. The relative magnitudes of the means for the three categories indicate that a high level of in-branch promotion leads to significantly higher sales of bank loans.

We now illustrate the analysis of variance procedure using a computer program. The results of conducting the same analysis by computer are presented in Table 19.4.

Table 19.4 **One-way ANOVA: effect of in-branch promotion on the sale of bank loans**

Source	df	Sum of squares	Mean square	F ratio	F probability
Between groups (in-branch promotion)	2	106.067	53.033	17.944	0.000
Within groups (error)	27	79.800	2.956		
Total	29	185.867	6.409		

Cell means		
Level of in-branch promotion	Count	Mean
High (1)	10	8.300
Medium (2)	10	6.200
Low (3)	10	3.700
Total	30	6.067

The value of SS_x denoted by main effects is 106.067 with two df; that of SS_{error} (within-group sums of squares) is 79.80 with 27 df. Therefore, $MS_x = 106.067/2 = 53.033$ and $MS_{error} = 79.80/27 = 2.956$. The value of $F = 53.033/2.956 = 17.944$ with 2 and 27 degrees of freedom, resulting in a probability of 0.000. Since the associated probability is less than the significance level of 0.05, the null hypothesis of equal population means is rejected. Alternatively, it can be seen from Table 5 in the Statistical Appendix that the critical value of F for 2 and 27 degrees of freedom is 3.35. Since the calculated value of F (17.944) is larger than the critical value, the null hypothesis is rejected. As can be seen from Table 19.4, the sample means with values of 8.3, 6.2 and 3.7 are quite different.

Assumptions in analysis of variance

The procedure for conducting one-way analysis of variance and the illustrative applications help us understand the assumptions involved. The salient assumptions in analysis of variance can be summarised as follows.

1 Ordinarily, the categories of the independent variable are assumed to be fixed. Inferences are made only to the specific categories considered. This is referred to as the *fixed-effects model*. Other models are also available. In the *random-effects model*, the categories or treatments are considered to be random samples from a universe of treatments. Inferences are made to other categories not examined in the analysis. A *mixed-effects model* results if some treatments are considered fixed and others random.[6]

2 The error term is normally distributed, with a zero mean and a constant variance. The error is not related to any of the categories of X. Modest departures from these assumptions do not seriously affect the validity of the analysis. Furthermore, the data can be transformed to satisfy the assumption of normality or equal variances.

3 The error terms are uncorrelated. If the error terms are correlated (i.e. the observations are not independent), the *F* ratio can be seriously distorted.

In many data analysis situations, these assumptions are reasonably met. Analysis of variance is therefore a common procedure.

N-way analysis of variance

In marketing research, one is often concerned with the effect of more than one factor simultaneously.[7] For example:

- How do consumers' intentions to buy a brand vary with different levels of price and different levels of distribution?
- How do advertising levels (high, medium and low) interact with price levels (high, medium and low) to influence a brand's sale?
- Do income levels (high, medium and low) and age (younger than 35, 35–55, older than 55) affect consumption of a brand?
- What is the effect of consumers' familiarity with a bank (high, medium and low) and bank image (positive, neutral and negative) on preference for taking a loan out with that bank?

Interaction
When assessing the relationship between two variables, an interaction occurs if the effect of X_1 depends on the level of X_2, and vice versa.

In determining such effects, *n*-way analysis of variance can be used. A major advantage of this technique is that it enables the researcher to examine **interactions** between the factors. Interactions occur when the effects of one factor on the dependent variable depend on the level (category) of the other factors. The procedure for conducting *n*-way analysis of variance is similar to that for one-way analysis of variance. The statistics associated with *n*-way analysis of variance are also defined similarly. Consider the simple case of two factors X_1 and X_2 having categories c_1 and c_2. The total variation in this case is partitioned as follows:

$$SS_{total} = SS \text{ due to } X_1 + SS \text{ due to } X_2 + SS \text{ due to interaction of } X_1 \text{ and } X_2 + SS_{within}$$

or

$$SS_y = SS_{x_1} + SS_{x_2} + SS_{x_1 x_2} + SS_{error}$$

A larger effect of X_1 will be reflected in a greater mean difference in the levels of X_1 and a larger SS_{x1}. The same is true for the effect of X_2. The larger the interaction between X_1 and X_2, the larger SS_{x1x2} will be. On the other hand, if X_1 and X_2 are independent, the value of SS_{x1x2} will be close to zero.[8]

Multiple η^2
The strength of the joint effect of two (or more) factors, or the overall effect.

The strength of the joint effect of two factors, called the overall effect, or **multiple η^2**, is measured as follows:

$$\text{multiple } \eta^2 = (SS_{x_1} + SS_{x_2} + SS_{x_1 x_2})/SS_y$$

Significance of the overall effect
A test that some differences exist between some of the treatment groups.

The **significance of the overall effect** may be tested by an *F* test, as follows:

$$F = \frac{(SS_{x_1} + SS_{x_2} + SS_{x_1 x_2})/df_n}{SS_{error}/df_d}$$

$$= \frac{SS_{x_1, x_2, x_1 x_2}/df_n}{SS_{error}/df_d}$$

$$= \frac{MS_{x_1, x_2, x_1 x_2}}{MS_{error}}$$

where df_n = degrees of freedom for the numerator

$$= (c_1 - 1) + (c_2 - 1) + (c_1 - 1)(c_2 - 1)$$
$$= c_1 c_2 - 1$$

df_d = degrees of freedom for the denominator

$$= N - c_1 c_2$$

MS = mean square.

Significance of the interaction effect
A test of the significance of the interaction between two or more independent variables.

If the overall effect is significant, the next step is to examine the **significance of the interaction effect**.[9] Under the null hypothesis of no interaction, the appropriate F test is:

$$F = \frac{SS_{x_1 x_2}/df_n}{SS_{error}/df_d}$$

$$= \frac{MS_{x_1 x_2}}{MS_{error}}$$

where $df_n = (c_1 - 1)(c_2 - 1)$
$df_d = N - c_1 c_2$

Significance of the main effect of each factor
A test of the significance of the main effect for each individual factor.

If the interaction effect is found to be significant, then the effect of X_1 depends on the level of X_2, and vice versa. Since the effect of one factor is not uniform but varies with the level of the other factor, it is not generally meaningful to test the **significance of the main effect of each factor**. It is meaningful to test the significance of each main effect of each factor, if the interaction effect is not significant.[10]

The significance of the main effect of each factor may be tested as follows for X_1:

$$F = \frac{SS_{x_1}/df_n}{SS_{error}/df_d}$$

$$= \frac{MS_{x_1}}{MS_{error}}$$

where $df_n = c_1 - 1$
$df_d = N - c_1 c_2$

The foregoing analysis assumes that the design was orthogonal, or balanced (the number of cases in each cell was the same). If the cell size varies, the analysis becomes more complex.

Returning to the data in Table 19.2, let us now examine the effect of the level of in-branch promotion and direct mail efforts on the sales of personal loans. The results of running a 3×2 ANOVA on the computer are presented in Table 19.5.

For the main effect of level of promotion, the sum of squares SS_{xp}, degrees of freedom, and mean square MS_{xp} are the same as earlier determined in Table 19.4. The sum of squares for direct mail $SS_{xd} = 53.333$ with 1 df, resulting in an identical value for the mean square MS_{xd}. The combined main effect is determined by adding the sum of squares due to the two main effects ($SS_{xp} + SS_{xd} = 106.067 + 53.333 = 159.400$) as well as adding the degrees of freedom ($2 + 1 = 3$). For the promotion and direct mail interaction effect, the sum of squares $SS_{xpxd} = 3.267$ with $(3 - 1) \times (2 - 1) = 2$ degrees of freedom, resulting in $MS_{xpxd} = 3.267/2 = 1.633$. For the overall (model) effect, the sum of squares is the sum of squares for promotion main effect, direct mail

Table 19.5 Two-way analysis of variance

Source of variation	Sum of squares	df	Mean square	F	Sig. of F	ω^2
Main effects						
In-branch promotion	106.067	2	53.033	54.862	0.000	0.557
Direct mail	53.333	1	53.333	55.172	0.000	0.280
Combined	159.400	3	53.133	54.966	0.000	
Two-way interaction	3.267	2	1.633	1.690	0.206	
Model	162.667	5	32.533	33.655	0.000	
Residual (error)	23.200	24	0.967			
Total	185.867	29	6.409			
Cell means						
In-branch promotion	*Direct mail*	*Count*	*Mean*			
High	Yes	5	9.200			
High	No	5	7.400			
Medium	Yes	5	7.600			
Medium	No	5	4.800			
Low	Yes	5	5.400			
Low	No	5	2.000			
Factor level means						
In-branch promotion	*Direct mail*	*Count*	*Mean*			
High		10	8.300			
Medium		10	6.200			
Low		10	3.700			
	Yes	15	7.400			
	No	15	4.733			
Grand mean		30	6.067			

main effect, and interaction effect = 106.067 + 53.333 + 3.267 = 162.667 with 2 + 1 + 2 = 5 degrees of freedom, resulting in a mean square of 162.667/5 = 32.533. Note, however, the error statistics are now different from those in Table 19.4. This is due to the fact that we now have two factors instead of one, SS_{error} = 23.2 with $(30 - (3 \times 2))$ or 24 degrees of freedom resulting in MS_{error} = 23.2/24 = 0.967.

The test statistic for the significance of the overall effect is

$$F = \frac{32.533}{0.967}$$

$$= 33.643$$

with 5 and 24 degrees of freedom, which is significant at the 0.05 level.

The test statistic for the significance of the interaction effect is

$$F = \frac{1.633}{0.967}$$

$$= 1.690$$

with 2 and 24 degrees of freedom, which is not significant at the 0.05 level.

As the interaction effect is not significant, the significance of the main effects can be evaluated. The test statistic for the significance of the main effect of promotion is

$$F = \frac{53.033}{0.967}$$
$$= 54.842$$

with 2 and 24 degrees of freedom, which is significant at the 0.05 level.

The test statistic for the significance of the main effect of direct mail is

$$F = \frac{53.333}{0.967}$$
$$= 55.153$$

with 1 and 24 degrees of freedom, which is significant at the 0.05 level. Thus, higher levels of promotions result in higher sales. The use of a direct mail campaign results in higher sales. The effect of each is independent of the other.

The following example illustrates the use of *n*-way analysis.

example

Country affects TV reception[11]

A study examined the impact of country affiliation on the credibility of product attribute claims for televisions. The dependent variables were the following product-attribute claims: good sound, reliability, crisp-clear picture and stylish design. The independent variables which were manipulated consisted of price, country affiliation and store distribution. A 2 × 2 × 2 between-subjects design was used. Two levels of price, 'low' and 'high', two levels of country affiliation, South Korea and Germany, and two levels of store distribution, Kaufhof and without Kaufhof, were specified.

Data were collected from two shopping centres in a large German city. Thirty respondents were randomly assigned to each of the eight treatment cells for a total of 240 subjects. Table 1 presents the results for manipulations that had significant effects on each of the dependent variables.

Table 1 Analyses for significant manipulations

Effect	Univariate			
	Dependent variable	F	df	p
Country × price	Good sound	7.57	1.232	0.006
Country × price	Reliability	6.57	1.232	0.011
Country × distribution	Crisp-clear picture	6.17	1.232	0.014
Country × distribution	Reliability	6.57	1.232	0.011
Country × distribution	Stylish design	10.31	1.232	0.002

The directions of country-by-distribution interaction effects for the three dependent variables are shown in Table 2. Although the credibility ratings for the crisp-clear picture, reliability and stylish design claims are improved by distributing the Korean-made TV set through Kaufhof rather than some other distributor, the same is not true of a German-made set. Similarly, the directions of country-by-price interaction effects for the two dependent variables are shown in Table 3. At the 'high' price level, the credibility ratings for the 'good sound' and 'reliability' claims are higher for the German-made TV set than for its Korean counterpart, but there is little difference related to country affiliation when the product is at the 'low' price.

This study demonstrates that credibility of attribute claims, for products traditionally exported to Germany by a company in a newly industrialised country, can be significantly improved if the same company distributes the product through a prestigious German retailer

Table 2 Country by distribution interaction means

Country × distribution	Crisp clear picture	Reliability	Stylish design
South Korea			
Kaufhof	3.67	3.42	3.82
Without Kaufhof	3.18	2.88	3.15
Germany			
Kaufhof	3.60	3.47	3.53
Without Kaufhof	3.77	3.65	3.75

Table 3 Country by price interaction means

Country × price	Good sound	Reliability
Low price		
Kaufhof	3.75	3.40
Without Kaufhof	3.53	3.45
High price		
Kaufhof	3.15	2.90
Without Kaufhof	3.73	3.67

and considers making manufacturing investments in Europe. Specifically, three product attribute claims (crisp-clear picture, reliability and stylish design) are perceived as more credible when the TVs are made in South Korea if they are also distributed through a prestigious German retailer. Also, the 'good sound' and 'reliability' claims for TVs are perceived to be more credible for a German-made set sold at a higher price, possibly offsetting the potential disadvantage of higher manufacturing costs in Europe. ■

Analysis of covariance

When examining the differences in the mean values of the dependent variable related to the effect of the controlled independent variables, it is often necessary to take into account the influence of uncontrolled independent variables. For example:

- In determining how consumers' intentions to buy a brand vary with different levels of price, attitude towards the brand may have to be taken into consideration.
- In determining how different groups exposed to different commercials evaluate a brand, it may be necessary to control for prior knowledge.
- In determining how different price levels will affect a household's breakfast cereal consumption, it may be essential to take household size into account.

In such cases, analysis of covariance should be used. Analysis of covariance includes at least one categorical independent variable and at least one interval or metric-independent variable. The categorical independent variable is called a *factor*, whereas the metric-independent variable is called a *covariate*. The most common use of the covariate is to remove extraneous variation from the dependent variable, because the effects of the factors are of major concern. The variation in the dependent variable due to the covariates is removed by an adjustment of the dependent variable's mean value within each treatment condition.

An analysis of variance is then performed on the adjusted scores.[12] The significance of the combined effect of the covariates, as well as the effect of each covariate, is tested by using the appropriate F tests. The coefficients for the covariates provide insights

into the effect that the covariates exert on the dependent variable. Analysis of covariance is most useful when the covariate is linearly related to the dependent variable and is not related to the factors.[13]

Illustrative application of covariance

We again use the data of Table 19.2 to illustrate analysis of covariance. Suppose that we wanted to determine the effect of in-branch promotion and direct mail on sales while controlling for the affluence of clientele. It is felt that the affluence of the clientele may also have an effect on the sales of personal loans (recognising that there may be certain clients who are so affluent that they may never need to take out a loan). The dependent variable consists of loan sales. As before, promotion has three levels and direct mail has two. Clientele affluence is measured on an interval scale and serves as the covariate. The results are shown in Table 19.6.

Table 19.6 Analysis of covariance

Source of variation	Sum of squares	df	Mean square	F	Sig. of F
Covariates					
Clientele	0.838	1	0.838	0.862	0.363
Main effects					
Promotion	106.067	2	53.033	54.546	0.000
Direct mail	53.333	1	53.333	54.855	0.000
Combined	159.400	3	53.133	54.649	0.000
Two-way interaction					
*Promotion*Direct mail*	3.267	2	1.633	1.680	.208
Model	163.505	6	27.251	28.028	.000
Residual (error)	22.362	23	0.972		
Total	185.867	29	6.409		
Covariate	Raw coefficient				
Clientele	−0.078				

As can be seen, the sum of squares attributable to the covariate is very small (0.838) with 1 *df*, resulting in an identical value for the mean square. The associated *F* value is 0.838/0.972 = 0.862, with 1 and 23 degrees of freedom, which is not significant at the 0.05 level. Thus, the conclusion is that the affluence of the clientele does not have an effect on the sales of personal loans. If the effect of the covariate is significant, the sign of the raw coefficient can be used to interpret the direction of the effect on the dependent variable.

Issues in interpretation

Important issues involved in the interpretation of ANOVA results include interactions, relative importance of factors, and multiple comparisons.

Interactions

The different interactions that can arise when conducting ANOVA on two or more factors are shown in Figure 19.3.
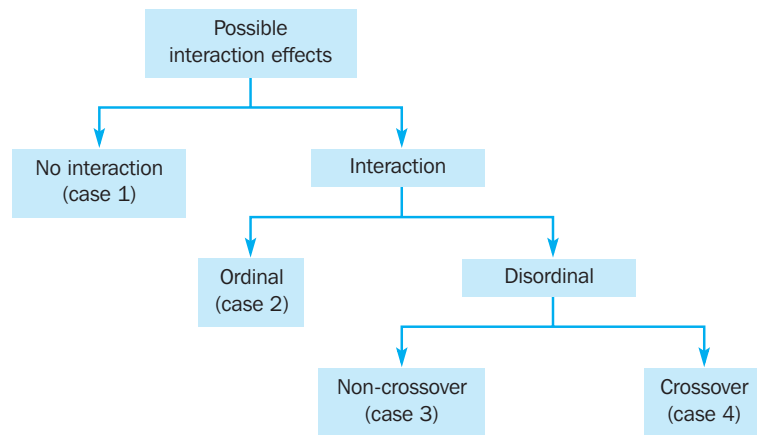
Figure 19.3
A classification of interaction effects

One outcome is that ANOVA may indicate that there are no interactions (the interaction effects are not found to be significant). The other possibility is that the interaction is significant. An interaction effect occurs when the effect of an independent variable on a dependent variable is different for different categories or levels of another independent variable. The interaction may be ordinal or disordinal. In **ordinal interaction**, the rank order of the effects related to one factor does not change across the levels of the second factor. **Disordinal interaction**, on the other hand, involves a change in the rank order of the effects of one factor across the levels of another. If the interaction is disordinal, it could be of a non-crossover or crossover type.[14] These interaction cases are displayed in Figure 19.4, which assumes that there are two factors, X_1 with three levels (X_{11}, X_{12} and X_{13}) and X_2 with two levels (X_{21} and X_{22}). Case 1 depicts no interaction.

The effects of X_1 on Y are parallel over the two levels of X_2. Although there is some departure from parallelism, this is not beyond what might be expected from chance. Parallelism implies that the net effect of X_{22} over X_{21} is the same across the three levels of X_1. In the absence of interaction, the joint effect of X_1 and X_2 is simply the sum of their individual main effects.

Ordinal interaction
An interaction where the rank order of the effects attributable to one factor does not change across the levels of the second factor.

Disordinal interaction
The change in the rank order of the effects of one factor across the levels of another.
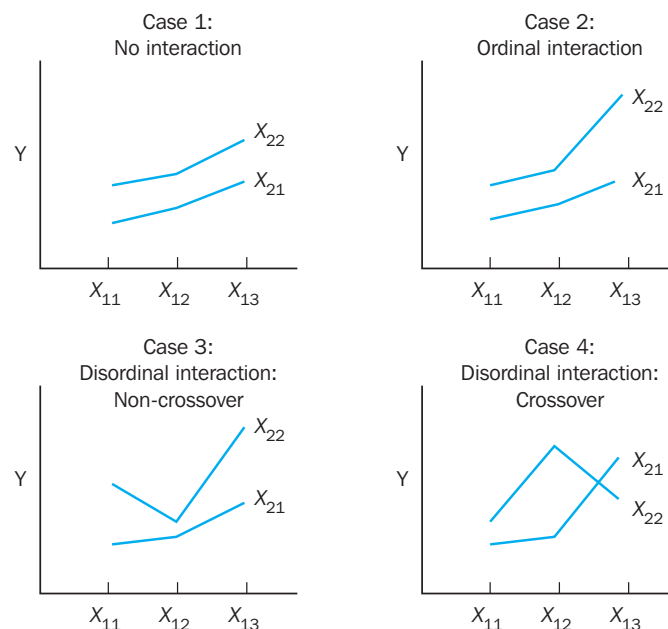
Figure 19.4
Patterns of interaction

Case 2 depicts an ordinal interaction. The line segments depicting the effects of X_1 and X_2 are not parallel. The difference between X_{22} and X_{21} increases as we move from X_{11} to X_{12} and from X_{12} to X_{13}, but the rank order of the effects of X_1 is the same over the two levels of X_2. This rank order, in ascending order, is X_{11}, X_{12}, X_{13}, and it remains the same for X_{21} and X_{22}.

Disordinal interaction of a non-crossover type is displayed by case 3. The lowest effect of X_1 at level X_{21} occurs at X_{11}, and the rank order of effects is X_{11}, X_{12}, X_{13}. At level X_{22}, however, the lowest effect of X_1 occurs at X_{12}, and the rank order is changed to X_{12}, X_{11}, X_{13}. Because it involves a change in rank order, disordinal interaction is stronger than ordinal interaction.

In disordinal interactions of a crossover type, the line segments cross each other, as shown by case 4 in Figure 19.4. In this case, the relative effect of the levels of one factor changes with the levels of the other. Note that X_{22} has a greater effect than X_{21} when the levels of X_1 are X_{11} and X_{12}. When the level of X_1 is X_{13}, the situation is reversed, and X_{21} has a greater effect than X_{22}. (Note that in cases 1, 2 and 3, X_{22} had a greater impact than X_{21} across all three levels of X_1.) Hence, disordinal interactions of a crossover type represent the strongest interactions.[15]

Relative importance of factors

Experimental designs are usually balanced in that each cell contains the same number of respondents. This results in an orthogonal design in which the factors are uncorrelated. Hence, it is possible to determine unambiguously the relative importance of each factor in explaining the variation in the dependent variable.[16] The most commonly used measure in ANOVA is **omega squared, ω^2**. This measure indicates what proportion of the variation in the dependent variable is related to a particular independent variable or factor. The relative contribution of a factor X is calculated as follows:[17]

$$\omega_x^2 = \frac{SS_x - (df_x \times MS_{error})}{SS_{total} + MS_{error}}$$

Normally, ω^2 is interpreted only for statistically significant effects.[18] In Table 19.4, ω^2 associated with level of in-branch promotion is calculated as follows:

$$\omega_p^2 = \frac{106.067 - (2 \times 0.967)}{185.867 + 0.967}$$

$$= \frac{104.133}{186.834}$$

$$= 0.557$$

In Table 19.4 note that:

$$SS_{total} = 106.067 + 53.333 + 3.267 + 23.2$$
$$= 185.867$$

Likewise, the ω^2 associated with direct mail is:

$$\omega_d^2 = \frac{53.333 - (1 \times 0.967)}{185.867 + 0.967}$$

$$= \frac{52.366}{186.834}$$

$$= 0.280$$

As a guide to interpreting ω, a large experimental effect produces an ω^2 of 0.15 or greater, a medium effect produces an index of around 0.06, and a small effect produces an index of 0.01.[19] In Table 19.5, while the effect of promotion and direct mail are both large, the effect of promotion is much larger.

Multiple comparisons

The ANOVA *F* test examines only the overall difference in means. If the null hypothesis of equal means is rejected, we can only conclude that not all the group means are equal. Only some of the means may be statistically different, however, and we may wish to examine differences among specific means. This can be done by specifying appropriate **contrasts,** or comparisons used to determine which of the means are statistically different. Contrasts may be *a priori* or *a posteriori*. **A priori contrasts** are determined before conducting the analysis, based on the researcher's theoretical framework. Generally, *a priori* contrasts are used in lieu of the ANOVA *F* test. The contrasts selected are orthogonal (they are independent in a statistical sense).

A posteriori contrasts are made after the analysis. These are generally **multiple comparison tests.** They enable the researcher to construct generalised confidence intervals that can be used to make pairwise comparisons of all treatment means. These tests, listed in order of decreasing power, include least significant difference, Duncan's multiple range test, Student-Newman-Keuls, Tukey's alternate procedure, honestly significant difference, modified least significant difference, and Scheffe's tests. Of these tests, least significant difference is the most powerful and Scheffe's the most conservative. For further discussion on *a priori* and *a posteriori* contrasts, refer to the literature.[20]

Our discussion so far has assumed that each subject is exposed to only one treatment or experimental condition. Sometimes subjects are exposed to more than one experimental condition, in which case repeated measures ANOVA should be used.

Repeated measures ANOVA

In marketing research, there are often large differences in the background and individual characteristics of respondents. If this source of variability can be separated from treatment effects (effects of the independent variable) and experimental error, then the sensitivity of the experiment can be enhanced. One way of controlling the differences between subjects is by observing each subject under each experimental condition (see Table 19.7).

In this sense, each subject serves as its own control. For example, in a survey attempting to determine differences in evaluations of various airlines, each respondent evaluates all the major competing airlines. In a study examining the differences among heavy users, medium users, light users and non-users of a brand, each respondent provides ratings on the relative importance of each attribute. Because repeated measurements are obtained from each respondent, this design is referred to as within-subjects design or **repeated measures analysis of variance.** This differs from the assumption we made in our earlier discussion that each respondent is exposed to only one treatment condition, also referred to as between-subjects design.[21] Repeated measures analysis of variance may be thought of as an extension of the paired-samples *t* test to the case of more than two related samples.

In the case of a single factor with repeated measures, the total variation, with $nc - 1$ degrees of freedom, may be split into between-people variation and within-people variation.

$$SS_{total} = SS_{between\ people} + SS_{within\ people}$$

Contrasts
In ANOVA, a method of examining differences among two or more means of the treatment groups.

A priori contrasts
Contrasts determined before conducting the analysis, based on the researcher's theoretical framework.

A posteriori contrasts
Contrasts made after conducting the analysis. These are generally multiple comparison tests.

Multiple comparison tests
A posteriori contrasts that enable the researcher to construct generalised confidence intervals that can be used to make pairwise comparisons of all treatment means.

Repeated measures analysis of variance
An ANOVA technique used when respondents are exposed to more than one treatment condition and repeated measurements are obtained.

Table 19.7 **Decomposition of the total variation: repeated measures ANOVA**

The between-people variation, which is related to the differences between the means of people, has $n - 1$ degrees of freedom. The within-people variation has $n(c - 1)$ degrees of freedom. The within-people variation may, in turn, be divided into two different sources of variation. One source is related to the differences between treatment means, and the second consists of residual or error variation. The degrees of freedom corresponding to the treatment variation are $c - 1$ and that corresponding to residual variation are $(c - 1)(n - 1)$. Thus,

$$SS_{within\ people} = SS_x + SS_{error}$$

A test of the null hypothesis of equal means may now be constructed in the usual way:

$$F = \frac{SS_x/(c - 1)}{SS_{error}/(n - 1(c - 1))}$$

$$= \frac{MS_x}{MS_{error}}$$

So far we have assumed that the dependent variable is measured on an interval or ratio scale. If the dependent variable is non-metric, however, a different procedure should be used.

Non-metric analysis of variance
An ANOVA technique for examining the difference in the central tendencies of more than two groups when the dependent variable is measured on an ordinal scale.

k-sample median test
A non-parametric test used to examine differences among more than two groups when the dependent variable is measured on an ordinal scale.

Kruskal-Wallis one-way analysis of variance
A non-metric ANOVA test that uses the rank value of each case, not merely its location relative to the median.

Non-metric analysis of variance

Non-metric analysis of variance examines the difference in the central tendencies of more than two groups when the dependent variable is measured on an ordinal scale. One such procedure is the **k-sample median test**. As its name implies, this is an extension of the median test for two groups, which was considered in Chapter 18. The null hypothesis is that the medians of the k populations are equal. The test involves the computation of a common median over the k samples. Then, a $2 \times k$ table of cell counts based on cases above or below the common median is generated. A chi-square statistic is computed. The significance of the chi-square implies a rejection of the null hypothesis.

A more powerful test is the **Kruskal-Wallis one-way analysis of variance**. This is an extension of the Mann-Whitney test (Chapter 18). This test also examines the difference in medians. The null hypothesis is the same as in the k-sample median test,

but the testing procedure is different. All cases from the k groups are ordered in a single ranking. If the k populations are the same, the groups should be similar in terms of ranks within each group. The rank sum is calculated for each group. From these, the Kruskal-Wallis H statistic, which has a chi-square distribution, is computed.

The Kruskal-Wallis test is more powerful than the k-sample median test because it uses the rank value of each case, not merely its location relative to the median. If there are a large number of tied rankings in the data, however, the k-sample median test may be a better choice.

Non-metric analysis of variance is not popular in marketing research. Another procedure that is also only rarely used is multivariate analysis of variance.

Multivariate analysis of variance

Multivariate analysis of variance (MANOVA)
An ANOVA technique using two or more metric dependent variables.

Multivariate analysis of variance (MANOVA) is similar to analysis of variance (ANOVA) except that instead of one metric-dependent variable we have two or more. The objective is the same, since MANOVA is also concerned with examining differences between groups. Although ANOVA examines group differences on a single dependent variable, MANOVA examines group differences across multiple dependent variables simultaneously. In ANOVA, the null hypothesis is that the means of the dependent variable are equal across the groups. In MANOVA, the null hypothesis is that the vector of the means of multiple dependent variables is equal across groups. Multivariate analysis of variance is appropriate when there are two or more dependent variables that are correlated. If there are multiple dependent variables that are uncorrelated or orthogonal, ANOVA on each of the dependent variables is more appropriate than MANOVA.[22]

As an example, suppose that four groups, each consisting of 100 randomly selected individuals, were exposed to four different commercials about the 'Series 3' BMW. After seeing the commercial, each individual provided ratings on preference for the 'Series 3', preference for BMW, and preference for the commercial itself. Because these three preference variables are correlated, multivariate analysis of variance should be conducted to determine which commercial is the most effective (produced the highest preference across the three preference variables). The following example illustrates the application of ANOVA and MANOVA in international marketing research.

e x a m p l e

The commonality of unethical research practices worldwide[23]

A study examined marketing professionals' perceptions of how common unethical practices in marketing research were across different countries, i.e. 'the commonality of unethical marketing research practices'. A sample of marketing professionals was drawn from Australia, Britain, Canada and the United States.

Respondents' evaluations were analysed using MANOVA and ANOVA techniques. The predictor variable was the 'country of respondent' and 15 evaluations of 'commonality' served as the criterion variables. The F values from the ANOVA analyses indicated that only two of the 15 commonality evaluations achieved significance ($p < 0.05$ or better). Further, the MANOVA F value was not statistically significant, implying the lack of overall differences in commonality evaluations across respondents of the four countries. It was concluded that marketing professionals in the four countries demonstrate similar perceptions of the commonality of unethical research practices. This finding is not surprising, given other research evidence that organisations in the four countries reflect similar corporate cultures. ■

Internet and computer applications

The computer packages SPSS and SAS have programs for conducting analysis of variance and covariance. In addition to the basic analysis that we have considered, these programs can also perform more complex analysis. Minitab and Excel also offer some programs. Given the importance of analysis of variance and covariance, several programs are available in each package.

SPSS

One-way ANOVA can be efficiently performed using the program ONEWAY. This program also allows the user to test *a priori* and *a posteriori* contrasts. For performing *n*-way analysis of variance, the program ANOVA can be used. Although covariates can be specified, ANOVA does not perform a full analysis of covariance. For comprehensive analysis of variance or analysis of covariance, including repeated measures and multiple dependent measures, the MANOVA procedure is recommended. For non-metric analysis of variance, including the *k*-sample median test and Kruskal-Wallis one-way analysis of variance, the program NPAR TESTS should be used.

SAS

The main program for performing analysis of variance in the case of a balanced design is ANOVA. This program can handle data from a wide variety of experimental designs, including multivariate analysis of variance and repeated measures. Both *a priori* and *a posteriori* contrasts can be tested. For unbalanced designs, the more general GLM procedure can be used. This program performs analysis of variance, analysis of covariance, repeated measures analysis of variance, and multivariate analysis of variance. It also allows the testing of *a priori* and *a posteriori* contrasts. Whereas GLM can also be used for analysing balanced designs, it is not as efficient as ANOVA for such models. The VARCOMP procedure computes variance components. For non-metric analysis of variance, the NPAR1WAY procedure can be used. For constructing designs and randomised plans, the PLAN procedure can be used.

Minitab

Analysis of variance and covariance can be accessed from the Stats>ANOVA function. This function performs one way ANOVA, one-way unstacked ANOVA, two-way ANOVA, analysis of means, balanced ANOVA, analysis of covariance, general linear model, main effects plot, interactions plot and residual plots. In order to compute the mean and standard deviation, the crosstab function must be used. To obtain *F* and *p* values, use the balanced ANOVA.

Excel

Both a one-way ANOVA and two-way ANOVA can be performed under the Tools>Data Analysis function. The two-way ANOVA has the features of a two-factor with replication and a two-factor without replication. The two-factor with replication includes more than one sample for each group of data, while the two-factor without replication does not include more than one sampling per group.

Summary

In ANOVA and ANCOVA, the dependent variable is metric and the independent variables are all categorical and metric variables. One-way ANOVA involves a single independent categorical variable. Interest lies in testing the null hypothesis that the category means are equal in the population. The total variation in the dependent variable may be decomposed into two components: variation related to the independent variable and variation related to error. The variation is measured in terms of the sum of squares corrected for the mean (SS). The mean square is obtained by dividing the SS by the corresponding degrees of freedom (df). The null hypothesis of equal means is tested by an F statistic, which is the ratio of the mean square related to the independent variable to the mean square related to error.

N-way analysis of variance involves the simultaneous examination of two or more categorical independent variables. A major advantage is that the interactions between the independent variables can be examined. The significance of the overall effect, interaction terms, and the main effects of individual factors are examined by appropriate F tests. It is meaningful to the significance of main effects only if the corresponding interaction terms are not significant.

ANCOVA includes at least one categorical independent variable and at least one interval or metric-independent variable. The metric-independent variable, or covariate, is commonly used to remove extraneous variation from the dependent variable.

When analysis of variance is conducted on two or more factors, interactions can arise. An interaction occurs when the effect of an independent variable on a dependent variable is different for different categories or levels of another independent variable. If the interaction is significant, it may be ordinal or disordinal. Disordinal interaction may be of a non-crossover or crossover type. In balanced designs, the relative importance of factors in explaining the variation in the dependent variable is measured by omega squared (ω^2). Multiple comparisons in the form of *a priori* or *a posteriori* contrasts can be used for examining differences among specific means.

In repeated measures analysis of variance, observations on each subject are obtained under each treatment condition. This design is useful for controlling for the differences in subjects that exist prior to the experiment. Non-metric analysis of variance involves examining the differences in the central tendencies of two or more groups when the dependent variable is measured on an ordinal scale. Multivariate analysis of variance (MANOVA) involves two or more metric dependent variables.

Questions

1 Discuss the similarities and differences between analysis of variance and analysis of covariance.

2 What is the relationship between analysis of variance and the *t* test?

3 What is total variation? How is it decomposed in a one-way analysis of variance?

4 What is the null hypothesis in one-way ANOVA? What basic statistic is used to test the null hypothesis in one-way ANOVA? How is this statistic computed?

5 How does *n*-way analysis of variance differ from the one-way procedure?

6 How is the total variation decomposed in *n*-way analysis of variance?

7 What is the most common use of the covariate in ANCOVA?

8 What is the difference between ordinal and disordinal interaction?

9 How is the relative importance of factors measured in a balanced design?

10 What is an *a priori* contrast?

11 What is the most powerful test for making *a posteriori* contrasts? Which test is the most conservative?

12 What is meant by repeated measures ANOVA? Describe the decomposition of variation in repeated measures ANOVA.

13 What are the differences between metric and non-metric analyses of variance?

14 Describe two tests used for examining differences in central tendencies in non-metric ANOVA.

15 What is multivariate analysis of variance? When is it appropriate?

Notes

1 Wilke, M., 'Health reports in vogue again for drug advertisers', *Advertising Age* 68(33) (18 August 1997), 31; Iyer, E.S., 'The influence of verbal content and relative newness on the effectiveness of comparative advertising', *Journal of Advertising* 17, (1988), 15–21.

2 For applications of ANOVA, see Varki, S. and Rust, R.T., 'Satisfaction is relative', *Marketing Research: A Magazine of Management and Applications* 9(2) (Summer 1997), 14–19; Deshpande, R. and Stayman, D.M., 'A tale of two cities: distinctiveness theory and advertising effectiveness', *Journal of Marketing Research* 31 (February 1994), 57–64.

3 Wright, D.B., *Understanding Statistics* (Thousand Oaks, CA: Sage, 1993); Norusis, M.J., *The SPSS Guide to Data Analysis for SPSS/PC+* (Chicago, IL: SPSS Inc., 1991).

4 Driscoll, W.C., 'Robustness of the ANOVA and Tukey-Kramer statistical tests', *Computers and Industrial Engineering* 31(1,2) (October 1996), 265–8; Burdick, R.K., 'Statement of hypotheses in the analysis of variance', *Journal of Marketing Research* (August 1983), 320–4.

5 The *F* test is a generalised form of the *t* test. If a random variable is *t* distributed with *N* degrees of freedom, then t^2 is *F* distributed with 1 and *N* degrees of freedom. Where there are two factor levels or treatments, ANOVA is equivalent to the two-sided *t* test.

6 Although computations for the fixed-effects and random-effects models are similar, interpretations of results differ. A comparison of these approaches is found in Erez, A., Bloom, M.C. and Wells, M.T., 'Using random rather than fixed effects models in meta-analysis: implications for situational specificity and validity generalization', *Personnel Psychology* 49(2) (Summer 1996), 275–306; Neter, J.W., *Applied Linear Statistical Models*, 4th edn (Burr Ridge, IL: Irwin, 1996).

7 We consider only the full factorial designs, which incorporate all possible combinations of factor levels. For example, see Menon, G., 'Are the parts better than the whole? The effects of decompositional questions on judgments of frequent behaviors', *Journal of Marketing Research* 34 (August 1997) 335–46.

8 Jaccard, J., *Interaction Effects in Factorial Analysis of Variance* (Thousand Oaks, CA: Sage, 1997); Mayers, J.L., *Fundamentals of Experimental Design*, 3rd edn (Boston, MA: Allyn & Bacon, 1979). See also Spence, M.T. and Brucks, M., 'The moderating effects of problem characteristics on experts' and novices'

judgments', *Journal of Marketing Research* 34 (February 1997), 233–47.

9 Tacq, J., *Multivariate Analysis Techniques in Social Science Research* (Thousand Oaks, CA: Sage, 1997); Daniel, W.W. and Terrell, J.C., *Business Statistics*, 7th edn (Boston, MA: Houghton Mifflin, 1995).

10 See Jaccard, J., *Interaction Effects in Factorial Analysis of Variance* (Thousand Oaks, CA: Sage, 1997).

11 Peterson, R.A. and Jolibert, A.J.P., 'A meta-analysis of country-of-origin effects', *Journal of International Business Studies* 26(4) (Fourth Quarter 1995), 883–900; Chao, P., 'The impact of country affiliation on the credibility of product attribute claims', *Journal of Advertising Research* (April–May 1989), 35–41.

12 Although this is the most common way in which analysis of covariance is performed, other situations are also possible. For example, covariate and factor effects may be of equal interest, or the set of covariates may be of major concern. For a recent application, see Lane Keller, K. and Aaker, D.A., 'The effects of sequential introduction of brand extensions', *Journal of Marketing Research* 29 (February 1992), 35–50.

13 For a more detailed discussion, see Neter, J.W., *Applied Linear Statistical Models*, 4th edn (Burr Ridge, IL: Irwin, 1996); Wildt, A.R. and Ahtola, O.T., *Analysis of Covariance* (Beverly Hills, CA: Sage, 1978).

14 See Umesh, U.N., Peterson, R.A., McCann-Nelson, M. and Vaidyanathan, R., 'Type IV error in marketing research: the investigation of ANOVA interactions', *Journal of the Academy of Marketing Science* 24(1) (Winter 1966), 17–26; Ross Jr., W. T. and Creyer, E.H., 'Interpreting interactions: raw means or residual means', *Journal of Consumer Research* 20(2) (September 1993), 330–8; Leigh, J.H. and Kinnear, T.C., 'On interaction classification', *Educational and Psychological Measurement* 40 (Winter 1980), 841–3.

15 For an examination of interactions using an ANOVA framework, see Jaccard, J., *Interaction Effects in Factorial Analysis of Variance* (Thousand Oaks, CA: Sage, 1997); Wansink, B., 'Advertising's impact on category substitution', *Journal of Marketing Research* 31 (November 1994), 505–15; Peracchio, L.A. and Meyers-Levy, J., 'How ambiguous cropped objects in ad photos can affect product evaluations', *Journal of Consumer Research* 21 (June 1994), 190–204.

16 Verma, R. and Goodale, J.C., 'Statistical power in operations management', *Journal of Operations Management* 13(2) (August 1995) 139–52; Wyner, G.A., 'The significance of marketing research', *Marketing Research: A Magazine of Management and Applications* 5(1) (Winter 1993), 43–5; Sawyer, A. and Peter, J.P., 'The significance of statistical significance tests in marketing research', *Journal of Marketing Research* 20 (May 1983), 125; Beltramini, R.F., 'A meta-analysis of effect sizes in consumer behavior experiments', *Journal of Consumer Research* 12 (June 1985), 97–103.

17 This formula does not hold if repeated measurements are made on the dependent variable. See Fern, E.F. and Monroe, K.B., 'Effect-size estimates: issues and problems in interpretation', *Journal of Consumer Research* 23(2) (September 1996), 89–105; Dodd, D.H. and Schultz Jr, R.E., 'Computational procedures for estimating magnitude of effect for some analysis of variance designs', *Psychological Bulletin* (June 1973), 391–5.

18 The ω^2 formula is attributed to Hays. See Hays, W.L., *Statistics for Psychologists* (New York: Holt, Rinehart & Winston, 1963). For an application, see Ratneshwar, S. and Chaiken, S., 'Comprehension's role in persuasion: the case of its moderating effect on the persuasive impact of source cues', *Journal of Consumer Research* 18 (June 1991), 52–62. For an alternative approach, see also Finn, A. and Kayande, U., 'Reliability assessment and optimisation of marketing measurement', *Journal of Marketing Research* 34 (February 1997) 262–75.

19 Fern, E.F. and Monroe, K.B., 'Effect-size estimates: issues and problems in interpretation', *Journal of Consumer Research* 23(2) (September 1996), 89–105; Cohen, J., *Statistical Power Analysis for the Behavioral Sciences* (New York: Academic Press, 1969).

20 Neter, J.W., *Applied Linear Statistical Models*, 4th edn (Burr Ridge, IL: Irwin, 1996); Winer, B.J., Brown, D.R. and Michels, K.M., *Statistical Principles in Experimental Design*, 3rd edn (New York: McGraw-Hill, 1991).

21 It is possible to combine between-subjects and within-subjects factors in a single design. See, for example, Mount, M.K., Sytsma, M.A., Hazucha, J.F. and Holt, K.E., 'Rater-ratee effects in developmental performance ratings of managers', *Personnel Psychology* 50(1) (Spring 1997), 51–69; Broniarczyk, S.M. and Alba, J.W., 'The importance of the brand in brand extension', *Journal of Marketing Research* 31 (May 1994), 214–28; Krishna, A., 'The effect of deal knowledge on consumer purchase behavior', *Journal of Marketing Research* 31 (February 1994), 76–91.

22 See Novak, N.P., 'MANOVAMAP: geographical representation of MANOVA in marketing research', *Journal of Marketing Research* 32(3) (August 1995), 357–74; Bray, J.H. and Maxwell, S.E., *Multivariate Analysis of Variance* (Beverly Hills, CA: Sage, 1985). For an application of MANOVA, see Varki, S., 'Satisfaction is relative', *Marketing Research: A Magazine of Management and Applications* 9(2) (Summer 1997), 14–19.

23 Abramson, N.R., Keating, R.J. and Lane, H.W., 'Cross-national cognitive process differences: a comparison of Canadian, American and Japanese managers', *Management International Review* 36(2) (Second Quarter 1996), 123–47; Akaah, I.P., 'A cross-national analysis of the perceived commonality of unethical practices in marketing research', in Lazer, L., Shaw, E. and Wee, C-H. (eds), *World Marketing Congress*, International Conference Series, Vol. 4 (Boca Raton, FL: Academy of Marketing Science, 1989), 2–9.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Correlation and regression

Objectives

After reading this chapter, you should be able to:

1 discuss the concepts of product moment correlation, partial correlation and part correlation, and show how they provide a foundation for regression analysis;

2 explain the nature and methods of bivariate regression analysis and describe the general model, estimation of parameters, standardised regression coefficient, significance testing, prediction accuracy, residual analysis and model cross-validation;

3 explain the nature and methods of multiple regression analysis and the meaning of partial regression coefficients;

4 describe specialised techniques used in multiple regression analysis, particularly stepwise regression, regression with dummy variables, and analysis of variance and covariance with regression;

5 discuss non-metric correlation and measures such as Spearman's rho and Kendall's tau.

Correlation is the simplest way to understand the association between two metric variables. When extended to multiple regression, the relationship between one variable and several others becomes more clear.

Overview

Chapter 19 examined the relationship among the *t* test, analysis of variance and covariance, and regression. This chapter describes regression analysis, which is widely used for explaining variation in market share, sales, brand preference and other marketing results. This is done in terms of marketing management variables such as advertising, price, distribution and product quality. Before discussing regression, however, we describe the concepts of product moment correlation and partial correlation coefficient, which lay the conceptual foundation for regression analysis.

In introducing regression analysis, we discuss the simple bivariate case first. We describe estimation, standardisation of the regression coefficients, and testing and examination of the strength and significance of association between variables, prediction accuracy, and the assumptions underlying the regression model. Next, we discuss the multiple regression model, emphasising the interpretation of parameters, strength of association, significance tests and examination of residuals.

We then cover topics of special interest in regression analysis, such as stepwise regression, multicollinearity, relative importance of predictor variables, and cross-validation. We describe regression with dummy variables and the use of this procedure to conduct analysis of variance and covariance. We begin with two examples that illustrate applications of regression analysis.

e x a m p l e

GlobalCash Project

Multiple regression

In the GlobalCash Project, multiple regression analysis was used to develop a model that explained 'bank preference' in terms of respondents' evaluations of the banks in their own countries, through four choice criteria. The dependent variable was the preference for individual banks. The independent variables were the evaluations of each bank on balance reporting, domestic payments and collections, international payments and collections, and managing currencies. The results indicated that all the factors of the choice criteria, except managing currencies, were significant in explaining bank preference. The coefficients of all the variables were positive, indicating that higher evaluations on each of the significant factors led to higher preference for that bank. The model had a good fit and good ability to predict bank preference. ■

e x a m p l e

Regression rings the right bell for Avon[1]

Avon Products were having significant problems with their sales staff. The company's business, dependent on sales representatives, was facing a shortage of sales representatives without much hope of getting new ones. Regression models were developed to reveal the possible variables that were fuelling this situation. The models revealed that the most significant variable was the level of the appointment fee that reps paid for materials. With data to back up its actions, the company lowered the fee. This resulted in an improvement in the recruitment and retention of sales reps. ■

These examples illustrate some of the uses of regression analysis in determining which independent variables explain a significant variation in the dependent variable of interest, the structure and form of the relationship, the strength of the relationship, and predicted values of the dependent variable. Fundamental to regression analysis is an understanding of the product moment correlation.

Product moment correlation

In marketing research, we are often interested in summarising the strength of association between two metric variables, as in the following situations:

- How strongly are sales related to advertising expenditures?
- Is there an association between market share and size of the sales force?
- Are consumers' perceptions of quality related to their perceptions of prices?

Product moment correlation (r)
A statistic summarising the strength of association between two metric variables.

In situations like these, the **product moment correlation (r),** is the most widely used statistic, summarising the strength of association between two metric (interval or ratio scaled) variables, say X and Y. It is an index used to determine whether a linear or straight line relationship exists between X and Y. It indicates the degree to which the variation in one variable, X, is related to the variation in another variable, Y. Because it was originally proposed by Karl Pearson, it is also known as the *Pearson correlation coefficient* and also referred to as *simple correlation, bivariate correlation* or merely the *correlation coefficient.* From a sample of n observations, X and Y, the product moment correlation, r, can be calculated as

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})^2}}$$

Division of the numerator and denominator by $n - 1$ gives

$$r = \frac{\dfrac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{n-1}}{\sqrt{\sum_{i=1}^{n}\dfrac{(X_i - \overline{X})^2}{n-1} \sum_{i=1}^{n}\dfrac{(Y_i - \overline{Y})^2}{n-1}}}$$

$$= \frac{COV_{xy}}{S_x S_y}$$

Covariance
A systematic relationship between two variables in which a change in one implies a corresponding change in the other (COV_{xy})

In these equations, X and Y denote the sample means, and S_x and S_y the standard deviations. COV_{xy}, the **covariance** between X and Y, measures the extent to which X and Y are related. The covariance may be either positive or negative. Division by $S_x S_y$ achieves standardisation so that r varies between -1.0 and $+1.0$. Note that the correlation coefficient is an absolute number and is not expressed in any unit of measurement. The correlation coefficient between two variables will be the same regardless of their underlying units of measurement.

As an example, suppose that a researcher wants to explain attitudes towards a respondent's city of residence in terms of duration of residence in the city. The attitude is measured on an 11-point scale (1 = do not like the city, 11 = very much like the city), and the duration of residence is measured in terms of the number of years the respondent has lived in the city. In a pre-test of 12 respondents, the data shown in Table 20.1 are obtained.

Table 20.1 Explaining attitude towards the city of residence

Respondent number	Attitude toward the city	Duration of residence
1	6	10
2	9	12
3	8	12
4	3	4
5	10	12
6	4	6
7	5	8
8	2	2
9	11	18
10	9	9
11	10	17
12	2	2

The correlation coefficient may be calculated as follows:

$$\bar{X} = (10 + 12 + 12 + 4 + 12 + 6 + 8 + 2 + 18 + 9 + 17 + 2)/12$$
$$= 9.333$$

$$\bar{Y} = (6 + 9 + 8 + 3 + 10 + 4 + 5 + 2 + 11 + 9 + 10 + 2)/12$$
$$= 6.583$$

$$
\begin{aligned}
\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) &= (10 - 9.33)(6 - 6.58) + (12 - 9.33)(9 - 6.58) \\
&\quad + (12 - 9.33)(8 - 6.58) + (4 - 9.33)(3 - 6.58) \\
&\quad + (12 - 9.33)(10 - 6.58) + (6 - 9.33)(4 - 6.58) \\
&\quad + (8 - 9.33)(5 - 6.58) + (2 - 9.33)(2 - 6.58) \\
&\quad + (18 - 9.33)(11 - 6.58) + (9 - 9.33)(9 - 6.58) \\
&\quad + (17 - 9.33)(10 - 6.58) + (2 - 9.33)(2 - 6.58) \\
&= -0.3886 + 6.4614 + 3.7914 + 19.0814 + 9.1314 + 8.5914 \\
&\quad + 2.1014 + 33.5714 + 38.3214 - 0.7986 + 26.2314 + 33.5714 \\
&= 179.6668
\end{aligned}
$$

$$
\begin{aligned}
\sum_{i=1}^{n} (X_i - \bar{X})^2 &= (10 - 9.33)^2 + (12 - 9.33)^2 + (12 - 9.33)^2 + (4 - 9.33)^2 \\
&\quad + (12 - 9.33)^2 + (6 - 9.33)^2 + (8 - 9.33)^2 + (2 - 9.33)^2 + (18 - 9.33)^2 \\
&\quad + (9 - 9.33)^2 + (17 - 9.33)^2 + (2 - 9.33)^2 \\
&= 0.4489 + 7.1289 + 7.1289 + 28.4089 + 7.1289 + 11.0889 + 1.7689 \\
&\quad + 53.7289 + 75.1689 + 0.1089 + 58.8289 + 53.7289 \\
&= 304.6668
\end{aligned}
$$

$$
\begin{aligned}
\sum_{i=1}^{n} (Y_i - \bar{Y})^2 &= (6 - 6.58)^2 + (9 - 6.58)^2 + (8 - 6.58)^2 + (3 - 6.58)^2 + (3 - 6.58)^2 \\
&\quad + (10 - 6.58)^2 + (4 - 6.58)^2 + (5 - 6.58)^2 + (2 - 6.58)^2 + (11 - 6.58)^2 \\
&\quad + (9 - 6.58)^2 + (10 - 6.58)^2 + (2 - 6.58)^2 \\
&= 0.3364 + 5.8564 + 2.0164 + 12.8164 + 11.6964 + 6.6564 + 2.4964 \\
&\quad + 20.9764 + 19.5364 + 5.8564 + 11.6964 + 20.9764 \\
&= 120.9168
\end{aligned}
$$

513

Thus,

$$r = \frac{179.6668}{\sqrt{(304.6668)(120.9168)}}$$

$$= 0.9361$$

In this example, $r = 0.9361$, a value close to 1.0. This means that respondents' duration of residence in the city is strongly associated with their attitude towards the city. Furthermore, the positive sign of r implies a positive relationship; the longer the duration of residence, the more favourable the attitude and vice versa.

Since r indicates the degree to which variation in one variable is related to variation in another, it can also be expressed in terms of the decomposition of the total variation (see Chapter 19). In other words,

$$r^2 = \frac{\text{explained variation}}{\text{total variation}}$$

$$= \frac{SS_x}{SS_y}$$

$$= \frac{\text{total variation} - \text{error variation}}{\text{total variation}}$$

$$= \frac{SS_y - SS_{error}}{SS_y}$$

Hence, r^2 measures the proportion of variation in one variable that is explained by the other. Both r and r^2 are symmetric measures of association. In other words, the correlation of X with Y is the same as the correlation of Y with X. It does not matter which variable is considered to be the dependent variable and which the independent. The product moment coefficient measures the strength of the linear relationship and is not designed to measure non-linear relationships. Thus $r = 0$ merely indicates that there is no linear relationship between X and Y. It does not mean that X and Y are unrelated. There could well be a non-linear relationship between them, which would not be captured by r (see Figure 20.1).

When computed for a population rather than a sample, the product moment correlation is denoted by the Greek letter rho, ρ. The coefficient r is an estimator of ρ. Note that the calculation of r assumes that X and Y are metric variables whose distributions have the same shape. If these assumptions are not met, r is deflated and
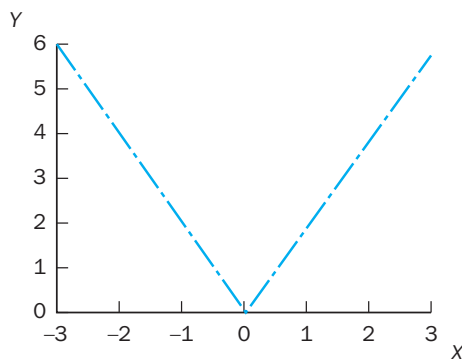
Figure 20.1
A non-linear relationship for which $r = 0$

underestimates ρ. In marketing research, data obtained by using rating scales with a small number of categories may not be strictly interval. This tends to deflate r, resulting in an underestimation of ρ.[2]

The statistical significance of the relationship between two variables measured by using r can be conveniently tested. The hypotheses are

$$H_0: \rho = 0$$
$$H_1: \rho \neq 0$$

The test statistic is

$$t = r \left[\frac{n-1}{1-r^2} \right]^{\frac{1}{2}}$$

which has a t distribution with $n - 2$ degrees of freedom.[3] For the correlation coefficient calculated based on the data given in Table 20.1,

$$t = 0.9361 \left[\frac{12-2}{1-(0.9361^2)} \right]^{\frac{1}{2}}$$

$$= 8.414$$

and the degrees of freedom $= 12 - 2 = 10$. From the t distribution table (Table 4 in the Statistical Appendix), the critical value of t for a two-tailed test and $\alpha = 0.05$ is 2.228. Hence, the null hypothesis of no relationship between X and Y is rejected. This, along with the positive sign of r, indicates that attitude towards the city is positively related to the duration of residence in the city. Moreover, the high value of r indicates that this relationship is strong.

In conducting multivariate data analysis, it is often useful to examine the simple correlation between each pair of variables. These results are presented in the form of a correlation matrix, which indicates the coefficient of correlation between each pair of variables. Usually, only the lower triangular portion of the matrix is considered. The diagonal elements all equal 1.00, since a variable correlates perfectly with itself. The upper triangular portion of the matrix is a mirror image of the lower triangular portion, since r is a symmetric measure of association. The form of a correlation matrix for five variables, V_1 to V_5 is as follows:

	V_1	V_2	V_3	V_4	V_5
V_1					
V_2	0.5				
V_3	0.3	0.4			
V_4	0.1	0.3	0.6		
V_5	0.2	0.5	0.3	0.7	

Although a matrix of simple correlations provides insights into pairwise associations, sometimes researchers want to examine the association between two variables after controlling for one or more other variables. In the latter case, partial correlation should be estimated.

Partial correlation

Whereas the product moment or simple correlation is a measure of association describing the linear association between two variables, a **partial correlation coefficient** measures the association between two variables after controlling for or adjusting for the effects of one or more additional variables. This statistic is used to answer the following questions:

■ How strongly are sales related to advertising expenditures when the effect of price is controlled?
■ Is there an association between market share and size of the sales force after adjusting for the effect of sales promotion?
■ Are consumers' perceptions of quality related to their perceptions of prices when the effect of brand image is controlled?

As in these situations, suppose that a researcher wanted to calculate the association between X and Y after controlling for a third variable, Z. Conceptually, one would first remove the effect of Z from X. To do this, one would predict the values of X based on a knowledge of Z by using the product moment correlation between X and Z, r_{xz}. The predicted value of X is then subtracted from the actual value of X to construct an adjusted value of X. In a similar manner, the values of Y are adjusted to remove the effects of Z. The product moment correlation between the adjusted values of X and the adjusted values of Y is the partial correlation coefficient between X and Y, after controlling for the effect of Z, and is denoted by $r_{xy \cdot z}$. Statistically, since the simple correlation between two variables completely describes the linear relationship between them, the partial correlation coefficient can be calculated by a knowledge of the simple correlations alone, without using individual observations.

$$r_{xy \cdot z} = \frac{r_{xy} - (r_{xz})(r_{yz})}{\sqrt{1 - r_{xz}^2} \ \sqrt{1 - r_{yz}^2}}$$

To continue our example, suppose that the researcher wanted to calculate the association between attitude towards the city, Y, and duration of residence, X_1, after controlling for a third variable, importance attached to weather, X_2. These data are presented in Table 20.2.

Table 20.2 Explaining attitude towards the city of residence, including 'importance attached to weather'

Respondent number	Attitude towards the city	Duration of residence	Importance attached to weather
1	6	10	3
2	9	12	11
3	8	12	4
4	3	4	1
5	10	12	11
6	4	6	1
7	5	8	7
8	2	2	4
9	11	18	8
10	9	9	10
11	10	17	8
12	2	2	5

The simple correlations between the variables are

$$r_{yx_1} = 0.9361 \qquad r_{yx_2} = 0.7334 \qquad r_{x_1 x_2} = 0.5495$$

The required partial correlation may be calculated as follows:

$$r_{xy_1 \cdot x_2} = \frac{0.9361 - (0.5495)(0.7334)}{\sqrt{1 - (0.5495)^2} \ \sqrt{1 - (0.7334)^2}}$$

$$= 0.9386$$

As can be seen, controlling for the effect of importance attached to weather has little effect on the association between attitude towards the city and duration of residence.

Partial correlations have an *order* associated with them that indicates how many variables are being adjusted or controlled for. The simple correlation coefficient, *r*, has a zero order, because it does not control for any additional variables while measuring the association between two variables. The coefficient $r_{xy \cdot z}$ is a first-order partial correlation coefficient, because it controls for the effect of one additional variable, *Z*. A second-order partial correlation coefficient controls for the effects of two variables, a third-order for the effects of three variables, and so on. The higher-order partial correlations are calculated similarly. The $(n + 1)$th order partial coefficient may be calculated by replacing the simple correlation coefficients on the right side of the preceding equation with the *n*th order partial coefficients.

Partial correlations can be helpful for detecting spurious relationships (see Chapter 18). The relationship between *X* and *Y* is spurious if it is solely because *X* is associated with *Z*, which is indeed the true predictor of *Y*. In this case, the correlation between *X* and *Y* disappears when the effect of *Z* is controlled. Consider a case in which consumption of a breakfast cereal brand (*C*) is positively associated with income (*I*), with $r_{ci} = 0.28$. Because this brand was popularly priced, income was not expected to be a significant factor. Therefore, the researcher suspected that this relationship was spurious. The sample results also indicated that income is positively associated with household size (*H*), $r_{hi} = 0.48$, and that household size is associated with breakfast cereal consumption, $r_{ch} = 0.56$. These figures seem to indicate that the real predictor of breakfast cereal consumption is not income but household size. To test this assertion, the first-order partial correlation between cereal consumption and income is calculated, controlling for the effect of household size. The reader can verify that this partial correlation, $r_{ci \cdot h}$, is 0.02, and the initial correlation between cereal consumption and income vanishes when the household size is controlled. Therefore, the correlation between income and cereal consumption is spurious. The special case when a partial correlation is larger than its respective zero-order correlation involves a suppressor effect (see Chapter 18).[4]

Part correlation coefficient
A measure of the correlation between *Y* and *X* when the linear effects of the other independent variables have been removed from *X* (but not from *Y*).

Another correlation coefficient of interest is the **part correlation coefficient**. This coefficient represents the correlation between *Y* and *X* when the linear effects of the other independent variables have been removed from *X* but not from *Y*. The part correlation coefficient, $r_{y(x \cdot z)}$, is calculated as follows:

$$r_{y(x \cdot z)} = \frac{r_{xy} - r_{yz} r_{xz}}{\sqrt{} 1}$$

The part correlation between attitude towards the city and the duration of residence, when the linear effects of the importance attached to weather have been removed from the duration of residence, can be calculated as

$$r_{y(x_1 \cdot x_2)} = \frac{0.9361 - (0.5495)(0.7334)}{\sqrt{1 - (0.5495)^2}}$$

$$= 0.63806$$

The partial correlation coefficient is generally viewed as more important than the part correlation coefficient. The product moment correlation, partial correlation and part correlation coefficient all assume that the data are interval or ratio scaled. If the data do not meet these requirements, the researcher should consider the use of non-metric correlation.

---

e x a m p l e

Selling ads to home shoppers[5]

Advertisements play a very important role in forming attitudes and preferences for brands. In general, it has been found that for low-involvement products, attitude towards the advertisement mediates brand cognition (beliefs about the brand) and attitude towards the brand. What would happen to the effect of this mediating variable when products are purchased through a home shopping network? Home Shopping Budapest in Hungary conducted research to assess the impact of advertisements towards purchase. A survey was conducted in which several measures were taken, such as attitude towards the product, attitude towards the brand, attitude towards the ad characteristics and brand cognitions. It was hypothesised that in a home shopping network, advertisements largely determined attitude towards the brand. To find the degree of association of attitude towards the ad with both attitude towards the brand and brand cognition, a partial correlation coefficient could be computed. The partial correlation would be calculated between attitude towards the brand and brand cognitions after controlling for the effects of attitude towards the ad on the two variables. If attitude towards the ad is significantly high, then the partial correlation coefficient should be significantly less than the product moment correlation between brand cognition and attitude towards the brand. Research was conducted which supported this hypothesis. Then Saatchi & Saatchi designed the ads aired on Home Shopping Budapest to generate positive attitude towards the advertising. This turned out to be a major competitive weapon for the network. ■

Non-metric correlation

At times the researcher may have to compute the correlation coefficient between two variables that are non-metric. It may be recalled that non-metric variables do not have interval or ratio scale properties and do not assume a normal distribution. If the non-metric variables are ordinal and numeric, Spearman's rho, ρ_s, and Kendall's tau, τ, are two measures of **non-metric correlation** which can be used to examine the correlation between them. Both these measures use rankings rather than the absolute values of the variables, and the basic concepts underlying them are quite similar. Both vary from -1.0 to $+1.0$.

Non-metric correlation
A correlation measure for two non-metric variables that relies on rankings to compute the correlation.

In the absence of ties, Spearman's ρ_s yields a closer approximation to the Pearson product moment correlation coefficient, r, than does Kendall's τ. In these cases, the absolute magnitude of τ tends to be smaller than Pearson's r. On the other hand, when the data contain a large number of tied ranks, Kendall's τ seems more appropriate. As a rule of thumb, Kendall's τ is to be preferred when a large number of cases fall into a relatively small number of categories (thereby leading to a large number of ties). Conversely, the use of Spearman's ρ_s is preferable when we have a relatively larger number of categories (thereby having fewer ties).[6]

The product moment as well as the partial and part correlation coefficients provide a conceptual foundation for bivariate as well as multiple regression analysis.

Regression analysis

Regression analysis
A statistical procedure for analysing associative relationships between a metric-dependent variable and one or more independent variables.

Regression analysis is a powerful and flexible procedure for analysing associative relationships between a metric-dependent variable and one or more independent variables. It can be used in the following ways:

1 To determine whether the independent variables explain a significant variation in the dependent variable: whether a relationship exists.
2 To determine how much of the variation in the dependent variable can be explained by the independent variables: strength of the relationship.
3 To determine the structure or form of the relationship: the mathematical equation relating the independent and dependent variables.
4 To predict the values of the dependent variable.
5 To control for other independent variables when evaluating the contributions of a specific variable or set of variables.

Although the independent variables may explain the variation in the dependent variable, this does not necessarily imply causation. The use of the terms dependent or criterion variables and independent or predictor variables in regression analysis arises from the mathematical relationship between the variables. These terms do not imply that the criterion variable is dependent on the independent variables in a causal sense. Regression analysis is concerned with the nature and degree of association between variables and does not imply or assume any causality. Bivariate regression is discussed first, followed by multiple regression.

Bivariate regression

Bivariate regression
A procedure for deriving a mathematical relationship, in the form of an equation, between a single metric-dependent variable and a single metric-independent variable.

Bivariate regression is a procedure for deriving a mathematical relationship, in the form of an equation, between a single metric-dependent or criterion variable and a single metric-independent or predictor variable. The analysis is similar in many ways to determining the simple correlation between two variables. Since an equation has to be derived, however, one variable must be identified as the dependent variable and the other as the independent variable. The examples given earlier in the context of simple correlation can be translated into the regression context.

■ Can variation in sales be explained in terms of variation in advertising expenditures? What is the structure and form of this relationship, and can it be modelled mathematically by an equation describing a straight line?
■ Can the variation in market share be accounted for by the size of the sales force?
■ Are consumers' perceptions of quality determined by their perceptions of price?

Before discussing the procedure for conducting bivariate regression, we define some important statistics associated with bivariate regression analysis.

Bivariate regression model. The basic regression equation is $Y_i = \beta_0 + \beta_1 X_i + e_i$, where Y = dependent or criterion variable, X = independent or predictor variable, β_0 = intercept of the line, β_1 = slope of the line, and e_i is the error term associated with the ith observation.

Coefficient of determination. The strength of association is measured by the coefficient of determination, r^2. It varies between 0 and 1 and signifies the proportion of the total variation in Y that is accounted for by the variation in X.

Estimated or predicted value. The estimated or predicted value of Y_i is $\hat{Y}_i = a + bx$, where $\hat{Y}_i$ is the predicted value of Y_i, and a and b are estimators of β_0 and β_1, respectively.

519

Regression coefficient. The estimated parameter b is usually referred to as the non-standardised regression coefficient.

Scattergram. A scatter diagram, or scattergram, is a plot of the values of two variables for all the cases or observations.

Standard error of estimate. This statistic, the SEE, is the standard deviation of the actual Y values from the predicted $\hat{Y}$ values.

Standard error. The standard deviation of b, SE_b, is called the standard error.

Standardised regression coefficient. Also termed the beta coefficient or beta weight, this is the slope obtained by the regression of Y on X when the data are standardised.

Sum of squared errors. The distances of all the points from the regression line are squared and added together to arrive at the sum of squared errors, which is a measure of total error, Σe_j^2.

t statistic. A t statistic with $n - 2$ degrees of freedom can be used to test the null hypothesis that no linear relationship exists between X and Y, or $H_0 : \beta_1 = 0$, where

$$t = \frac{b}{SE_b}$$

Conducting bivariate regression analysis

The steps involved in conducting bivariate regression analysis are described in Figure 20.2.

Plot the scatter diagram

Suppose that the researcher wants to explain attitudes towards the city of residence in terms of the duration of residence (see Table 20.2). In deriving such relationships, it is often useful to first examine a scatter diagram. A scatter diagram, or scattergram, is a

Plot the scatter diagram

Formulate the general model

Estimate the parameters

Estimate standardised regression coefficient

Test for significance

Determine strength & significance of association

Check prediction accuracy

Examine the residuals

Cross-validate the model
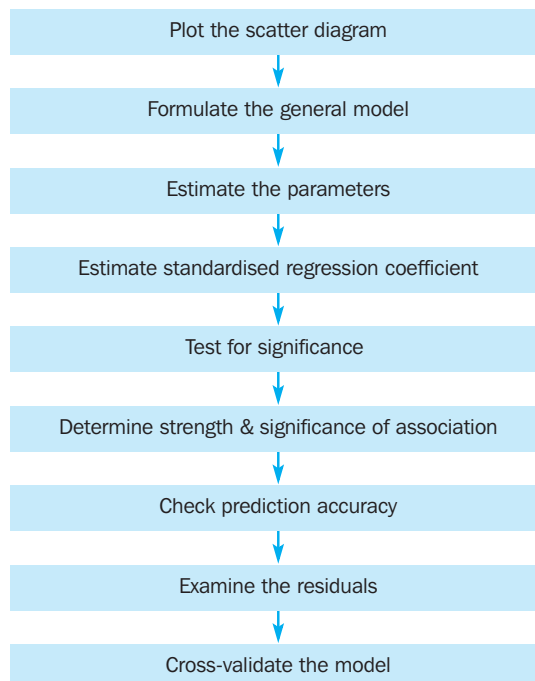
Figure 20.2
Conducting bivariate regression analysis
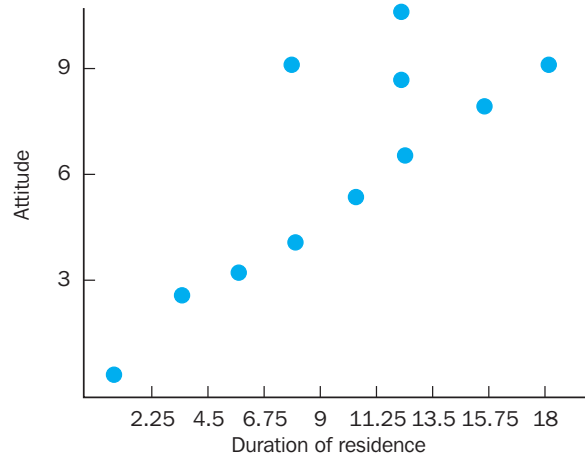
Figure 20.3
Plot of attitude with duration

plot of the values of two variables for all the cases or observations. It is customary to plot the dependent variable on the vertical axis and the independent variable on the horizontal axis. A scatter diagram is useful for determining the form of the relationship between the variables. A plot can alert the researcher to patterns in the data or to possible problems. Any unusual combinations of the two variables can be easily identified. A plot of Y (attitude towards the city) against X (duration of residence) is given in Figure 20.3. The points seem to be arranged in a band running from the bottom left to the top right. One can see the pattern: as one variable increases, so does the other. It appears from this scattergram that the relationship between X and Y is linear and could be well described by a straight line. How should the straight line be fitted to best describe the data?

The most commonly used technique for fitting a straight line to a scattergram is the **least squares procedure**. This technique determines the best-fitting line by minimising the vertical distances of all the points from the line. The best-fitting line is called the regression line. Any point that does not fall on the regression line is not fully accounted for. The vertical distance from the point to the line is the error, e_j (see Figure 20.4). The distances of all the points from the line are squared and added together to arrive at the sum of squared errors, which is a measure of total error, Σe_j^2.

Least squares procedure
A technique for fitting a straight line to a scattergram by minimising the vertical distances of all the points from the line.
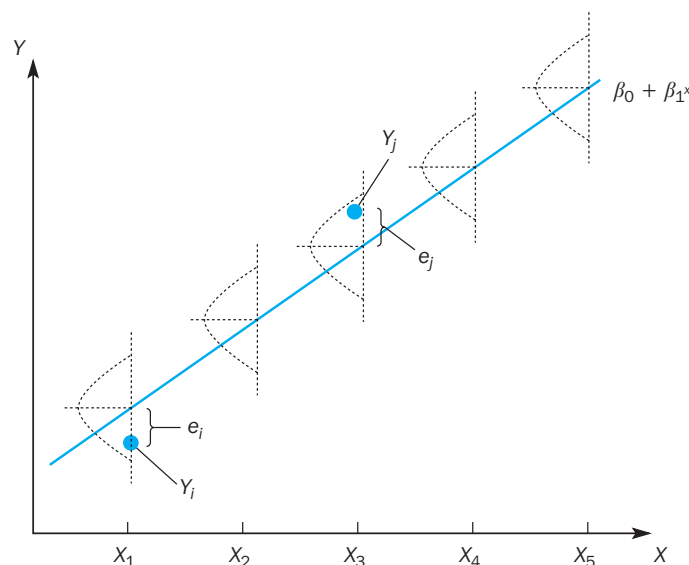
Figure 20.4
Bivariate regression

In fitting the line, the least squares procedure minimises the sum of squared errors. If Y is plotted on the vertical axis and X on the horizontal axis, as in Figure 20.4, the best fitting line is called the regression of Y on X, since the vertical distances are minimised. The scatter diagram indicates whether the relationship between Y and X can be modelled as a straight line and, consequently, whether the bivariate regression model is appropriate.

Formulate the general model

In the bivariate regression model, the general form of a straight line is

$$Y = \beta_0 + \beta_1 X$$

where Y = dependent or criterion variable
$\quad\quad\;\; X$ = independent or predictor variable
$\quad\quad\;\; \beta_0$ = intercept of the line
$\quad\quad\;\; \beta_1$ = slope of the line.

This model implies a deterministic relationship in that Y is completely determined by X. The value of Y can be perfectly predicted if β_0 and β_1 are known. In marketing research, however, very few relationships are deterministic. Thus, the regression procedure adds an error term to account for the probabilistic or stochastic nature of the relationship. The basic regression equation becomes

$$Y_i = \beta_0 + \beta_1 X_i + e_i$$

where e_i is the error term associated with the ith observation.[7] Estimation of the regression parameters, β_0 and β_1, is relatively simple.

Estimate the parameters

In most cases, β_0 and β_1 are unknown and are estimated from the sample observations using the equation

$$\hat{Y}_i = a + bx_i$$

Estimated or predicted value
The value $Y_i = a + b_x$, where a and b are, respectively, estimators of β_0 and β_1, the corresponding population parameters.

where $\hat{Y}_i$ is the **estimated or predicted value** of Y_i, and a and b are estimators of β_0 and β_1 respectively. The constant b is usually referred to as the non-standardised regression coefficient. It is the slope of the regression line, and it indicates the expected change in Y when X is changed by one unit. The formulae for calculating a and b are simple.[8] The slope, b, may be computed in terms of the covariance between X and Y (COV_{xy}) and the variance of X as

$$b = \frac{COV_{xy}}{S_x^2}$$

$$= \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$= \frac{\sum_{i=1}^{n}X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n}X_i^2 - n\bar{X}^2}$$

The intercept, a, may then be calculated using

$$a = \bar{Y} - b\bar{X}$$

For the data in Table 20.2, the estimation of parameters may be illustrated as follows:

$$\sum_{i=1}^{12} X_i Y_i = (10)(6) + (12)(9) + (12)(8) + (4)(3) + (12)(10) + (6)(4) + (8)(5) + (2)(2)$$
$$+ (18)(11) + (9)(9) + (17)(10) + (2)(2)$$
$$= 917$$

$$\sum_{i=1}^{12} X_i^2 = 10^2 + 12^2 + 12^2 + 4^2 + 12^2 + 6^2 + 8^2 + 2^2 + 18^2 + 9^2 + 17^2 + 2^2$$
$$= 1350$$

It may be recalled from earlier calculations of the simple correlation that

$$\bar{X} = 9.333$$
$$\bar{Y} = 6.583$$

Given $n = 12$, b can be calculated as

$$b = \frac{917 - (12)(9.333)(6.583)}{1350 - (12)(9.333)^2}$$
$$= 0.5897$$
$$a = \bar{Y} - b\bar{X}$$
$$= 6.583 - (0.5897)(9.333)$$
$$= 1.0793$$

Note that these coefficients have been estimated on the raw (untransformed) data. Should standardisation of the data be considered desirable, the calculation of the standardised coefficients is also straightforward.

Estimate the standardised regression coefficient

Standardisation is the process by which the raw data are transformed into new variables that have a mean of 0 and a variance of 1 (Chapter 17). When the data are standardised, the intercept assumes a value of 0. The term *beta coefficient* or *beta weight* is used to denote the standardised regression coefficient. In this case, the slope obtained by the regression of Y on X, B_{yx}, is the same as the slope obtained by the regression of X on Y, B_{xy}. Moreover, each of these regression coefficients is equal to the simple correlation between X and Y:

$$B_{yx} = B_{xy} = r_{xy}$$

There is a simple relationship between the standardised and non-standardised regression coefficients:

$$B_{yx} = b_{yx}\left(\frac{S_x}{S_y}\right)$$

For the regression results given in Table 20.3, the value of the beta coefficient is estimated as 0.9361.

Once the parameters have been estimated, they can be tested for significance.

Table 20.3 Bivariate regression

Multiple R	0.93608
R^2	0.87624
Adjusted R^2	0.86387
Standard error	1.22329

Analysis of variance			
	df	Sum of squares	Mean square
Regression	1	105.95222	105.95222
Residual	10	14.96444	1.49644
$F = 70.80266$	Significance of $F = 0.0000$		

Variables in the equation					
Variable	b	SE_b	β	T	Sig. of T
Duration	0.58972	0.07008	0.93608	8.414	0.0000
(Constant)	1.07932	0.74335		1.452	0.1772

Test for significance

The statistical significance of the linear relationship between X and Y may be tested by examining the hypotheses

$$H_0: \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

The null hypothesis implies that there is no linear relationship between X and Y. The alternative hypothesis is that there is a relationship, positive or negative, between X and Y. Typically, a two-tailed test is done. A t statistic with $n - 2$ degrees of freedom can be used, where

$$t = \frac{b}{SE_b}$$

and SE_b denotes the standard deviation of b, called the *standard error*.[9] The t distribution was discussed in Chapter 18.

Using a software package, the regression of attitude on duration of residence, using the data shown in Table 20.2, yielded the results shown in Table 20.3. The intercept, a, equals 1.0793, and the slope, b, equals 0.5897. Therefore, the estimated equation is

$$\text{attitude } (\hat{Y}) = 1.0793 + 0.5897 \text{ (duration of residence)}$$

The standard error or standard deviation of b is estimated as 0.07008, and the value of the t statistic, $t = 0.5897/0.0701 = 8.414$, with $n - 2 = 10$ degrees of freedom. From Table 4 in the Statistical Appendix, we see that the critical value of t with 10 degrees of freedom and $\alpha = 0.05$ is 2.228 for a two-tailed test. Since the calculated value of t is larger than the critical value, the null hypothesis is rejected. Hence, there is a significant linear relationship between attitude towards the city and duration of residence in the city. The positive sign of the slope coefficient indicates that this relationship is positive. In other words, those who have lived in the city for a longer time have more positive attitudes towards the city.

Determine strength and significance of association

A related inference involves determining the strength and significance of the association between Y and X. The strength of association is measured by the coefficient of determination, r^2. In bivariate regression, r^2 is the square of the simple correlation coefficient obtained by correlating the two variables. The coefficient r^2 varies between 0 and 1. It signifies the proportion of the total variation in Y that is accounted for by the variation in X. The decomposition of the total variation in Y is similar to that for analysis of variance (Chapter 19). As shown in Figure 20.5, the total variation SS_y may be decomposed into the variation accounted for by the regression line, SS_{reg}, and the error or residual variation, SS_{error} or SS_{res}, as follows:

$$SS_y = SS_{reg} + SS_{res}$$

$$SS_y = \sum_{i=1}^{n} (Y_i - \overline{Y})^2$$

$$SS_{reg} = \sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2$$

$$SS_{res} = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

The strength of the association may then be calculated as follows:

$$r^2 = \frac{SS_{reg}}{SS_y}$$

$$= \frac{SS_y - SS_{res}}{SS_y}$$

To illustrate the calculations of r^2, let us consider again the effect of attitude towards the city on the duration of residence. It may be recalled from earlier calculations of the simple correlation coefficient that

$$SS_y = \sum_{i=1}^{n} (Y_i - \overline{Y})^2$$

$$= 120.9168$$

The predicted values ($\hat{Y}$) can be calculated using the regression equation

$$\text{attitude } (\hat{Y}) = 1.0793 + 0.5897 \text{ (duration of residence)}$$

For the first observation in Table 20.2, this value is

$$(\hat{Y}) = 1.0793 + (0.5897 \times 10) = 6.9763$$
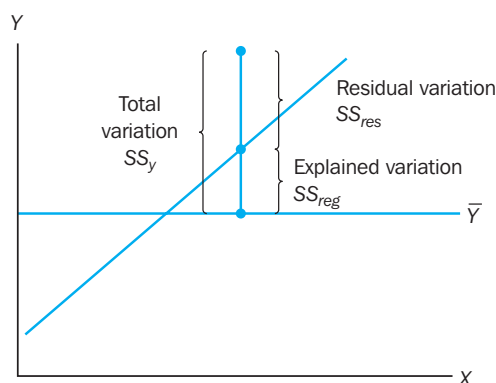
Figure 20.5
Decomposition of the total variation in bivariate regression

For each successive observation, the predicted values are, in order, 8.1557, 8.1557, 3.4381, 8.1557, 4.6175, 5.7969, 2.2587, 11.6939, 6.3866, 11.1042, 2.2587. Therefore,

$$SS_{reg} = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$$

$$
\begin{aligned}
&= (6.9763 - 6.5833)^2 + (8.1557 - 6.5833)^2 + (8.1557 - 6.5833)^2 \\
&\quad + (3.4381 - 6.5833)^2 + (8.1557 - 6.5833)^2 + (4.6175 - 6.5833)^2 \\
&\quad + (5.7969 - 6.5833)^2 + (2.2587 - 6.5833)^2 + (11.6939 - 6.5833)^2 \\
&\quad + (6.3866 - 6.5833)^2 + (11.1042 - 6.5833)^2 + (2.2587 - 6.5833)^2 \\
&= 0.1544 + 2.4724 + 2.4724 + 9.8922 + 2.4724 + 3.8643 + 0.6184 \\
&\quad + 18.7021 + 26.1182 + 0.0387 + 20.4385 + 18.7021 \\
&= 105.9466
\end{aligned}
$$

$$SS_{res} = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

$$
\begin{aligned}
&= (6 - 6.9763)^2 + (9 - 8.1557)^2 + (8 - 8.1557)^2 + (3 - 3.4381)^2 \\
&\quad + (10 - 8.1557)^2 + (4 - 4.6175)^2 + (5 - 5.7969)^2 + (2 - 2.2587)^2 \\
&\quad + (11 - 11.6939)^2 + (9 - 6.3866)^2 + (10 - 11.1042)^2 + (2 - 2.2587)^2 \\
&= 14.9644
\end{aligned}
$$

It can be seen that $SS_y = SS_{reg} + SS_{res}$. Furthermore,

$$r^2 = \frac{SS_{reg}}{SS_y}$$

$$= \frac{105.9466}{120.9168}$$

$$= 0.8762$$

Another equivalent test for examining the significance of the linear relationship between X and Y (significance of b) is the test for the significance of the coefficient of determination. The hypotheses in this case are

$$H_0: R^2_{pop} = 0$$
$$H_1: R^2_{pop} > 0$$

The appropriate test statistic is the F statistic

$$F = \frac{SS_{reg}}{SS_{res}/(n-2)}$$

which has an F distribution with 1 and $n - 2$ degrees of freedom. The F test is a generalised form of the t test (see Chapter 18). If a random variable is t distributed with n degrees of freedom, then t^2 is F distributed with 1 and n degrees of freedom. Hence, the F test for testing the significance of the coefficient of determination is equivalent to testing the following hypotheses:

$$H_0: \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

or

$$H_0: \rho = 0$$
$$H_1: \rho \neq 0$$

From Table 20.3, it can be seen that

$$r^2 = \frac{105.9522}{105.9522 + 14.9644}$$

$$= 0.8762$$

which is the same as the value calculated earlier. The value of the F statistic is

$$F = \frac{105.9522}{14.9644/10}$$

$$= 70.8027$$

with 1 and 10 degrees of freedom. The calculated F statistic exceeds the critical value of 4.96 determined from Table 5 in the Statistical Appendix. Therefore, the relationship is significant at $\alpha = 0.05$, corroborating the results of the t test. If the relationship between X and Y is significant, it is meaningful to predict the values of Y based on the values of X and to estimate prediction accuracy.

Check prediction accuracy

To estimate the accuracy of predicted values, $\hat{Y}$, it is useful to calculate the standard error of estimate, SEE. This statistic is the standard deviation of the actual Y values from the predicted $\hat{Y}$ values.

$$SEE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y})^2}{n-2}}$$

$$SEE = \sqrt{\frac{SS_{res}}{n-2}}$$

or, more generally, if there are k independent variables

$$SEE = \sqrt{\frac{SS_{res}}{n-k-1}}$$

SEE may be interpreted as a kind of average residual or average error in predicting Y from the regression equation.[10]

Two cases of prediction may arise. The researcher may want to predict the mean value of Y for all the cases with a given value of X, say X_0, or predict the value of Y for a single case. In both situations, the predicted value is the same and is given by $\hat{Y}$, where

$$\hat{Y} = a + bX_0$$

But the standard error is different in the two situations, although in both situations it is a function of SEE. For large samples, the standard error for predicting the mean value of Y is $SEE/\sqrt{n}$ and for predicting individual Y values it is SEE. Hence, the construction of confidence intervals (see Chapter 15) for the predicted value varies, depending upon whether the mean value or the value for a single observation is being predicted. For the data given in Table 20.3, SEE is estimated as follows:

$$SEE = \sqrt{\frac{14.9644}{12-2}}$$

$$= 1.22329$$

The final two steps in conducting bivariate regression, namely examination of residuals and model cross-validation, are considered later, and we now turn to the assumptions underlying the regression model.

Assumptions

The regression model makes a number of assumptions in estimating the parameters and in significance testing, as shown in Figure 20.4:

1 The error term is normally distributed. For each fixed value of X, the distribution of Y is normal.[11]
2 The means of all these normal distributions of Y, given X, lie on a straight line with slope b.
3 The mean of the error term is 0.
4 The variance of the error term is constant. This variance does not depend on the values assumed by X.
5 The error terms are uncorrelated. In other words, the observations have been drawn independently.

Insights into the extent to which these assumptions have been met can be gained by an examination of residuals, which is covered in the next section on multiple regression.[12]

Multiple regression

Multiple regression
A statistical technique that simultaneously develops a mathematical relationship between two or more independent variables and an interval-scaled dependent variable.

Multiple regression involves a single dependent variable and two or more independent variables. The questions raised in the context of bivariate regression can also be answered via multiple regression by considering additional independent variables:

- Can variation in sales be explained in terms of variation in advertising expenditures, prices and level of distribution?
- Can variation in market shares be accounted for by the size of the sales force, advertising expenditures and sales promotion budgets?
- Are consumers' perceptions of quality determined by their perceptions of prices, brand image and brand attributes?

Additional questions can also be answered by multiple regression:

- How much of the variation in sales can be explained by advertising expenditures, prices and level of distribution?
- What is the contribution of advertising expenditures in explaining the variation in sales when the levels of prices and distribution are controlled?
- What levels of sales may be expected given the levels of advertising expenditures, prices and level of distribution?

e x a m p l e

Global brands, local ads[13]

Europeans welcome brands from other countries, but when it comes to advertising, they seem to prefer brands from their own country. A survey conducted by Yankelovich and Partners and its affiliates found that most European consumers' favourite commercials were for local brands even though they were more than likely to buy foreign brands. Respondents in Britain, France and Germany named Coca-Cola as the most often purchased soft drink. The French, however, selected the famous award-winning spot for France's Perrier bottled water as their favourite commercial. Similarly, in Germany, the favourite advertising was for a German brand of non-alcoholic beer, Clausthaler. In Britain, though, Coca-Cola was the favourite soft drink and also the favourite advertising. In the light of such findings, the impor-

tant question was: does advertising help? Does it help increase the purchase probability of the brand or does it merely maintain the brand recognition rate high? One way of finding out was by running a regression where the dependent variable was the likelihood of brand purchase and the independent variables were brand attribute evaluations and advertising evaluations. Separate models with and without advertising could be run to assess any significant difference in the contribution. Individual t tests could also be examined to find out the significant contribution of both the brand attributes and advertising. The results could indicate the degree to which advertising plays an important part in brand purchase decisions. ■

Multiple regression model
An equation used to explain the results of multiple regression analysis.

The general form of the **multiple regression model** is as follows:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_k X_k + e$$

which is estimated by the following equation:

$$\hat{Y} = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots + b_k X_k$$

As before, the coefficient a represents the intercept, but the bs are now the partial regression coefficients. The least squares criterion estimates the parameters in such a way as to minimise the total error, SS_{res}. This process also maximises the correlation between the actual values of Y and the predicted values of $\hat{Y}$. All the assumptions made in bivariate regression also apply in multiple regression. We define some associated statistics and then describe the procedure for multiple regression analysis.[14]

Most of the statistics and statistical terms described under bivariate regression also apply to multiple regression. In addition, the following statistics are used:

Adjusted R^2. R^2, the coefficient of multiple determination, is adjusted for the number of independent variables and the sample size to account for the diminishing returns. After the first few variables, the additional independent variables do not make much contribution.

Coefficient of multiple determination. The strength of association in multiple regression is measured by the square of the multiple correlation coefficient, R^2, which is also called the *coefficient of multiple determination.*

F test. The F test is used to test the null hypothesis that the coefficient of multiple determination in the population, R^2_{pop}, is zero. This is equivalent to testing the null hypothesis $H_0: \beta_1 = \beta_2 = \beta_3 = \ldots = \beta_k = 0$. The test statistic has an F distribution with k and $(n-k-1)$ degrees of freedom.

Partial F test. The significance of a partial regression coefficient, β_i, of X_i may be tested using an incremental F statistic. The incremental F statistic is based on the increment in the explained sum of squares resulting from the addition of the independent variable X_i to the regression equation after all the other independent variables have been included.

Partial regression coefficient. The partial regression coefficient, b_1, denotes the change in the predicted value, $\hat{Y}$, per unit change in X_1 when the other independent variables, X_2 to X_k, are held constant.

Conducting multiple regression analysis

The steps involved in conducting multiple regression analysis are similar to those for bivariate regression analysis. The discussion focuses on partial regression coefficients, strength of association, significance testing and examination of residuals.

Estimating the partial regression coefficients

To understand the meaning of a partial regression coefficient, let us consider a case in which there are two independent variables, so that

$$\hat{Y} = a + b_1 X_1 + b_2 X_2$$

First, note that the relative magnitude of the partial regression coefficient of an independent variable is, in general, different from that of its bivariate regression coefficient. In other words, the partial regression coefficient, b_1, will be different from the regression coefficient, b, obtained by regressing Y on only X_1. This happens because X_1 and X_2 are usually correlated. In bivariate regression, X_2 was not considered, and any variation in Y that was shared by X_1 and X_2 was attributed to X_1. In the case of multiple independent variables, however, this is no longer true.

The interpretation of the partial regression coefficient, b_1, is that it represents the expected change in Y when X_1 is changed by one unit but X_2 is held constant or otherwise controlled. Likewise, b_2 represents the expected change in Y for a unit change in X_2 when X_1 is held constant. Thus, calling b_1 and b_2 partial regression coefficients is appropriate. It can also be seen that the combined effects of X_1 and X_2 on Y are additive. In other words, if X_1 and X_2 are each changed by one unit, the expected change in Y would be $(b_1 + b_2)$.

Conceptually, the relationship between the bivariate regression coefficient and the partial regression coefficient can be illustrated as follows. Suppose that one were to remove the effect of X_2 from X_1. This could be done by running a regression of X_1 on X_2. In other words, one would estimate the equation $\hat{Y}_1 = a + bX_2$ and calculate the residual $X_r = (X_1 - \hat{Y}_1)$. The partial regression coefficient, b_1, is then equal to the bivariate regression coefficient, b, obtained from the equation $\hat{Y} = a + bX_r$. In other words, the partial regression coefficient, b_1, is equal to the regression coefficient, b, between Y and the residuals of X_1 from which the effect of X_2 has been removed. The partial coefficient, b, can also be interpreted along similar lines.

Extension to the case of k variables is straightforward. The partial regression coefficient, b_1, represents the expected change in Y when X_1 is changed by one unit and X_2 to X_k are held constant. It can also be interpreted as the bivariate regression coefficient, b, for the regression of Y on the residuals of X_1, when the effect of X_2 to X_k has been removed from X_1.

The beta coefficients are the partial regression coefficients obtained when all the variables $(Y, X_1, X_2, \ldots X_k)$ have been standardised to a mean of 0 and a variance of 1 before estimating the regression equation. The relationship of the standardised to the non-standardised coefficients remains the same as before:

$$B_1 = b_1 \left(\frac{S_{x1}}{S_y} \right)$$

$$\vdots$$

$$B_k = b_k \left(\frac{S_{xk}}{S_y} \right)$$

The intercept and the partial regression coefficients are estimated by solving a system of simultaneous equations derived by differentiating and equating the partial derivatives to 0. Since these coefficients are automatically estimated by the various computer programs, we will not present the details. Yet it is worth noting that the equations cannot be solved if (1) the sample size, n, is smaller than or equal to the number of independent variables, k, or (2) one independent variable is perfectly correlated with another.

Suppose that in explaining the attitude towards the city we now introduce a second variable, importance attached to the weather. The data for the 12 pre-test respondents on attitude towards the city, duration of residence and importance attached to the weather are given in Table 20.2. The results of multiple regression analysis are depicted in Table 20.4. The partial regression coefficient for duration (X_1) is now 0.4811, different from what it was in the bivariate case. The corresponding beta coefficient is 0.7636.

Table 20.4 Multiple regression

Multiple R	0.97210
R^2	0.94498
Adjusted R^2	0.93276
Standard error	0.85974

Analysis of variance			
	df	Sum of squares	Mean square
Regression	2	114.26425	57.13213
Residual	9	6.65241	0.73916
$F = 77.29364$	Significance of $F = 0.0000$		

Variables in the equation					
Variable	b	SE_b	β	T	Sig. of T
Importance	0.28865	0.08608	0.31382	3.353	0.0085
Duration	0.48108	0.05895	0.76363	8.160	0.0000
(Constant)	0.33732	0.56736		0.595	0.5668

The partial regression coefficient for importance attached to weather (X_2) is 0.2887, with a beta coefficient of 0.3138. The estimated regression equation is

$$(\hat{Y}) = 0.33732 + 0.48108X_1 + 0.28865X_2$$

or

$$\text{attitude} = 0.33732 + 0.48108 \,(\text{duration}) + 0.28865 \,(\text{importance})$$

This equation can be used for a variety of purposes, including predicting attitudes towards the city, given a knowledge of the respondents' duration of residence in the city and the importance they attach to weather.

Strength of association

The strength of the relationship stipulated by the regression equation can be determined by using appropriate measures of association. The total variation is decomposed as in the bivariate case:

$$SS_y = SS_{reg} + SS_{res}$$

where $SS_y = \sum_{i=1}^{n} (Y_i - \overline{Y})^2$

$SS_{reg} = \sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2$

$SS_{res} = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$

531

The strength of association is measured by the square of the multiple correlation coefficient, R^2, which is also called the coefficient of multiple determination:

$$R^2 = \frac{SS_{reg}}{SS_y}$$

The multiple correlation coefficient, R, can also be viewed as the simple correlation coefficient, r, between Y and $\hat{Y}$. Several points about the characteristics of R^2 are worth noting. The coefficient of multiple determination, R^2, cannot be less than the highest bivariate, r^2, of any individual independent variable with the dependent variable. R^2 will be larger when the correlations between the independent variables are low. If the independent variables are statistically independent (uncorrelated), then R^2 will be the sum of bivariate r^2 of each independent variable with the dependent variable. R^2 cannot decrease as more independent variables are added to the regression equation. Yet diminishing returns set in, so that after the first few variables, the additional independent variables do not make much of a contribution.[15] For this reason, R^2 is adjusted for the number of independent variables and the sample size by using the following formula:

$$\text{adjusted } R^2 = R^2 - \frac{k(1 - R^2)}{n - k - 1}$$

For the regression results given in Table 20.4, the value of R^2 is

$$R^2 = \frac{114.2643}{114.2643 + 6.6524}$$

$$= 0.9450$$

This is higher than the r^2 value of 0.8762 obtained in the bivariate case. The r^2 in the bivariate case is the square of the simple (product moment) correlation between attitude toward the city and duration of residence. The R^2 obtained in multiple regression is also higher than the square of the simple correlation between attitude and importance attached to weather (which can be estimated as 0.5379). The adjusted R^2 is estimated as

$$\text{adjusted } R^2 = 0.9450 - \frac{2(1.0 - 0.9450)}{12 - 2 - 1}$$

$$= 0.9328$$

Note that the value of adjusted R^2 is close to R^2 and both are higher than r^2 for the bivariate case. This suggests that the addition of the second independent variable, importance attached to weather, makes a contribution in explaining the variation in attitude towards the city.

Test for significance

Significance testing involves testing the significance of the overall regression equation as well as specific partial regression coefficients. The null hypothesis for the overall test is that the coefficient of multiple determination in the population, R^2_{pop}, is zero:

$$H_0: R^2_{pop} = 0$$

This is equivalent to the following null hypothesis:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \ldots = \beta_k = 0$$

The overall test can be conducted by using an F statistic

$$F = \frac{SS_{reg}/k}{SS_{reg}/(n-k-1)}$$

$$= \frac{R^2/k}{(1-R^2)/(n-k-1)}$$

which has an F distribution with k and $n-k-1$ degrees of freedom.[16] For the multiple regression results given in Table 20.4,

$$F = \frac{114.2642 / 2}{6.6524 / 9} = 77.2938$$

which is significant at $\alpha = 0.05$.

If the overall null hypothesis is rejected, one or more population partial regression coefficients have a value different from 0. To determine which specific coefficients (β_is) are non-zero, additional tests are necessary. Testing for the significance of the β_is can be done in a manner similar to that in the bivariate case by using t tests. The significance of the partial coefficient for importance attached to weather may be tested by the following equation:

$$t = \frac{b}{SE_b}$$

$$= \frac{0.2887}{0.08608} = 3.353$$

which has a t distribution with $n-k-1$ degrees of freedom. This coefficient is significant at $\alpha = 0.05$. The significance of the coefficient for duration of residence is tested in a similar way and found to be significant. Therefore, both the duration of residence and importance attached to weather are important in explaining attitude towards the city.

Some computer programs provide an equivalent F test, often called the partial F test, which involves a decomposition of the total regression sum of squares, SS_{reg}, into components related to each independent variable. In the standard approach, this is done by assuming that each independent variable has been added to the regression equation after all the other independent variables have been included. The increment in the explained sum of squares, resulting from the addition of an independent variable, X_i, is the component of the variation attributed to that variable and is denoted SS_{xi}.[17] The significance of the partial regression coefficient for this variable, β_i, is tested using an incremental F statistic

$$F = \frac{SS_{xi} / 1}{SS_{res} / (n-k-1)}$$

which has an F distribution with 1 and $(n-k-1)$ degrees of freedom.

While high R^2 and significant partial regression coefficients are comforting, the efficacy of the regression model should be evaluated further by an examination of the residuals.

Examine the residuals

Residual
The difference between the observed value of Y_i and the value predicted by the regression equation $\hat{Y}_i$.

A **residual** is the difference between the observed value of Y_i and the value predicted by the regression equation $\hat{Y}_i$. Residuals are used in the calculation of several statistics associated with regression. In addition, scattergrams of the residuals – in which the residuals are plotted against the predicted values, $\hat{Y}_i$, time, or predictor variables – provide useful insights in examining the appropriateness of the underlying assumptions and regression model fitted.[18]

Examining scattergrams of residuals provides useful insights.
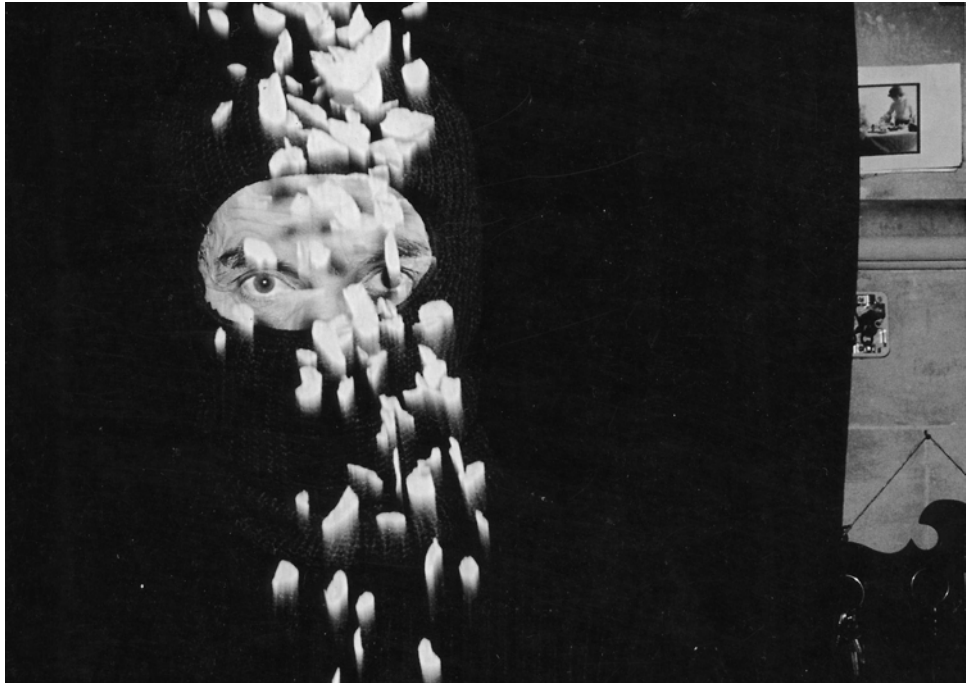
The assumption of a normally distributed error term can be examined by constructing a histogram of the residuals. A visual check reveals whether the distribution is normal. Additional evidence can be obtained by determining the percentages of residuals falling within ±1 SE or ±2 SE. These percentages can be compared with what would be expected under the normal distribution (68% and 95%, respectively). More formal assessment can be made by running the K-S one-sample test.

The assumption of constant variance of the error term can be examined by plotting the residuals against the predicted values of the dependent variable, $\hat{Y}_i$. If the pattern is not random, the variance of the error term is not constant. Figure 20.6 shows a pattern whose variance is dependent on the $\hat{Y}_i$ values.

A plot of residuals against time, or the sequence of observations, will throw some light on the assumption that the error terms are uncorrelated. A random pattern should be seen if this assumption is true. A plot like the one in Figure 20.7 indicates a linear relationship between residuals and time. A more formal procedure for examining the correlations between the error terms is the Durbin-Watson test.[19]

Plotting the residuals against the independent variables provides evidence of the appropriateness or inappropriateness of using a linear model. Again, the plot should
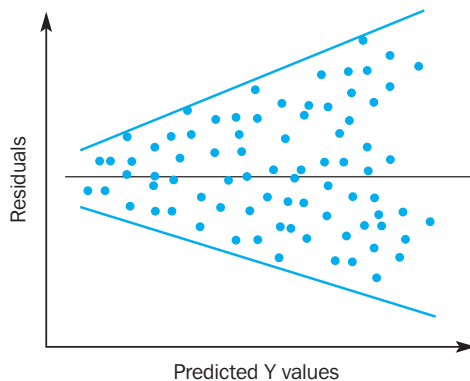
Figure 20.6
Residual plot indicating that variance is not constant
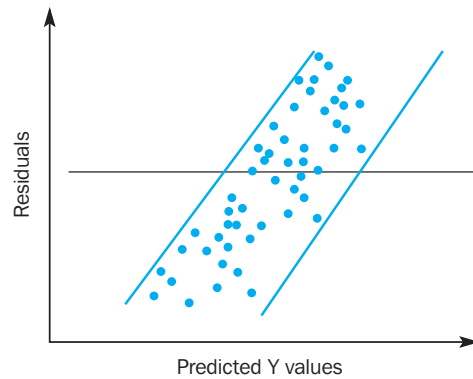
Figure 20.7
Plot indicating a linear relationship between residuals and time

result in a random pattern. The residuals should fall randomly, with relatively equal distribution dispersion about 0. They should not display any tendency to be either positive or negative.

To examine whether any additional variables should be included in the regression equation, one could run a regression of the residuals on the proposed variables. If any variable explains a significant proportion of the residual variation, it should be considered for inclusion. Inclusion of variables in the regression equation should be strongly guided by the researcher's theory. Thus, an examination of the residuals provides valuable insights into the appropriateness of the underlying assumptions and the model that is fitted. Figure 20.8 shows a plot that indicates that the underlying assumptions are met and that the linear model is appropriate.

If an examination of the residuals indicates that the assumptions underlying linear regression are not met, the researcher can transform the variables in an attempt to satisfy the assumptions. Transformations, such as taking logs, square roots or reciprocals, can stabilise the variance, make the distribution normal or make the relationship linear. We further illustrate the application of multiple regression with an example.
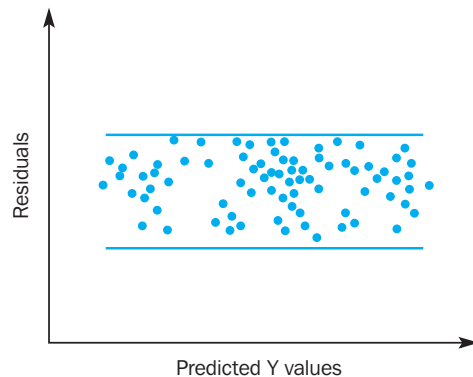
Figure 20.8
Plot of residuals indicating that a fitted model is appropriate

example

At no 'Ad' ditional cost[20]

It is widely believed that consumer magazines' prices are subsidised by the advertising carried within the magazines. A study examined the contribution of advertising to the price per copy of magazines.

Multiple regression analysis was used to examine the relationships among price per copy and editorial pages, circulation, percentage of news-stand circulation, promotional expenditures, percentage of colour pages, and per copy advertising revenues. The form of the analysis was

$$PPC = b_0 + b_1(\text{ed. pages}) + b_2(\text{circ.}) + b_3(\% \text{ news circ.}) + b_4(\text{PE}) + b_5(\% \text{ colour}) + b_6(\text{ad revs.})$$

where

$$
\begin{aligned}
\text{PPC} &= \text{price per copy (in €)} \\
\text{ed. pages} &= \text{editorial pages per average issue} \\
\text{circ.} &= \text{the log of average paid circulation (in thousands)} \\
\% \text{ news circ.} &= \text{percentage news-stand circulation} \\
\text{PE} &= \text{promotional expenditures (in €)} \\
\% \text{ colour} &= \text{percentage of pages printed in colour} \\
\text{ad revs.} &= \text{per copy advertising revenues (in €)}
\end{aligned}
$$

Table 1 shows the zero-order Pearson product moment correlations among the variables. The correlations provide directional support for the predicted relationships and show that collinearity among the independent variables is sufficiently low so as not to affect the stability of the regression analysis. The highest correlation among the independent variables was between promotional expenditures and circulation ($r = 0.42$).

Table 1 Zero-order correlation matrix of variables in analyses

	Price per copy	Price per editorial page	Circulation	Editorial pages	Promotional expenditures	% colour pages	% news-stand circ.
Price per editorial page	0.60[a]						
Circulation	−0.21[a]	−0.42[a]					
Editorial pages	0.52[a]	−0.30[a]	0.29[a]				
Promotional expenditures	−0.22[a]	−0.06	0.42[a]	−0.19			
% colour pages	0.01	−0.15	0.33[a]	0.19	−0.15		
% news-stand circ.	0.46[a]	0.17	0.09	0.31[a]	0.26[a]	0.02	
Ad. revenues per copy	0.29[a]	−0.04	−0.25[a]	0.30[a]	−0.14	0.15	0.08

[a] $p < 0.05$.

The results of the regression analysis using price per copy as the dependent variable are given in Table 2. Of the six independent variables, three were significant ($p < 0.05$): the number of editorial pages, average circulation, and percentage news-stand circulation. The three variables accounted for virtually all of the explained variance ($R^2 = 0.51$; adjusted $R^2 = 0.48$). The direction of the coefficients was consistent with prior expectations: the number of editorial pages was positive, circulation was negative, and percentage news-stand circulation was positive. This was expected, given the structure of the magazine publishing industry, and it confirmed the hypothesised relationship.

Table 2 Regression analysis using price per copy as dependent variable

	b	SE	F
Dependent variable: price per copy			
Independent variables:			
Editorial pages	0.0084	0.0017	23.04[a]
Circulation	−0.4180	0.1372	9.29[a]
Percentage news-stand circulation	0.0067	0.0016	18.46 [a]
Promotional expenditures	0.13–0.04[b]	0.0000	0.59
Percentage colour pages	0.0227	0.0092	0.01
Per copy ad. revenues	0.1070	0.0412	0.07
Overall $R^2 = 0.51$	$df = 6, 93$	Overall $F = 16.19$[a]	

[a] $p < 0.05$.
[b] Decimal moved in by four zeros.

Promotional expenditures, use of colour and per copy advertising revenues were found to have no relationship with price per copy, after the effects of circulation, percentage news-stand circulation and editorial pages were controlled in the regression analysis.

Because the effect of per copy advertising revenue was not significant, no support was found for the contention that advertising decreases the price per copy of consumer magazines. It was concluded that advertising in magazines is provided free to consumers, but does not subsidise prices. ■

In the preceding example, promotional expenditures, percentage of colour pages and per copy advertising revenues were not found to be significantly related to the price per copy of magazines. Some of the independent variables considered in a study often turn out to be non-significant. When there are a large number of independent variables and the researcher suspects that not all of them are significant, stepwise regression should be used.

Stepwise regression

Stepwise regression
A regression procedure in which the predictor variables enter or leave the regression equation one at a time.

The purpose of **stepwise regression** is to select, from a large number of predictor variables, a small subset of variables that account for most of the variation in the dependent or criterion variable. In this procedure, the predictor variables enter or are removed from the regression equation one at a time.[21] There are several approaches to stepwise regression.

1 *Forward inclusion.* Initially, there are no predictor variables in the regression equation. Predictor variables are entered one at a time, only if they meet certain criteria specified in terms of F ratio. The order in which the variables are included is based on the contribution to the explained variance.
2 *Backward elimination.* Initially, all the predictor variables are included in the regression equation. Predictors are then removed one at a time based on the F ratio for removal.
3 *Stepwise solution.* Forward inclusion is combined with the removal of predictors that no longer meet the specified criterion at each step.

Stepwise procedures do not result in regression equations which are optimal, in the sense of producing the largest R^2, for a given number of predictors.[22] Because of the correlations between predictors, an important variable may never be included or less important variables may enter the equation. To identify an optimal regression equation, one would have to compute combinatorial solutions in which all possible combinations are examined. Nevertheless, stepwise regression can be useful when the sample size is large in relation to the number of predictors, as shown in the following example.

e x a m p l e

Browsers step out[23]

A profile of browsers in regional shopping centres was constructed using three sets of independent variables: demographics, shopping behaviour, and psychological and attitudinal variables. The dependent variable consisted of a browsing index. In a stepwise regression including all three sets of variables, demographics were found to be the most powerful predictors of browsing behaviour. The final regression equation, which contained 20 of the possible 36 variables, included all the demographics. The table presents the regression coefficients, standard errors of the coefficients, and their significance levels.

In interpreting the coefficients, it should be recalled that the smaller the browsing index (the dependent variable), the greater the tendency to exhibit behaviours associated with browsing. The two predictors with the largest coefficients were gender and employment status. Browsers were more likely to be employed females. They also tend to be somewhat

'downscale', compared with other shopping centre patrons, exhibiting lower levels of education and income, after accounting for the effects of gender and employment status. Although browsers tend to be somewhat younger than non-browsers, they are not necessarily single; those who reported larger family sizes tended to be associated with smaller values of the browsing index.

The 'downscale' profile of browsers relative to other patrons indicates that speciality stores in shopping centres should emphasise moderately priced products. This may explain the historically low rate of failure in shopping centres among such stores and the tendency of high-priced speciality shops to be located in only the prestigious shopping centres or 'upscale' non-enclosed shopping centres. ■

Regression of browsing index on descriptive and attitudinal variables by order of entry into stepwise regression

Variable description	Coefficient	SE	Significance
Gender (0 = male, 1 = female)	−0.485	0.164	0.001
Employment status (0 = employed)	0.391	0.182	0.003
Self-confidence	−0.152	0.128	0.234
Education	0.079	0.072	0.271
Brand intention	−0.063	0.028	0.024
Watch daytime TV? (0 = yes)	0.232	0.144	0.107
Tension	−0.182	0.069	0.008
Income	0.089	0.061	0.144
Frequency of shopping centre visits	−0.130	0.059	0.028
Fewer friends than most	0.162	0.084	0.054
Good shopper	−0.122	0.090	0.174
Others' opinions important	−0.147	0.065	0.024
Control over life	−0.069	0.069	0.317
Family size	−0.086	0.062	0.165
Enthusiastic person	−0.143	0.099	0.150
Age	0.036	0.069	0.603
Number of purchases made	−0.068	0.043	0.150
Purchases per store	0.209	0.152	0.167
Shop on tight budget	−0.055	0.067	0.412
Excellent judge of quality	−0.070	0.089	0.435
Constant	3.250		
Overall R^2 = 0.477			

Multicollinearity

Multicollinearity
A state of high intercorrelations among independent variables.

Stepwise regression and multiple regression are complicated by the presence of **multi-collinearity**. Virtually all multiple regression analyses done in marketing research involve predictors or independent variables that are related. Multicollinearity, however, arises when intercorrelations amongst the predictors are very high.[24] Multicollinearity can result in several problems, including the following:

1 The partial regression coefficients may not be estimated precisely. The standard errors are likely to be high.

2 The magnitudes as well as the signs of the partial regression coefficients may change from sample to sample.

3 It becomes difficult to assess the relative importance of the independent variables in explaining the variation in the dependent variable.

4 Predictor variables may be incorrectly included or removed in stepwise regression.

What constitutes serious multicollinearity is not always clear, although several rules of thumb and procedures have been suggested in the literature. Procedures of varying complexity have also been suggested to cope with multicollinearity.[25] A simple procedure consists of using only one of the variables in a highly correlated set of variables.

Alternatively, the set of independent variables can be transformed into a new set of predictors that are mutually independent by using techniques such as principal components analysis (see Chapter 22). More specialised techniques, such as ridge regression and latent root regression, can also be used.[26]

Relative importance of predictors

When multicollinearity is present, special care is required in assessing the relative importance of independent variables. In marketing research, it is valuable to determine the relative importance of the predictors. In other words, how important are the independent variables in accounting for the variation in the criterion or dependent variable?[27] Unfortunately, because the predictors are correlated, there is no unambiguous measure of relative importance of the predictors in regression analysis.[28] Several approaches, however, are commonly used to assess the relative importance of predictor variables.

1 *Statistical significance.* If the partial regression coefficient of a variable is not significant, as determined by an incremental F test, that variable is judged to be unimportant. An exception to this rule is made if there are strong theoretical reasons for believing that the variable is important.

2 *Square of the simple correlation coefficient.* This measure, r^2, represents the proportion of the variation in the dependent variable explained by the independent variable in a bivariate relationship.

3 *Square of the partial correlation coefficient.* This measure, $R^2yx_i.x_j.x_k$, is the coefficient of determination between the dependent variable and the independent variable, controlling for the effects of the other independent variables.

4 *Square of the part correlation coefficient.* This coefficient represents an increase in R^2 when a variable is entered into a regression equation that already contains the other independent variables.

5 *Measures based on standardised coefficients or beta weights.* The most commonly used measures are the absolute values of the beta weights, $|\beta_i|$, or the squared values, β_i^2. Because they are partial coefficients, beta weights take into account the effect of the other independent variables. These measures become increasingly unreliable as the correlations among the predictor variables increase (multicollinearity increases).

6 *Stepwise regression.* The order in which the predictors enter or are removed from the regression equation is used to infer their relative importance.

Given that the predictors are correlated, at least to some extent, in virtually all regression situations, none of these measures is satisfactory. It is also possible that the different measures may indicate a different order of importance of the predictors.[29] Yet if all the measures are examined collectively, useful insights may be obtained into the relative importance of the predictors.

Cross-validation

Before assessing the relative importance of the predictors or drawing any other inferences, it is necessary to cross-validate the regression model. Regression and other multivariate procedures tend to capitalise on chance variations in the data. This could result in a regression model or equation that is unduly sensitive to the specific data used to estimate the model. One approach for evaluating the model for this and other problems associated with regression is cross-validation. **Cross-validation** examines whether the regression model continues to hold on comparable data not used in the estimation. The typical cross-validation procedure used in marketing research is as follows.

1 The regression model is estimated using the entire data set.
2 The available data are split into two parts, the *estimation sample* and the *validation sample*. The estimation sample generally contains 50 – 90% of the total sample.
3 The regression model is estimated using the data from the estimation sample only. This model is compared with the model estimated on the entire sample to determine the agreement in terms of the signs and magnitudes of the partial regression coefficients.
4 The estimated model is applied to the data in the validation sample to predict the values of the dependent variable, $\hat{Y}_i$ for the observations in the validation sample.
5 The observed values, Y_i, and the predicted values, $\hat{Y}_i$, in the validation sample are correlated to determine the simple r^2. This measure, r^2, is compared with R^2 for the total sample and with R^2 for the estimation sample to assess the degree of shrinkage.

A special form of validation is called **double cross-validation**. In double cross-validation the sample is split into halves. One half serves as the estimation sample, and the other is used as a validation sample in conducting cross-validation. The roles of the estimation and validation halves are then reversed, and the cross-validation is repeated.[30]

Regression with dummy variables

Cross-validation is a general procedure that can be applied even in some special applications of regression, such as regression with dummy variables. Nominal or categorical variables may be used as predictors or independent variables by coding them as dummy variables. The concept of dummy variables was introduced in Chapter 17. In that chapter, we explained how a categorical variable with four categories (heavy users, medium users, light users and non-users) can be coded in terms of three dummy variables, D_1, D_2 and D_3, as shown.

Suppose that the researcher was interested in running a regression analysis of the effect of attitude towards the brand on product use. The dummy variables D_1, D_2 and D_3 would be used as predictors. Regression with dummy variables would be modelled as

$$\hat{Y}_i = a + b_1 D_1 + b_2 D_2 + b_3 D_3$$

Product usage category	Original variable code	Dummy variable code		
		D_1	D_2	D_3
Non-users	1	1	0	0
Light users	2	0	1	0
Medium users	3	0	0	1
Heavy users	4	0	0	0

In this case, 'heavy users' have been selected as a reference category and have not been directly included in the regression equation. Note that for heavy users, D_1, D_2 and D_3 assume a value of 0, and the regression equation becomes

$$\hat{Y}_i = a$$

For non-users, $D_1 = 1$, and $D_2 = D_3 = 0$, and the regression equation becomes

$$\hat{Y}_i = a + b_1$$

Thus, the coefficient b_1 is the difference in predicted Y_i for non-users, as compared with heavy users. The coefficients b_2 and b_3 have similar interpretations. Although heavy users was selected as a reference category, any of the other three categories could have been selected for this purpose.[31]

Analysis of variance and covariance with regression

Regression with dummy variables provides a framework for understanding the analysis of variance and covariance. Although multiple regression with dummy variables provides a general procedure for the analysis of variance and covariance, we show only the equivalence of regression with dummy variables to one-way analysis of variance. In regression with dummy variables, the predicted $\hat{Y}$ for each category is the mean of Y for each category. To illustrate using the dummy variable coding of product use we just considered, the predicted $\hat{Y}$ and mean values for each category are as follows:

Product usage category	Predicted value	Mean value
	$\hat{Y}$	$\overline{Y}$
Non-users	$a + b_1$	$a + b_1$
Light users	$a + b_2$	$a + b_2$
Medium users	$a + b_3$	$a + b_3$
Heavy users	a	a

Given this equivalence, it is easy to see further relationships between dummy variable regression and one-way ANOVA.[32]

Thus, we see that regression in which the single independent variable with c categories has been recoded into $c - 1$ dummy variables is equivalent to one-way analysis of variance. Using similar correspondences, one can also illustrate how n-way analysis of variance and analysis of covariance can be performed using regression with dummy variables.

Dummy variable regression	**One-way ANOVA**
$SS_{res} = \sum\limits_{i=1}^{n} (Y_i - \hat{Y}_i)^2$	$= SS_{within}$
$SS_{reg} = \sum\limits_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2$	$= SS_{between}$
R^2	$= \eta^2$
Overall F test	$= F$ test

Internet and computer applications

The computer packages contain several programs to perform correlation analysis and regression analysis, calculating the associated statistics, performing tests for significance and plotting the residuals.

SPSS

CORRELATIONS can be used for computing Pearson product moment correlations, PARTIAL CORR for partial correlations, and NONPAR CORR for Spearman's ρs and Kendall's τ. The main program is REGRESSION which calculates bivariate and multiple regression equations, associated statistics and plots. It allows for easy examination of residuals. Stepwise regression can also be conducted. Regression statistics can be requested with PLOT, which produces simple scattergrams and some other types of plots.

SAS

The program CORR can be used for calculating Pearson, Spearman's, Kendall's and partial correlations. REG is a general-purpose regression procedure that fits bivariate and multiple regression models using the least-squares procedure. All the associated statistics are computed, and residuals can be plotted. Stepwise methods can be implemented. RSREG is a more specialised procedure that fits a quadratic response surface model using least squares regression. It is useful for determining factor levels that optimise a response. The ORTHOREG procedure is recommended for regression when the data are ill-conditioned. GLM uses the method of least squares to fit general linear models and can also be used for regression analysis. NLIN computes the parameters of a non-linear model using least squares or weighted least squares procedures.

Minitab

Correlation can be computed using the Stat>Basic statistics>Correlation function. It calculates Pearson's product moment. The Spearman's procedure ranks the columns first and then performs the correlation, on the ranked columns. To compute partial correlation, use the menu commands Stat>Basic Statistics>Correlation and Stat>Regression>Regression. Regression analysis, under the Stats>Regression function, can perform simple, polynomial and multiple analysis. The output includes a linear regression equation, table of coefficients, R^2, adjusted R^2, analysis of variance table, a table of fits and residuals that provide unusual observations. Other available features include stepwise, best subsets, fitted line plot and residual plots.

Excel

Correlations can be determined in Excel by using the Tools>Data analysis>Correlation function. Utilise the Correlation Worksheet function when a correlation coefficient for two cell ranges is needed. There is no separate function for partial correlations. Regression can be accessed from the Tools>Data analysis menu. Depending on the features selected, the output can consist of a summary output table, including an ANOVA table, a standard error of Y estimate, coefficients, standard error of coefficients, R^2 values and the number of observations. In addition, the function computes a residual output table, a residual plot, a line fit plot, a normal probability plot and a two-column probability data output table.

Summary

The product moment correlation coefficient, r, measures the linear association between two metric (interval or ratio scaled) variables. Its square, r^2, measures the proportion of variation in one variable explained by the other. The partial correlation coefficient measures the association between two variables after controlling, or adjusting for, the effects of one or more additional variables. The order of a partial correlation indicates how many variables are being adjusted or controlled. Partial correlations can be very helpful for detecting spurious relationships.

Bivariate regression derives a mathematical equation between a single metric criterion variable and a single metric predictor variable. The equation is derived in the form of a straight line by using the least squares procedure. When the regression is run on standardised data, the intercept assumes a value of 0, and the regression coefficients are called beta weights. The strength of association is measured by the coefficient of determination, r^2, which is obtained by computing a ratio of SS_{reg} to SS_y. The standard error of estimate is used to assess the accuracy of prediction and may be interpreted as a kind of average error made in predicting Y from the regression equation.

Multiple regression involves a single dependent variable and two or more independent variables. The partial regression coefficient, b_1, represents the expected change in Y when X_1 is changed by one unit and X_2 to X_k are held constant. The strength of association is measured by the coefficient of multiple determination, R^2. The significance of the overall regression equation may be tested by the overall F test. Individual partial regression coefficients may be tested for significance using the incremental F test. Scattergrams of the residuals, in which the residuals are plotted against the predicted values, $\hat{Y}_i$, time, or predictor variables, are useful for examining the appropriateness of the underlying assumptions and the regression model fitted.

In stepwise regression, the predictor variables are entered or removed from the regression equation one at a time for the purpose of selecting a smaller subset of predictors that account for most of the variation in the criterion variable. Multicollinearity, or very high intercorrelations among the predictor variables, can result in several problems. Because the predictors are correlated, regression analysis provides no unambiguous measure of relative importance of the predictors. Cross-validation examines whether the regression model continues to hold true for comparable data not used in estimation. It is a useful procedure for evaluating the regression model.

Nominal or categorical variables may be used as predictors by coding them as dummy variables. Multiple regression with dummy variables provides a general procedure for the analysis of variance and covariance.

Questions

1 What is the product moment correlation coefficient? Does a product moment correlation of 0 between two variables imply that the variables are not related to each other?

2 What are the main uses of regression analysis?

3 What is the least squares procedure?

4 Explain the meaning of standardised regression coefficients.

5 How is the strength of association measured in bivariate regression? In multiple regression?

6 What is meant by prediction accuracy?

7 What is the standard error of the estimate?

8 What is multiple regression? How is it different from bivariate regression?

9 Explain the meaning of a partial regression coefficient. Why is it called that?

10 State the null hypothesis in testing the significance of the overall multiple regression equation. How is this null hypothesis tested?

11 What is gained by an examination of residuals?

12 Explain the stepwise regression approach. What is its purpose?

13 What is multicollinearity? What problems can arise because of multicollinearity?

14 Describe the cross-validation procedure. Describe double cross-validation.

15 Demonstrate the equivalence of regression with dummy variables to one-way ANOVA.

Notes

1 Zajac, J., 'Avon's finally glowing thanks to global sales – and new lip-shtic', *Money* 26(9) (September 1997), 60; Miller, C., 'Computer modelling rings the right bell for Avon', *Marketing News* (9 May), 14.

2 Draper, N.R. and Smith, H., *Applied Regression Analysis*, 3rd edn (New York: Wiley, 1998); Doherty, M.E. and Sullivan, J.A., 'rho = p', *Organizational Behaviour and Human Decision Processes* 43(1) (February 1989), 136–44; Martin, W.S., 'Effects of scaling on the correlation coefficient: additional considerations', *Journal of Marketing Research* 15 (May 1978), 304–8; Bollen, K.A. and Barb, R.H., 'Pearson's *R* and coarsely categorized measures', *American Sociological Review* 46 (1981), 232–9.

3 Tacq, J., *Multivariate Analysis Techniques in Social Science Research* (Thousand Oaks, CA: Sage, 1997); Neter, J., Wasserman, W. and Kutner, M.J., *Applied Linear Statistical Models*, 3rd edn (Burr Ridge, IL: Irwin, 1990).

4 Although the topic is not discussed here, partial correlations can also be helpful in locating intervening variables and making certain types of causal inferences.

5 'Bates Saatchi & Saatchi, Budapest: accounting for change', *Accountancy* 116(224) (August 1995), 31; Kasriel, K., 'Hungary's million-dollar slap', *Advertising Age* (8 June 1992).

6 Another advantage to τ is that it can be generalised to a partial correlation coefficient. Pett, M.A., *Nonparametric Statistics for Health Care Research* (Thousand Oaks, CA: Sage, 1997); Siegel, S. and Castellan, N.J., *Nonparametric Statistics*, 2nd edn (New York: McGraw-Hill, 1988).

7 In a strict sense, the regression model requires that errors of measurement be associated only with the criterion variable and that the predictor variables be measured without error. For serially correlated errors, see Canjels, E. and Watson, M.W., 'Estimating deterministic trends in the presence of serially correlated errors', *Review of Economics and Statistics* 79(2) (May 1997), 184–200.

8 See any text on regression, such as Draper, N.R. and Smith, H., *Applied Regression Analysis*, 3rd edn (New York: Wiley, 1998); Neter, J., Wasserman, W. and Kutner, M.J., *Applied Linear Statistical Models*, 3rd edn (Burr Ridge, IL: Irwin, 1990).

9 Technically, the numerator is $b - \beta$. Since it has been hypothesised that $\beta = 0.0$, however, it can be omitted from the formula.

10 The larger the SEE, the poorer the fit of the regression.

11 The assumption of fixed levels of predictors applies to the 'classical' regression model. It is possible, if certain conditions are met, for the predictors to be random variables. Their distribution is not allowed to depend on the parameters of the regression equation. See Draper, N.R. and Smith, H., *Applied Regression Analysis*, 3rd edn (New York: Wiley, 1998).

12 For an approach to handling the violations of these assumptions, see Dispensa, G.S., 'Use logistic regression with customer satisfaction data', *Marketing News* 31(1) (6 January 1997), 13; Reddy, S.K., Holak, S.L. and Bhat, S., 'To extend or not to extend: success determinants of line extensions', *Journal of Marketing Research* 31 (May 1994), 243–62.

13 Rees, J., 'Tight ship keeps Coke on top of the world', *Marketing Week* 20(6) (8 May 1997), 28–9; Giges, N., 'Europeans buy outside goods, but like local ads', *Advertising Age International* (27 April 1992).

14 For other applications of multiple regression see Griffin, A., 'The effect of project and process characteristics on product development cycle time', *Journal of Marketing Research* 34 (February 1997), 24–35; Gatignon, H. and Xuereb, J.M., 'Strategic orientation of the firm and new product performance', *Journal of Marketing Research* 34 (February 1997), 77–90; Kumar, N., Scheer, L.K. and Steenkamp, J.B., 'The effects of supplier fairness on vulnerable resellers', *Journal of Marketing Research* 32 (February 1995), 54–65.

15 Yet another reason for adjusting R^2 is that as a result of the optimising properties of the least squares approach it is a maximum. Thus, to some extent, R^2 always overestimates the magnitude of a relationship. For applications of adjusted R^2 see Smith, N.C. and Cooper-Martin, E., 'Ethics and target marketing: the role of product harm and consumer vulnerability', *Journal of Marketing* 61(3) (January 1997), 1–20; Cohen, M.A. and Ho, T.H., 'An anatomy of a decision support system for developing and launching line extensions', *Journal of Marketing Research* 34 (February 1997), 117–29.

16 If R^2_{pop} is zero, then the sample R^2 reflects only sampling error, and the *F* ratio will tend to be equal to unity.

17 Another approach is the hierarchical method, in which the variables are added to the regression equation in an order specified by the researcher. For an application of multiple regression, see Corfman, K.P., 'Perceptions of relative influence: formation and measurement', *Journal of Marketing Research* 28 (May 1991), 125–36.

18 Atkinson, A.C., Koopman, S.J. and Shephard, N., 'Detecting shocks: outliers and breaks in time series', *Journal of Econometrics* 80(2) (October 1997), 387–422; Wang, G.C.S. and Akabay, C.K., 'Autocorrelation: problems and solutions in regression modeling', *Journal of Business Forecasting Methods and Systems* 13(4) (Winter 1994/95), 18–26; Belsley, D., Kuh, E. and Walsh, R.E., *Regression Diagnostics* (New York: Wiley, 1980).

19 The Durbin-Watson test is discussed in virtually all regression textbooks. See Draper, N.R. and Smith, H., *Applied Regression Analysis*, 3rd edn (New York: Wiley, 1998).

20 Berman, H., 'Selling the advertising/trade show partnership', *Folio: The Magazine for Magazine Management* (Special Sourcebook Issue for 1997 Supplement) 25(18) (1997), 214–15; Soley, L. and Krishnan, R., 'Does advertising subsidise consumer magazine prices?', *Journal of Advertising* 16 (Spring 1987), 4–9.

21 Fox, J., *Applied Regression Analysis, Linear Models and Related Methods* (Thousand Oaks, CA: Sage, 1997); McIntyre, S.H., Montgomery, D.B., Srinivasan, V. and Weitz, B.A., 'Evaluating the statistical significance of models developed by stepwise regression', *Journal of Marketing Research* 20 (February 1983), 1–11.

22 For applications of stepwise regression, see Ittner, D. and Larcker, D.F., 'Product development cycle time and organisational performance', *Journal of Marketing Research* 34 (February 1997), 13–23; Laroche, M. and Sadokierski, R., 'Role of confidence in a multi-brand model of intentions for a high-involvement service', *Journal of Business Research* 29 (January 1994), 1–12.

23 Crispell, D., 'Hispanics at the mall', *American Demographics* 19(10) (October 1997), 35–6; Jarboe, G.R. and McDaniel, C.D., 'A profile of browsers in regional shopping malls', *Journal of the Academy of Marketing Science* (Spring 1987), 46–53.

24 Greenberg, E. and Parks, R.P., 'A predictive approach to model selection and multicollinearity', *Journal of Applied Econometrics* 12(1) (January/February 1997), 67–75; Ofir, C. and Khuri, A., 'Multicollinearity in marketing models: diagnostics and remedial measures', *International Journal of Research in Marketing* 3 (1986), 181–205.

25 Possible procedures are given in Wang, G.C.S., 'How to handle multicollinearity in regression modelling', *Journal of Business Forecasting Methods and Systems* 15(1) (Spring 1996), 23–7; Mason, C.H. and Perreault Jr, W.D., 'Collinearity, power, and interpretation of multiple regression analysis', *Journal of Marketing Research* 28 (August 1991), 268–80; Hocking, R.R., 'Developments in linear regression methodology: 1959–1982', *Technometrics* 25 (August 1983), 219–30; Snee, R.D., 'Discussion', *Technometrics* 25 (August 1983), 230–7.

26 Holzworth, J.R., 'Policy capturing with ridge regression', *Organizational Behavior and Human Decision Processes* 68(2) (November 1996), 171–9; Wildt, A.R., 'Equity estimation and assessing market response', *Journal of Marketing Research* 31 (February 1994), 437–51; Sharma, S. and James, W.L., 'Latent root regression: an alternative procedure for estimating parameters in the presence of multicollinearity', *Journal of Marketing Research* (May 1981), 154–61.

27 Only relative importance can be determined, since the importance of an independent variable depends upon all the independent variables in the regression model.

28 Rugimbana, R., 'Predicting automated teller machine usage: the relative importance of perceptual and demographic factors', *International Journal of Bank Marketing* 13(4) (1995), 26–32; Green, P.E., Carroll, S.D. and DeSarbo, W.S., 'A new measure of predictor variable importance in multiple regression', *Journal of Marketing Research* (August 1978), 356–60; Jackson, B.B., 'Comment on "A new measure of predictor variable importance in multiple regression"', *Journal of Marketing Research* (February 1980), 116–18.

29 In the rare situation in which all the predictors are uncorrelated, simple correlations = partial correlations = part correlations = betas. Hence, the squares of these measures will yield the same rank order of the relative importance of the variables.

30 For more on cross-validation, see Litwin, M.S., *How to Measure Survey Reliability and Validity* (Thousand Oaks, CA: Sage, 1997); Song, X.M. and Perry, M.E., 'The determinants of Japanese new product success', *Journal of Marketing Research* 34 (February 1997), 64–76; Cooil, B., Winer, R.S. and Rados, D.L., 'Cross-validation for prediction', *Journal of Marketing Research* (August 1987), 271–9.

31 For an application of dummy variable regression, see Yavas, U., 'Demand forecasting in a service setting', *Journal of International Marketing and Marketing Research* 21(1) (February 1996), 3–11. For further discussion on dummy variable coding, see Cohen, J. and Cohen, P., *Applied Multiple Regression Correlation Analysis for the Behavioural Sciences*, 2nd edn (Hillsdale, NJ: Lawrence Erlbaum Associates, 1983), 181–222.

32 Fox, J., *Applied Regression Analysis, Linear Models and Related Methods* (Thousand Oaks, CA: Sage, 1997). For an application of regression analysis to conduct analysis of covariance, see Dancin, P.A. and Smith, D.C., 'The effect of brand portfolio characteristics on consumer evaluations of brand extensions', *Journal of Marketing Research* 31 (May 1994), 229–42.

Stage 1
Problem definition

Stage 2
Research approach
developed

Stage 3
Research design
developed

Stage 4
Fieldwork or data
collection

Stage 5
Data preparation
and analysis

Stage 6
Report preparation
and presentation

Discriminant analysis

Objectives

After reading this chapter, you should be able to:

1 describe the concept of discriminant analysis, its objectives and its applications in marketing research;

2 outline the procedures for conducting discriminant analysis, including the formulation of the problem, estimation of the discriminant function coefficients, determination of significance, interpretation and validation;

3 discuss multiple discriminant analysis and the distinction between two-group and multiple discriminant analysis;

4 explain stepwise discriminant analysis and describe the Mahalanobis procedure.

Discriminant analysis is used to estimate the relationship between a categorical dependent variable and a set of interval scaled, independent variables.

Overview

This chapter discusses the technique of discriminant analysis. We begin by examining the relationship of this procedure to regression analysis (Chapter 20) and analysis of variance (Chapter 19). We present a model and describe the general procedure for conducting discriminant analysis, with an emphasis on formulation, estimation, determination of significance, interpretation, and validation of the results. The procedure is illustrated with an example of two-group discriminant analysis, followed by an example of multiple (three-group) discriminant analysis. The stepwise discriminant analysis procedure is also covered.

We begin with examples illustrating the applications of two-group and multiple discriminant analysis.

e x a m p l e

GlobalCash Project

Two-group discriminant analysis

In the GlobalCash Project, two-group discriminant analysis was used. It helped to examine whether respondents who held accounts outside their home country, versus those who do not have accounts, attached different relative importance to nine factors of allocating business between banks. The dependent variable was whether the respondent's company held accounts outside their home country, and the independent variables were the importance attached to the nine factors of the choice criteria. The overall discriminant function was significant, indicating significant differences between the two groups. The results indicated that, compared with respondents who do not have accounts, those companies holding accounts outside their home country attached a greater relative importance to a good relationship with a bank, electronic banking systems, service quality and level of commitment to their business. ◼

e x a m p l e

An eye for a bargain[1]

A study of 294 consumers was undertaken to determine the correlates of 'discount proneness', in other words, the characteristics of consumers who respond favourably to direct mail promotions that offer a discount on the normal purchase price. The predictor variables were four factors related to household shopping attitudes and behaviour and selected demographic characteristics (gender, age and income). The dependent variable was the extent to which respondents were predisposed to take up the offer of a discount, of which three levels were identified. Respondents who reported no purchases triggered by a discount during the past 12 months were classified as *non-users*, those who reported one or two such purchases as *light users*, and those with more than two purchases as *frequent users* of discounts. Multiple discriminant analysis was used to analyse the data.

Two primary findings emerged. First, consumers' perception of the effort/value relationship was the most effective variable in discriminating among frequent users, light users and non-users of discount offers. Clearly, 'discount-sensitive' consumers associate less effort with fulfilling the requirements of the discounted purchase, and are willing to accept a relatively smaller refund, than other customers. Second, consumers who were aware of the regular prices of products, so that they recognise bargains, are more likely than others to respond to discount offers. ◼

In the GlobalCash example, there were two groups of respondents (accounts outside home country and no accounts outside home country), whereas the rebate predisposition example examined three groups (non-users, light users and frequent users of rebates). In both studies, significant inter-group differences were found using multiple predictor variables. An examination of differences across groups lies at the heart of the basic concept of discriminant analysis.

Basic concept

Discriminant analysis
A technique for analysing marketing research data when the criterion or dependent variable is categorical and the predictor or independent variables are interval in nature.

Discriminant analysis is a technique for analysing data when the criterion or dependent variable is categorical and the predictor or independent variables are interval in nature.[2] For example, the dependent variable may be the choice of a brand of personal computer (A, B or C) and the independent variables may be ratings of attributes of PCs on a seven-point Likert scale. The objectives of discriminant analysis are as follows:

1 Development of discriminant functions, or linear combinations of the predictor or independent variables, that best discriminate between the categories of the criterion or dependent variable (groups).
2 Examination of whether significant differences exist among the groups, in terms of the predictor variables.
3 Determination of which predictor variables contribute to most of the inter-group differences.
4 Classification of cases to one of the groups based on the values of the predictor variables.
5 Evaluation of the accuracy of classification.

Two-group discriminant analysis
Discriminant analysis technique where the criterion variable has two categories.

Multiple discriminant analysis
Discriminant analysis technique where the criterion variable involves three or more categories.

Discriminant function
The linear combination of independent variables developed by discriminant analysis that will best discriminate between the categories of the dependent variable.

Discriminant analysis techniques are described by the number of categories possessed by the criterion variable. When the criterion variable has two categories, the technique is known as **two-group discriminant analysis**. When three or more categories are involved, the technique is referred to as **multiple discriminant analysis**. The main distinction is that in the two-group case it is possible to derive only one **discriminant function**, but in multiple discriminant analysis more than one function may be computed.[3]

Examples of discriminant analysis abound in marketing research. This technique can be used to answer questions such as the following:[4]

■ In terms of demographic characteristics, how do customers who exhibit bank loyalty differ from those who do not?
■ Do heavy users, medium users and light users of soft drinks differ in terms of their consumption of frozen foods?
■ What psychographic characteristics help differentiate between price-sensitive and non-price-sensitive buyers of groceries?
■ Do market segments differ in their media consumption habits?
■ What are the distinguishing characteristics of consumers who respond to direct mail solicitations?

Relationship to regression and ANOVA

The relationships between discriminant analysis, analysis of variance (ANOVA) and regression analysis are shown in Table 21.1.

We explain these relationships with an example in which the researcher is attempting to explain the amount of life insurance purchased in terms of age and income. All three procedures involve a single criterion or dependent variable and multiple predictor or independent variables. The nature of these variables differs, however. In analysis of variance and regression analysis, the dependent variable is metric or interval scaled (amount of life insurance purchased in euros), whereas in discriminant analysis, it is categorical (amount of life insurance purchased classified as high, medium or low). The independent variables are categorical in the case of analysis of variance (age and income are each classified as high, medium or low) but metric in

Table 21.1 Similarities and differences among ANOVA, regression and discriminant analysis

	ANOVA	Regression	Discriminant analysis
Similarities			
Number of dependent variables	One	One	One
Number of independent variables	Multiple	Multiple	Multiple
Differences			
Nature of the dependent variable	Metric	Metric	Categorical
Nature of the independent variable	Categorical	Metric	Metric

the case of regression and discriminant analysis (age in years and income in euros, i.e. both measured on a ratio scale).

Two-group discriminant analysis, in which the dependent variable has only two categories, is closely related to multiple regression analysis. In this case, multiple regression, in which the dependent variable is coded as a 0 or 1 dummy variable, results in partial regression coefficients that are proportional to discriminant function coefficients.

Discriminant analysis model

Discriminant analysis model
The statistical model on which discriminant analysis is based.

The **discriminant analysis model** involves linear combinations of the following form:

$$D = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_kX_k$$

where D = discriminant score
   b = discriminant coefficients or weights
   X = predictor or independent variable.

The coefficients or weights (b) are estimated so that the groups differ as much as possible on the values of the discriminant function. This occurs when the ratio of between-group sum of squares to within-group sum of squares for the discriminant scores is at a maximum. Any other linear combination of the predictors will result in a smaller ratio. The technical details of estimation are described in Appendix 21A.

The following are important statistics associated with discriminant analysis.

Canonical correlation. Canonical correlation measures the extent of association between the discriminant scores and the groups. It is a measure of association between the single discriminant function and the set of dummy variables that define the group membership.

Centroid. The centroid is the mean values for the discriminant scores for a particular group. There are as many centroids as there are groups, as there is one for each group. The means for a group on all the functions are the group centroids.

Classification matrix. Sometimes also called confusion or prediction matrix, the classification matrix contains the number of correctly classified and misclassified cases. The correctly classified cases appear on the diagonal, because the predicted and actual groups are the same. The off-diagonal elements represent cases that have been incorrectly classified. The sum of the diagonal elements divided by the total number of cases represents the hit ratio.

Discriminant function coefficients. The discriminant function coefficients (unstandardised) are the multipliers of variables, when the variables are in the original units of measurement.

549

Discriminant scores. The unstandardised coefficients are multiplied by the values of the variables. These products are summed and added to the constant term to obtain the discriminant scores.

Eigenvalue. For each discriminant function, the eigenvalue is the ratio of between-group to within-group sums of squares. Large eigenvalues imply superior functions.

***F* values and their significance.** *F* values are calculated from a one-way ANOVA, with the grouping variable serving as the categorical independent variable. Each predictor, in turn, serves as the metric-dependent variable in the ANOVA.

Group means and group standard deviations. Group means and group standard deviations are computed for each predictor for each group.

Pooled within-group correlation matrix. The pooled within-group correlation matrix is computed by averaging the separate covariance matrices for all the groups.

Standardised discriminant function coefficients. The standardised discriminant function coefficients are the discriminant function coefficients that are used as the multipliers when the variables have been standardised to a mean of 0 and a variance of 1.

Structure correlations. Also referred to as discriminant loadings, the structure correlations represent the simple correlations between the predictors and the discriminant function.

Total correlation matrix. If the cases are treated as if they were from a single sample and the correlations are computed, a total correlation matrix is obtained.

Wilks' λ. Sometimes also called the *U* statistic, Wilks' λ for each predictor is the ratio of the within-group sum of squares to the total sum of squares. Its value varies between 0 and 1. Large values of λ (near 1) indicate that group means do not seem to be different. Small values of λ (near 0) indicate that the group means seem to be different.

The assumptions in discriminant analysis are that each of the groups is a sample from a multivariate normal population and that all the populations have the same covariance matrix. The role of these assumptions and the statistics just described can be better understood by examining the procedure for conducting discriminant analysis.

Conducting discriminant analysis

The steps involved in conducting discriminant analysis consist of formulation, estimation, determination of significance, interpretation and validation (see Figure 21.1). These steps are discussed and illustrated within the context of two-group discriminant analysis. Discriminant analysis with more than two groups is discussed later in this chapter.

Formulate the problem

The first step in discriminant analysis is to formulate the problem by identifying the objectives, the criterion variable and the independent variables. The criterion variable must consist of two or more mutually exclusive and collectively exhaustive categories. When the dependent variable is interval or ratio scaled, it must first be converted into categories. For example, attitude towards the brand, measured on a six-point scale, could be categorised as unfavourable (1, 2, 3) or favourable (4, 5, 6). Alternatively, one could plot the distribution of the dependent variable and form groups of equal size by determining the appropriate cut-off points for each category. The predictor variables should be selected based on a theoretical model or previous research, or in the case of exploratory research, the experience of the researcher should guide their selection.

Formulate the problem

Estimate the discriminant function coefficients

Determine the significance of the discriminant function

Interpret the results

Assess validity of discriminant analysis
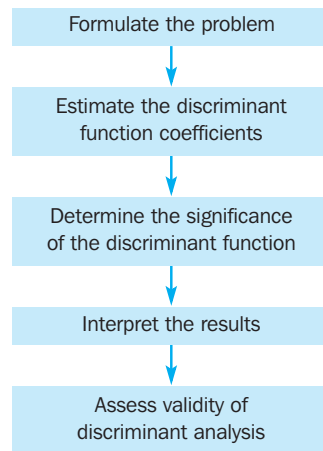
Figure 21.1
Conducting discriminant analysis

Analysis sample
Part of the total sample used to check the results of the discriminant function.

Validation sample
That part of the total sample used to check the results of the estimation sample.

The next step is to divide the sample into two parts. One part of the sample, called the *estimation* or **analysis sample**, is used for estimation of the discriminant function. The other part, called the *holdout* or validation sample, is reserved for validating the discriminant function. When the sample is large enough, it can be split in half. One half serves as the analysis sample, and the other is used for validation. The roles of the halves are then interchanged and the analysis is repeated. This is called double cross-validation and is similar to the procedure discussed in regression analysis (Chapter 20).

Often, the distribution of the number of cases in the analysis and **validation samples** follows the distribution in the total sample. For instance, if the total sample contained 50% loyal and 50% non-loyal consumers, then the analysis and validation samples would each contain 50% loyal and 50% non-loyal consumers. On the other hand, if the sample contained 25% loyal and 75% non-loyal consumers, the analysis and validation samples would be selected to reflect the same distribution (25% vs. 75%).

Finally, it has been suggested that the validation of the discriminant function should be conducted repeatedly. Each time, the sample should be split into different analysis and validation parts. The discriminant function should be estimated and the validation analysis carried out. Thus, the validation assessment is based on a number of trials. More rigorous methods have also been suggested.[5]

To illustrate two-group discriminant analysis better, let us look at an example. Suppose that we want to determine the salient characteristics of families that have visited a holiday resort during the last two years. Data were obtained from a pre-test sample of 42 households. Of these, 30 households, shown in Table 21.2, were included in the analysis sample and the remaining 12, shown in Table 21.3, were part of the validation sample. The households that visited a resort during the last two years were coded as 1; those that did not visit, as 2. Both the analysis and validation samples were balanced in terms of visit. As can be seen, the analysis sample contains 15 households in each category, whereas the validation sample had six in each category. Data were also obtained on annual family income (income), attitude towards travel (travel, measured on a nine-point scale), importance attached to family holiday (holiday, measured on a nine-point scale), household size (hsize), and age of the head of the household (age).

Estimate the discriminant function coefficients

Direct method
An approach to discriminant analysis that involves estimating the discriminant function so that all the predictors are included simultaneously.

Once the analysis sample has been identified, as in Table 21.2, we can estimate the discriminant function coefficients. Two broad approaches are available. The **direct method** involves estimating the discriminant function so that all the predictors are included simultaneously. In this case, each independent variable is included, regardless of its discriminating power. This method is appropriate when, based on previous research or a theoretical model, the researcher wants the discrimination to be based

Table 21.2 Information on resort visits: analysis sample

Number	Resort visit	Annual family income (€000)	Attitude towards travel	Importance attached to family holiday	Household size	Age of head of household	Amount spent on family holiday
1	1	50.2	5	8	3	43	M (2)
2	1	70.3	6	7	4	61	H (3)
3	1	62.9	7	5	6	52	H (3)
4	1	48.5	7	5	5	36	L (1)
5	1	52.7	6	6	4	55	H (3)
6	1	75.0	8	7	5	68	H (3)
7	1	46.2	5	3	3	62	M (2)
8	1	57.0	2	4	6	51	M (2)
9	1	64.1	7	5	4	57	H (3)
10	1	68.1	7	6	5	45	H (3)
11	1	73.4	6	7	5	44	H (3)
12	1	71.9	5	8	4	64	H (3)
13	1	56.2	1	8	6	54	M (2)
14	1	49.3	4	2	3	56	H (3)
15	1	62.0	5	6	2	58	H (3)
16	2	32.1	5	4	3	58	L (1)
17	2	36.2	4	3	2	55	L (1)
18	2	43.2	2	5	2	57	M (2)
19	2	50.4	5	2	4	37	M (2)
20	2	44.1	6	6	3	42	M (2)
21	2	38.3	6	6	2	45	L (1)
22	2	55.0	1	2	2	57	M (2)
23	2	46.1	3	5	3	51	L (1)
24	2	35.0	6	4	5	64	L (1)
25	2	37.3	2	7	4	54	L (1)
26	2	41.8	5	1	3	56	M (2)
27	2	57.0	8	3	2	36	M (2)
28	2	33.4	6	8	2	50	L (1)
29	2	37.5	3	2	3	48	L (1)
30	2	41.3	3	3	2	42	L (1)

Stepwise discriminant analysis
Discriminant analysis in which the predictors are entered sequentially based on their ability to discriminate between the groups.

on all the predictors. An alternative approach is the stepwise method. In **stepwise discriminant analysis**, the predictor variables are entered sequentially, based on their ability to discriminate among groups. This method, described in more detail later in this chapter, is appropriate when the researcher wants to select a subset of the predictors for inclusion in the discriminant function.

The results of running two-group discriminant analysis on the data of Table 21.2 using a popular statistical analysis package are presented in Table 21.4. Some intuitive feel for the results may be obtained by examining the group means and standard deviations. It appears that the two groups are more widely separated in terms of income than other variables, and there appears to be more of a separation on the

Table 21.3 Information on resort visits: validation sample

Number	Resort visit	Annual family income (€000)	Attitude towards travel	Importance attached to family holiday	Household size	Age of head of household	Amount spent on family holiday
1	1	50.8	4	7	3	45	M (2)
2	1	63.6	7	4	7	55	H (3)
3	1	54.0	6	7	4	58	M (2)
4	1	45.0	5	4	3	60	M (2)
5	1	68.0	6	6	6	46	H (3)
6	1	62.1	5	6	3	56	H (3)
7	2	35.0	4	3	4	54	L (1)
8	2	49.6	5	3	5	39	L (1)
9	2	39.4	6	5	3	44	H (3)
10	2	37.0	2	6	5	51	L (1)
11	2	54.5	7	3	3	37	M (2)
12	2	38.2	2	2	3	49	L (1)

Table 21.4 Results of two-group discriminant analysis

Group means

Visit	Income	Travel	Holiday	Hsize	Age
1	60.52000	5.40000	5.80000	4.33333	53.73333
2	41.91333	4.33333	4.06667	2.80000	50.13333
Total	51.21667	4.86667	4.93333	3.56667	51.93333

Group standard deviations

1	9.83065	1.91982	1.82052	1.23443	8.77062
2	7.55115	1.95180	2.05171	0.94112	8.27101
Total	12.79523	1.97804	2.09981	1.33089	8.57395

Pooled within-groups correlation matrix

Income	1.00000				
Travel	0.19745	1.00000			
Holiday	0.09148	0.08434	1.00000		
Hsize	0.08887	−0.01681	0.07046	1.00000	
Age	−0.01431	−0.19709	0.01742	−0.04301	1.00000

Wilks' λ (*U* statistic) and univariate *F* ratio with 1 and 28 degrees of freedom

Variable	Wilks' λ	F	Significance
Income	0.45310	33.800	0.0000
Travel	0.92479	2.277	0.1425
Holiday	0.82377	5.990	0.0209
Hsize	0.65672	14.640	0.0007
Age	0.95441	1.338	0.2572

Table 21.4 Continued

Canonical discriminant functions

Function	Eigenvalue	Per cent of variance	Cumulative percentage	Canonical correlation	After function	Wilks' λ	Chi-square	df	Sig
1*	1.7862	100.00	100.00	0.8007	0	0.3589	26.13	5	0.0001

* Marks the 1 canonical discriminant function remaining in the analysis

Standard canonical discriminant function coefficients

	Func 1
Income	0.74301
Travel	0.09611
Holiday	0.23329
Hsize	0.46911
Age	0.20922

Structure matrix: Pooled within-groups correlations between discriminating variables and canonical discriminant functions (variables ordered by size of correlation within function)

	Func 1
Income	0.82202
Hsize	0.54096
Holiday	0.34607
Travel	0.21337
Age	0.16354

Unstandardised canonical discriminant function coefficients

	Func 1
Income	0.8476710E-01
Travel	0.4964455E-01
Holiday	0.1202813
Hsize	0.4273893
Age	0.2454380E-01
(constant)	−7.975476

Canonical discriminant functions evaluated at group means (group centroids)

Group	Func 1
1	1.29118
2	−1.29118

Classification results for cases selected for use in analysis

			Predicted group membership	
	Actual group	No. of cases	1	2
Group	1	15	12	3
			80.0%	20.0%
Group	2	15	0	15
			0.0%	100%
Percentage of grouped cases correctly classified: 90%				

Table 21.4 **Continued**

Classification results for cases not selected for use in analysis (holdout sample)

			Predicted group membership	
	Actual group	*No. of cases*	*1*	*2*
Group	1	6	4	2
			66.7%	33.3%
Group	2	6	0	6
			0.0%	100%
Percentage of grouped cases correctly classified: 83.33%				

importance attached to the family holiday than on attitude towards travel. The difference between the two groups on age of the head of the household is small, and the standard deviation of this variable is large.

The pooled within-groups correlation matrix indicates low correlations between the predictors. Multicollinearity is unlikely to be a problem. The significance of the univariate F ratios indicates that, when the predictors are considered individually, only income, importance of holiday and household size significantly differentiate between those who visited a resort and those who did not.

Because there are two groups, only one discriminant function is estimated. The eigenvalue associated with this function is 1.7862, and it accounts for 100% of the explained variance. The canonical correlation associated with this function is 0.8007. The square of this correlation, $(0.8007)^2 = 0.64$, indicates that 64% of the variance in the dependent variable (visit) is explained or accounted for by this model. The next step is determination of significance.

Determine the significance of the discriminant function

It would not be meaningful to interpret the analysis if the discriminant functions estimated were not statistically significant. The null hypothesis that, in the population, the means of all discriminant functions in all groups are equal can be statistically tested. In SPSS, this test is based on Wilks' λ. If several functions are tested simultaneously (as in the case of multiple discriminant analysis), the Wilks' λ statistic is the product of the univariate λ for each function. The significance level is estimated based on a chi-square transformation of the statistic. In testing for significance in the holiday resort example (see Table 21.4), it may be noted that the Wilks' λ associated with the function is 0.3589, which transforms to a chi-square of 26.13 with 5 degrees of freedom. This is significant beyond the 0.05 level. In SAS, an approximate F statistic, based on an approximation to the distribution of the likelihood ratio, is calculated. If the null hypothesis is rejected, indicating significant discrimination, one can proceed to interpret the results.[6]

Interpret the results

The interpretation of the discriminant weights, or coefficients, is similar to that in multiple regression analysis. The value of the coefficient for a particular predictor depends on the other predictors included in the discriminant function. The signs of the coefficients are arbitrary, but they indicate which variable values result in large and small function values and associate them with particular groups.

Given the multicollinearity in the predictor variables, there is no unambiguous measure of the relative importance of the predictors in discriminating between the groups.[7] With this caveat in mind, we can obtain some idea of the relative importance of the variables by examining the absolute magnitude of the standardised discriminant function coefficients. Generally, predictors with relatively large standardised

coefficients contribute more to the discriminating power of the function, as compared with predictors with smaller coefficients.

Some idea of the relative importance of the predictors can also be obtained by examining the structure correlations, also called canonical loadings or discriminant loadings. These simple correlations between each predictor and the discriminant function represent the variance that the predictor shares with the function. Like the standardised coefficients, these correlations must also be interpreted with caution.

An examination of the standardised discriminant function coefficients for the holiday resort example is instructive. Given the low intercorrelations between the predictors, one might cautiously use the magnitudes of the standardised coefficients to suggest that income is the most important predictor in discriminating between the groups, followed by household size and importance attached to the family holiday. The same observation is obtained from examination of the structure correlations. These simple correlations between the predictors and the discriminant function are listed in order of magnitude.

The unstandardised discriminant function coefficients are also given. These can be applied to the raw values of the variables in the holdout set for classification purposes. The group centroids, giving the value of the discriminant function evaluated at the group means, are also shown. Group 1, those who have visited a resort, has a positive value, whereas Group 2 has an equal negative value. The signs of the coefficients associated with all the predictors are positive, which suggests that higher family income, household size, importance attached to family holiday, attitude towards travel, and age are more likely to result in the family visiting the resort. It would be reasonable to develop a profile of the two groups in terms of the three predictors that seem to be the most important: income, household size, and importance of holiday. The values of these three variables for the two groups are given at the beginning of Table 21.4.

The determination of relative importance of the predictors is further illustrated by the following example.

<table>
<tr><td>*e x a m p l e*</td></tr>
</table>

Satisfied salespeople stay[8]

Discriminant analysis was used to determine what factors explained the differences between salespeople who left a large computer manufacturing company and those who stayed. The independent variables were company rating, job security, seven job satisfaction dimensions, four role-conflict dimensions, four role-ambiguity dimensions, and nine measures of sales performance. The dependent variable was the dichotomy between those who stayed and those who left. The canonical correlation, an index of discrimination ($R = 0.4572$), was significant (Wilks' $\lambda = 0.7909$, $F(26,173) = 1.7588$, $p = 0.0180$). This result indicated that the variables discriminated between those who left and those who stayed.

The results from simultaneously entering all variables in discriminant analysis are presented in the table opposite. The rank order of importance, as determined by the relative magnitude of the canonical loadings, is presented in the first column. Satisfaction with the job and promotional opportunities were the two most important discriminators, followed by job security. Those who stayed in the company found the job to be more exciting, satisfying, challenging and interesting than those who left. ■

In this example, promotion was identified as the second most important variable based on the canonical loadings. However, it is not the second most important variable based on the absolute magnitude of the standardised discriminant function coefficients. This anomaly results from multicollinearity.

Another aid to interpreting discriminant analysis results is to develop a **characteristic profile** for each group by describing each group in terms of the group means for the predictor variables. If the important predictors have been identified, then a comparison of the group means on these variables can assist in understanding the intergroup differences. Before any findings can be interpreted with confidence, however, it is necessary to validate the results.

Characteristic profile
An aid to interpreting discriminant analysis results by describing each group in terms of the group means for the predictor variables.

Discriminant analysis results

Variable	Coefficients	Standardised coefficients	Canonical loadings
1 Work[a]	0.0903	0.3910	0.5446
2 Promotion[a]	0.0288	0.1515	0.5044
3 Job security	0.1567	0.1384	0.4958
4 Customer relations[b]	0.0086	0.1751	0.4906
5 Company rating	0.4059	0.3240	0.4824
6 Working with others[b]	0.0018	0.0365	0.4651
7 Overall performance[b]	4.0148	–0.3252	0.4518
8 Time-territory management[b]	0.0126	0.2899	0.4496
9 Sales produced[b]	0.0059	0.1404	0.4484
10 Presentation skill[b]	0.0118	0.2526	0.4387
11 Technical information[b]	0.0003	0.0065	0.4173
12 Pay-benefits[a]	0.0600	0.1843	0.3788
13 Quota achieved[b]	0.0035	0.2915	0.3780
14 Management[a]	0.0014	0.0138	0.3571
15 Information collection[b]	–0.0146	4.3327	0.3326
16 Family[c]	–0.0684	–0.3408	–0.3221
17 Sales manager[a]	4.0121	–0.1102	0.2909
18 Coworker[a]	0.0225	0.0893	0.2671
19 Customer[c]	–0.0625	4.2797	–0.2602
20 Family[d]	0.0473	0.1970	0.2180
21 Job[d]	0.1378	0.5312	0.2119
22 Job[c]	0.0410	0.5475	–0.1029
23 Customer[d]	–0.0060	4.0255	0.1004
24 Sales manager[c]	–0.0365	–0.2406	–0.0499
25 Sales manager[d]	–0.0606	–0.3333	0.0467
26 Customer[a]	–0.0338	–0.1488	0.0192

Note: Rank order of importance is based on the magnitude of the canonical loadings:
[a] Satisfaction
[b] Performance
[c] Ambiguity
[d] Conflict

Assess the validity of discriminant analysis

As explained earlier, the data are randomly divided into two sub-samples. One, the analysis sample, is used for estimating the discriminant function, and the validation sample is used for developing the classification matrix. The discriminant weights, estimated by using the analysis sample, are multiplied by the values of the predictor variables in the holdout sample to generate discriminant scores for the cases in the holdout sample. The cases are then assigned to groups based on their discriminant scores and an appropriate decision rule. For example, in two-group discriminant analysis, a case will be assigned to the group whose centroid is the closest. The **hit ratio**, or the percentage of cases correctly classified, can then be determined by summing the diagonal elements and dividing by the total number of cases.[9]

Hit ratio
The percentage of cases correctly classified by the discriminant analysis.

It is helpful to compare the percentage of cases correctly classified by discriminant analysis with the percentage that would be obtained by chance. When the groups are equal in size, the percentage of chance classification is 1 divided by the number of groups. How much improvement should be expected over chance? No general guidelines are available, although some authors have suggested that classification accuracy achieved by discriminant analysis should be at least 25% greater than that obtained by chance.[10]

Most discriminant analysis programs also estimate a classification matrix based on the analysis sample. Because they capitalise on chance variation in the data, such results are invariably better than the classification obtained on the holdout sample.[11]

Table 21.4, of the holiday resort example, also shows the classification results based on the analysis sample. The hit ratio, or the percentage of cases correctly classified, is $(12 + 15)/30 = 0.90$, or 90%. One might suspect that this hit ratio is artificially inflated, as the data used for estimation were also used for validation. Conducting classification analysis on an independent holdout set of data results in the classification matrix with a slightly lower hit ratio of $(4 + 6)/12 = 0.833$, or 83.3% (see Table 21.4). Given two groups of equal size, by chance one would expect a hit ratio of $1/2 = 0.50$, or 50%. Hence, the improvement over chance is more than 25%, and the validity of the discriminant analysis is judged as satisfactory.

Another application of two-group discriminant analysis is provided by the following example.

<div style="border:1px solid #000;">**e x a m p l e**</div>

Home bodies and couch potatoes[12]

Two-group discriminant analysis was used to assess the strength of each of five dimensions used in classifying individuals as TV users or non-users. The discriminant-analysis procedure was appropriate for this use because of the nature of the predefined categorical groups (users and non-users) and the interval scales used to generate individual factor scores.

Two equal groups of 185 elderly consumers, users and non-users (total $n = 370$), were created. The discriminant equation for the analysis was estimated by using a sub-sample of 142 respondents from the sample of 370. Of the remaining respondents, 198 were used as a validation sub-sample in a cross-validation of the equation. Thirty respondents were excluded from the analysis because of missing discriminant values.

The canonical correlation for the discriminant function was 0.4291, significant at the $p < 0.0001$ level. The eigenvalue was 0.2257. The table opposite summarises the standardised canonical discriminant coefficients. A substantial portion of the variance is explained by the discriminant function. In addition, as the table shows, the home-orientation dimension made a fairly strong contribution to classifying individuals as users or non-users of television. Morale, security and health, and respect also contributed significantly. The social factor appeared to make little contribution.

The cross-validation procedure using the discriminant function from the analysis sample gave support to the contention that the dimensions aided researchers in discriminating between users and non-users of television. As the table shows, the discriminant function was successful in classifying 75.76% of the cases. This suggests that consideration of the identified dimensions will help marketers understand the elderly market. ■

The extension from two-group discriminant analysis to multiple discriminant analysis involves similar steps and is illustrated with an application.

Conducting multiple discriminant analysis

Formulate the problem

The data presented in Tables 21.2 and 21.3 can also be used to illustrate three-group discriminant analysis. In the last column of these tables, the households are classified into three categories, based on the amount spent on their family holiday (high, medium or low). Ten households fall in each category. The question of interest is

Summary of discriminant analysis

Standard canonical discriminant function coefficients

Morale	0.27798
Security and health	0.39850
Home orientation	0.77496
Respect	0.32069
Social	−0.01996

Classification results for cases selected for use in the analysis

		Predicted group membership	
Actual group	*No. of cases*	*Non-users*	*Users*
TV non-users	77	56	21
		72.7%	27.3%
TV users	65	24	41
		36.9%	63.1%
Per cent of grouped cases correctly classified: 68.31%			

Classification results for cases selected for cross-validation

		Predicted group membership	
Actual group	*No. of cases*	*Non-users*	*Users*
TV non-users	108	85	23
		78.7%	21.3%
TV users	90	25	65
		27.8%	72.2%
Per cent of grouped cases correctly classified: 75.76%			

whether the households that spend high, medium or low amounts on their holidays (amount) can be differentiated in terms of annual family income (income), attitude towards travel (travel), importance attached to family holiday (holiday), household size (hsize), and age of the head of household (age).[13]

Estimate the discriminant function coefficients

Table 21.5 presents the results of estimating three-group discriminant analysis. An examination of group means indicates that income appears to separate the groups more widely than any other variable. There is some separation on travel and holiday. Groups 1 and 2 are very close in terms of household size and age. Age has a large standard deviation relative to the separation between the groups. The pooled within-groups correlation matrix indicates some correlation of holiday and household size with income. Age has some negative correlation with travel. Yet these correlations are on the lower side, indicating that, although multicollinearity may be of some concern, it is not likely to be a serious problem. The significance attached to the univariate F ratios indicates that, when the predictors are considered individually, only income and travel are significant in differentiating between the two groups.

In multiple discriminant analysis, if there are G groups, $G - 1$ discriminant functions can be estimated if the number of predictors is larger than this quantity. In general, with G groups and k predictors, it is possible to estimate up to the smaller of $G - 1$, or k, discriminant functions. The first function has the highest ratio of between-groups to within-groups sum of squares. The second function, uncorrelated

Table 21.5 Results of three-group discriminant analysis

Group means

Visit	Income	Travel	Holiday	Hsize	Age
1	38.57000	4.50000	4.70000	3.10000	50.30000
2	50.1100	4.00000	4.20000	3.40000	49.50000
3	64.97000	6.10000	5.90000	4.20000	56.00000
Total	51.21667	4.86667	4.93333	3.56667	51.93333

Group standard deviations

1	5.29718	1.71594	1.88856	1.19722	8.09732
2	6.00231	2.35702	2.48551	1.50555	9.25263
3	8.61434	1.19722	1.66333	1.13529	7.60117
Total	12.79523	1.97804	2.09981	1.33089	8.57395

Pooled within-groups correlation matrix

Income	1.00000				
Travel	0.05120	1.00000			
Holiday	0.30681	0.03588	1.00000		
Hsize	0.38050	0.00474	0.22080	1.00000	
Age	−0.20939	−0.34022	−0.01326	−0.02512	1.00000

Wilks' λ (U statistic) and univariate F ratio with 2 and 27 degrees of freedom

Variable	Wilks' λ	F	Significance
Income	0.26215	38.000	0.0000
Travel	0.78790	3.634	0.0400
Holiday	0.88060	1.830	0.1797
Hsize	0.87411	1.944	0.1626
Age	0.88214	1.804	0.1840

Canonical discriminant functions

Function	Eigenvalue	Per cent of variance	Cumulative percentage	Canonical correlation	After function	Wilks' λ	Chi-square	df.	Sig
					0	0.1664	44.831	10	0.00
1*	3.8190	93.93	93.93	0.8902	1	0.8020	5.517	4	0.24
2*	0.2469	6.07	100.00	0.4450					

* Marks the two canonical discriminant functions remaining in the analysis

Standardised canonical discriminant function coefficients

	Func 1	Func 2
Income	1.04740	−0.42076
Travel	0.33991	0.76851
Holiday	−0.14198	0.53354
Hsize	−0.16317	0.12932
Age	0.49474	0.52447

Table 21.5 **Continued**

Structure matrix: Pooled within-groups correlations between discriminating variables and canonical discriminant functions (variables ordered by size of correlation within function)

	Func 1	Func 2
Income	0.85556*	−0.27833
Hsize	0.19319*	0.07749
Holiday	0.21935	0.58829*
Travel	0.14899	0.45362*
Age	0.16576	0.34079*

Unstandardised canonical discriminant function coefficients

	Func 1	Func 2
Income	0.1542658	−0.6197148E-01
Travel	0.1867977	0.4223430
Holiday	−0.6952264E-01	0.2612652
Hsize	−0.1265334	0.1002796
Age	0.5928055E-01	0.6284206E-01
(constant)	−11.09442	−3.791600

Canonical discriminant functions evaluated at group means (group centriods)

Group	Func 1	Func 2
1	−2.04100	0.41847
2	−0.40479	−0.65867
3	2.44578	0.24020

Classification results for cases selected for use in analysis

	Actual group	No. of cases	Predicted group membership 1	2	3
Group	1	10	9	1	0
			90.0%	10.0%	.0%
Group	2	10	1	9	0
			10.0%	90.0%	.0%
Group	3	10	0	2	8
			.0%	20.0%	80.0%
Percentage of grouped cases correctly classified: 86.67%					

Classification results for cases selected for use in analysis

	Actual group	No. of cases	Predicted group membership 1	2	3
Group	1	4	3	1	0
			75.0%	25.0%	.0%
Group	2	4	0	3	1
			.0%	75.0%	25.0%
Group	3	4	1	0	3
			25.0%	.0%	75.0%
Percentage of grouped cases correctly classified: 75.00%					

with the first, has the second highest ratio, and so on. Not all the functions may be statistically significant, however.

Because there are three groups, a maximum of two functions can be extracted. The eigenvalue associated with the first function is 3.8190, and this function accounts for 93.93% of variance in the data. The eigenvalue is large, so the first function is likely to be superior. The second function has a small eigenvalue of 0.2469 and accounts for only 6.07% of the variance.

Determine the significance of the discriminant function

To test the null hypothesis of equal group centroids, both the functions must be considered simultaneously. It is possible to test the means of the functions successively by first testing all means simultaneously. Then one function is excluded at a time, and the means of the remaining functions are tested at each step. In Table 21.5, the 0 below the 'After function' heading indicates that no functions have been removed. The value of Wilks' λ is 0.1644. This transforms to a chi-square of 44.831, with 10 degrees of freedom, which is significant beyond the 0.05 level. Thus, the two functions together significantly discriminate among the three groups. When the first function is removed, however, the Wilks' λ associated with the second function is 0.8020, which is not significant at the 0.05 level. Therefore, the second function does not contribute significantly to group differences.

Interpret the results

The interpretation of the results is aided by an examination of the standardised discriminant function coefficients, the structure correlations and certain plots. The standardised coefficients indicate a large coefficient for income on function 1, whereas function 2 has relatively larger coefficients for travel, holiday and age. A similar conclusion is reached by an examination of the structure matrix (see Table 21.5). To help interpret the functions, variables with large coefficients for a particular function are grouped together. These groupings are shown with asterisks. Thus income and household size have asterisks for function 1 because these variables have coefficients which are larger for function 1 than for function 2. These variables are associated primarily with function 1. On the other hand, travel, holiday and age are predominantly associated with function 2, as indicated by the asterisks.

Figure 21.2 is a scattergram plot of all the groups on function 1 and function 2. It can be seen that group 3 has the highest value on function 1, and group 1 the lowest. Because function 1 is primarily associated with income and household size, one would expect the three groups to be ordered on these two variables. Those with higher incomes and higher household size are likely to spend large amounts of money on holidays. Conversely, those with low incomes and smaller household size are likely to spend small amounts on holidays. This interpretation is further strengthened by an examination of group means on income and household size.

Figure 21.2 further indicates that function 2 tends to separate group 1 (highest value) and group 2 (lowest value). This function is primarily associated with travel, holiday and age. Given the positive correlations of these variables with function 2 in the structure matrix, we expect to find group 1 to be higher than group 2 in terms of travel, holiday and age. This is indeed true for travel and holiday, as indicated by the group means of these variables. If families in group 1 have more favourable attitudes towards travel and attach more importance to family holiday than group 2, why do they spend less? Perhaps they would like to spend more on holidays but cannot afford it because they have low incomes.
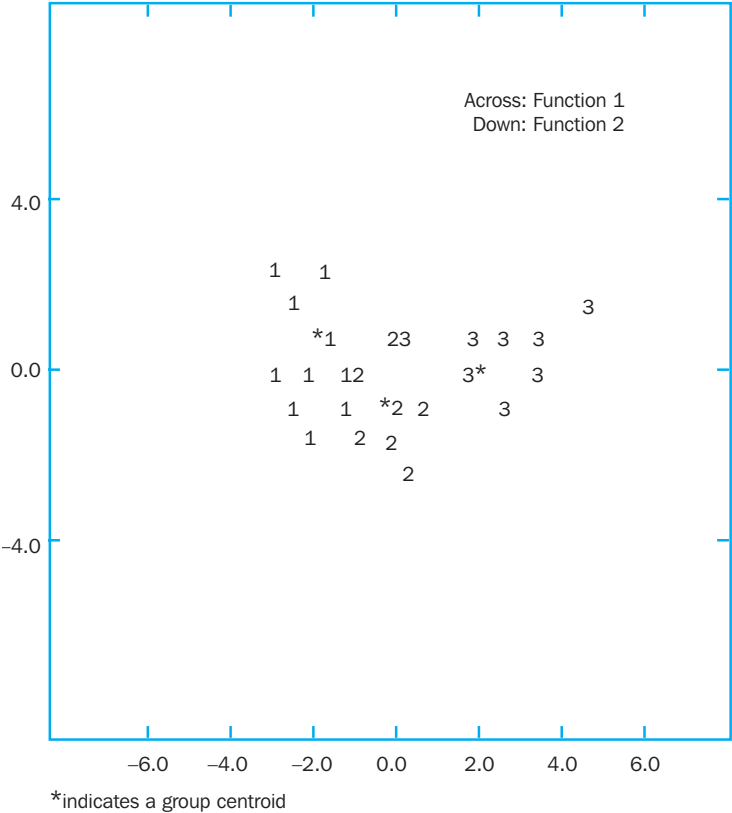
Figure 21.2
All-groups scattergram

*indicates a group centroid
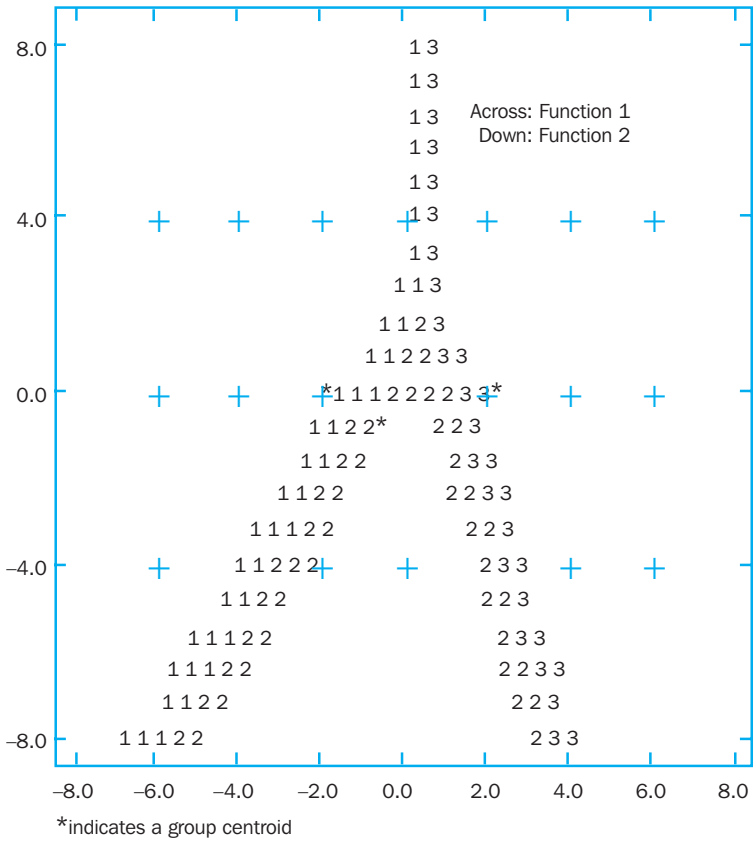
Figure 21.3
Territorial map

*indicates a group centroid

A similar interpretation is obtained by examining a **territorial map**, as shown in Figure 21.3. In a territorial map, each group centroid is indicated by an asterisk. The group boundaries are shown by numbers corresponding to the groups. Thus, group 1 centroid is bounded by ls, group 2 centroid by 2s, and group 3 centroid by 3s.

Assess the validity of discriminant analysis

The classification results based on the analysis sample indicate that $(9 + 9 + 8)/30 = 86.67\%$ of the cases are correctly classified. When the classification analysis is conducted on the independent holdout sample of Table 21.3, a slightly lower hit ratio of 83.3% is obtained. Given three groups of equal size, by chance alone one would expect a hit ratio of $1/3 = 0.333$ or 33.3%. The improvement over chance is more than 25%, indicating at least satisfactory validity.[14]

Further illustration of multiple group discriminant analysis is provided by the following example.

e x a m p l e

The home is where the patient's heart is[15]

Consumers were surveyed to determine their attitudes towards four systems of health care delivery (home health care, hospitals, nursing homes and outpatient clinics) along 10 attributes. A total of 102 responses were obtained, and the results were analysed using multiple discriminant analysis (Table 1 below). Three discriminant functions were identified. Chi-square tests performed on the results indicated that all three discriminant functions were significant at the 0.01 level. The first function accounted for 63% of the total discriminative power, and the remaining two functions contributed 29.4% and 7.6%, respectively.

Table 1 gives the standardised discriminant function coefficients of the 10 variables in the discriminant equations. Coefficients ranged in value from −1 to +1. In determining the ability of each attribute to classify the delivery system, absolute values were used. In the first discriminant function, the two variables with the largest coefficients were comfort (0.53) and privacy (0.40). Because both related to personal attention and care, the first dimension was labelled 'personalised care'. In the second function, the two variables with the largest coefficients were quality of medical care (0.67) and likelihood of faster recovery (0.32). Hence, this dimension was labelled 'quality of medical care'. In the third discriminant function, the

Table 1 Standardised discriminant function coefficients

Variable	*Discriminant function*		
	1	*2*	*3*
Safe	−0.20	−0.04	0.15
Convenient	0.08	0.08	0.07
Chance of medical complications[a]	−0.27	0.10	0.16
Expensive[a]	0.30	−0.28	0.52
Comfortable	0.53	0.27	−0.19
Sanitary	−0.27	−0.14	−0.70
Best medical care	−0.25	0.67	−0.10
Privacy	0.40	0.08	0.49
Faster recovery	0.30	0.32	−0.15
Staffed with best medical personnel	−0.17	−0.03	0.18
Percentage of variance explained	63.0	29.4	7.6
Chi-square	663.3[b]	289.2[b]	70.1[b]

[a]These two items were worded negatively on the questionnaire. They were reverse coded for purposes of data analysis.
[b] $p < 0.01$.

Table 2 Centroids of health care systems in discriminant space

System	Discriminant function		
	1	2	3
Hospital	−1.66	0.97	−0.08
Nursing home	−0.60	−1.36	−0.27
Outpatient clinic	0.54	−0.13	0.77
Home health care	1.77	0.50	−0.39

Table 3 Classification table

System	Hospital	Classification (%)		
		Nursing home	Outpatient clinic	Home health care
Hospital	86	6	6	2
Nursing home	9	78	10	3
Outpatient clinic	9	13	68	10
Home health care	5	4	13	78

most significant attributes were sanitation (− 0.70) and expense (0.52). Because these two attributes represent value and price, the third discriminant function was labelled 'value'.

The four group centroids are shown in Table 2. This table shows that home health care was evaluated most favourably along the dimension of personalised care, and hospitals were evaluated least favourably. Along the dimension of quality of medical care, there was a substantial separation between nursing homes and the other three systems. Also, home health care received higher evaluations on the quality of medical care than did outpatient clinics. Outpatient clinics, on the other hand, were judged to offer the best value.

Classification analysis of the 102 responses, reported in Table 3, showed correct classifications ranging from 86% for hospitals to 68% for outpatient clinics. The mis-classifications for hospitals were 6% each to nursing homes and outpatient clinics, and 2% to home health care. Nursing homes showed mis-classifications of 9% to hospitals, 10% to outpatient clinics, and 3% to home health care. For outpatient clinics, 9% mis-classifications were made to hospitals, 13% to nursing homes and 10% to home health care. For home health care, the mis-classifications were 5% to hospitals, 4% to nursing homes and 13% to outpatient clinics. The results demonstrated that the discriminant functions were fairly accurate in predicting group membership. ◼

Stepwise discriminant analysis

Stepwise discriminant analysis is analogous to stepwise multiple regression (see Chapter 20) in that the predictors are entered sequentially based on their ability to discriminate between the groups. An *F* ratio is calculated for each predictor by conducting a univariate analysis of variance in which the groups are treated as the categorical variable and the predictor as the criterion variable. The predictor with the highest *F* ratio is the first to be selected for inclusion in the discriminant function, if it meets certain significance and tolerance criteria. A second predictor is added based on the highest adjusted or partial *F* ratio, taking into account the predictor already selected.

Each predictor selected is tested for retention based on its association with other predictors selected. The process of selection and retention is continued until all predictors meeting the significance criteria for inclusion and retention have been entered in the discriminant function. Several statistics are computed at each stage. In addition, at the conclusion, a summary of the predictors entered or removed is provided. The standard output associated with the direct method is also available from the stepwise procedure.

The selection of the stepwise procedure is based on the optimising criterion adopted. The **Mahalanobis procedure** is based on maximising a generalised measure of the distance between the two closest groups. This procedure allows marketing researchers to make maximal use of the available information.[16]

The Mahalanobis method was used to conduct a two-group stepwise discriminant analysis on the data pertaining to the visit variable in Tables 21.2 and 21.3. The first predictor variable to be selected was income, followed by household size and then holiday. The order in which the variables were selected also indicates their importance in discriminating between the groups. This was further corroborated by an examination of the standardised discriminant function coefficients and the structure correlation coefficients. Note that the findings of the stepwise analysis agree with the conclusions reported earlier by the direct method.

Mahalanobis procedure
A stepwise procedure used in discriminant analysis to maximise a generalised measure of the distance between the two closest groups.

e x a m p l e

Satisfactory results of satisfaction programmes in Europe[17]

These days, more and more computer companies are emphasising customer service programmes rather than their erstwhile emphasis on computer features and capabilities. Hewlett-Packard learned this lesson in Europe. Research conducted in the European market revealed that there was a difference in emphasis on service requirements across age segments. Focus groups revealed that customers above 40 years of age had a hard time with the technical aspects of the computer and greatly required the customer service programmes. On the other hand, young customers appreciated the technical aspects of the product that added to their satisfaction. To uncover the factors leading to differences in the two segments, further research in the form of a large single cross-sectional survey was done. A two-group discriminant analysis was conducted with satisfied and dissatisfied customers as the two groups, with several inde-

Consumers over 40 years of age can have a hard time with the technical aspects of the computer.

pendent variables such as technical information, ease of operation, variety and scope of customer service programmes, etc. Results confirmed the fact that the variety and scope of customer satisfaction programmes was indeed a strong differentiating factor. This was a crucial finding because Hewlett-Packard could better handle dissatisfied customers by focusing more on customer services than on technical details. Consequently, Hewlett-Packard successfully started three programmes on customer satisfaction: customer feedback, customer satisfaction surveys, and total quality control. This effort resulted in increased customer satisfaction. ■

Internet and computer applications

SPSS

In the mainframe version, the DISCRIMINANT procedure is used for conducting discriminant analysis. This is a general program that can be used for two-group or multiple discriminant analysis. Furthermore, the direct or the stepwise method can be adopted. A similar program, DISCRIMINANT, is available in the PC version.

SAS

The DISCRIM procedure can be used for performing two-group or multiple discriminant analysis. If the assumption of a multivariate normal distribution cannot be met, the NEIGHBOR procedure can be used. In this procedure, a non-parametric nearest-neighbour rule is used for classifying the observations. CANDISC performs canonical discriminant analysis and is related to principal components analysis and canonical correlation. The STEPDISC procedure can be used for performing stepwise discriminant analysis. The mainframe and microcomputer versions are similar, except that the program NEIGHBOR is not available on the microcomputer version.

Minitab

Discriminant analysis can be conducted using the Stats>Multivariate>Discriminant Analysis function. It computes both linear and quadratic discriminant analysis in the classification of observations into two or more groups.

Excel

At the time of writing, discriminant analysis was not available.

Summary

Discriminant analysis is useful for analysing data when the criterion or dependent variable is categorical and the predictor or independent variables are interval scaled. When the criterion variable has two categories, the technique is known as two-group discriminant analysis. Multiple discriminant analysis refers to the case when three or more categories are involved.

Conducting discriminant analysis is a five-step procedure. First, formulating the discriminant problem requires identification of the objectives and the criterion and predictor variables. The sample is divided into two parts. One part, the analysis sample, is used to estimate the discriminant function. The other part, the holdout sample, is reserved for validation. Estimation, the second step, involves developing a linear combination of the predictors, called discriminant functions, so that the groups differ as much as possible on the predictor values.

Determination of statistical significance is the third step. It involves testing the null hypothesis that, in the population, the means of all discriminant functions in all groups are equal. If the null hypothesis is rejected, it is meaningful to interpret the results.

The fourth step, the interpretation of discriminant weights or coefficients, is similar to that in multiple regression analysis. Given the multicollinearity in the predictor variables, there is no unambiguous measure of the relative importance of the predictors in discriminating between the groups. Some idea of the relative importance of the variables, however, may be obtained by examining the absolute magnitude of the standardised discriminant function coefficients and by examining the structure correlations or discriminant loadings. These simple correlations between each predictor and the discriminant function represent the variance that the predictor shares with the function. Another aid to interpreting discriminant analysis results is to develop a characteristic profile for each group, based on the group means for the predictor variables.

Validation, the fifth step, involves developing the classification matrix. The discriminant weights estimated by using the analysis sample are multiplied by the values of the predictor variables in the holdout sample to generate discriminant scores for the cases in the holdout sample. The cases are then assigned to groups based on their discriminant scores and an appropriate decision rule. The percentage of cases correctly classified is determined and compared with the rate that would be expected by chance classification.

Two broad approaches are available for estimating the coefficients. The direct method involves estimating the discriminant function so that all the predictors are included simultaneously. An alternative is the stepwise method in which the predictor variables are entered sequentially, based on their ability to discriminate among groups.

In multiple discriminant analysis, if there are G groups and k predictors, it is possible to estimate up to the smaller of $G - 1$, or k, discriminant functions. The first function has the highest ratio of between-group to within-group sums of squares; the second function, uncorrelated with the first, has the second highest ratio; and so on.

Questions

1. What are the objectives of discriminant analysis?

2. Describe four examples of the application of discriminant analysis.

3. What is the main distinction between two-group and multiple discriminant analysis?

4. Describe the relationship of discriminant analysis to regression and ANOVA.

5. What are the steps involved in conducting discriminant analysis?

6. How should the total sample be split for estimation and validation purposes?

7. What is Wilks' λ? For what purpose is it used?

8. Define discriminant scores.

9. Explain what is meant by an eigenvalue.

10. What is a classification matrix?

11. Explain the concept of structure correlations.

12. How is the statistical significance of discriminant analysis determined?

13. Describe a common procedure for determining the validity of discriminant analysis.

14. When the groups are of equal size, how is the accuracy of chance classification determined?

15. How does the stepwise discriminant procedure differ from the direct method?

Appendix: Estimation of discriminant function coefficients

Suppose that there are G groups, $i = 1, 2, 3, \ldots, G$, each containing n_i observations on K independent variables, $X_1, X_2, \ldots, X_k$. The following notations are used:

N = total sample size $= \sum_{i=1}^{G} n_i$

W_i = matrix of mean corrected sum of squares and cross-products for the ith group

W = matrix of pooled within-groups mean corrected sum of squares and cross-products

B = matrix of between-groups mean corrected sum of squares and cross-products

T = matrix of total mean corrected sum of squares and cross-products for all the N observations

$= W + B$

$\bar{X}_i$ = vector of means of observations in the ith group

$\bar{X}$ = vector of grand means for all the N observations

λ = ratio of between-groups to within-group sums of squares

b = vector of discriminant coefficients or weights

Then,

$$T = \sum_{i=1}^{G} \sum_{j=1}^{n_i} (X_{ij} - \bar{X})(X_{ij} - \bar{X})'$$

$$W_i = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(X_{ij} - \bar{X}_i)'$$

$$W = W_1 + W_2 + W_3 + \ldots + W_G$$

$$B = T - W$$

Define the linear composite $D = b_1'X$. Then, with reference to D, the between-groups and within-groups sums of squares are $b_1'Bb$ and $b_1'Wb$, respectively. To maximally discriminate the groups, the discriminant functions are estimated to maximise the between-group variability. The coefficients b are calculated to maximise λ, by solving

$$\text{Max } \lambda = \frac{b'Bb}{b'Wb}$$

Taking the partial derivative with respect to λ and setting it equal to zero, with some simplification, yields:

$$(B - \lambda W)b = 0$$

To solve for b, it is more convenient to premultiply by W^{-1} and solve the following characteristic equation:

$$(W^{-1}B - \lambda I)b = 0$$

The maximum value of λ is the largest eigenvalue of the matrix $W^{-1}B$, and b is the associated eigenvector. The elements of b are the discriminant coefficients, or weights, associated with the first discriminant function. In general, it is possible to estimate up to the smaller of $G - 1$ or k discriminant functions, each with its associated eigenvalue. The discriminant functions are estimated sequentially. In other words, the first discriminant function exhausts most of the between-group variability, the second function maximises the between-group variation that was not explained by the first one, and so on.

Notes

1 Lichtenstein, D.R., Burton, S. and Netemeyer, R.G., 'An examination of deal proneness across sales promotion types: a consumer segmentation perspective', *Journal of Retailing* 73(2) (Summer 1997), 283–97; Jolson, M.A., Wiener, J.L. and Rosecky, R.B., 'Correlates of rebate proneness', *Journal of Advertising Research* (February–March 1987), 33–43.

2 A detailed discussion of discriminant analysis may be found in Tacq, J., *Multivariate Analysis Techniques in Social Science Research* (Thousand Oaks, CA: Sage, 1996); Lachenbruch, P.A., *Discriminant Analysis* (New York: Hafner Press, 1975). For an application, see Deal, K., 'Determining success criteria for financial products: a comparative analysis of CART, logit and factor/discriminant analysis', *Service Industries Journal* 17(3) (July 1997), 489–506.

3 Johnson, R.A. and Wichern, D.W., *Applied Multivariate Statistical Analysis*, 4th edn (Upper Saddle River, NJ: Prentice Hall, 1998); Klecka, W.A., *Discriminant Analysis* (Beverly Hills, CA: Sage, 1980). See also Sinclair, S.A. and Stalling, E.C., 'How to identify differences between market segments with attribute analysis', *Industrial Marketing Management* 19 (February 1990), 31–40.

4 For an application, see Khan, Z., Chawla, S.K. and Cianciolo, T.A., 'Multiple discriminant analysis: tool for effective marketing of computer information systems to small business clients', *Journal of Professional Systems Marketing* 12(2) (1995), 153–62; Sager, J.K. and Menon, A., 'The role of behavioural intentions in turnover of salespeople', *Journal of Business Research* 29 (March 1994), 179–88; Kijewski, V., Yoon, E. and Young, G., 'How exhibitors select trade shows', *Industrial Marketing Management* 22 (November 1993), 287–98.

5 Mitchell, V.-W., 'How to identify psychographic segments: Part 2', *Marketing Intelligence and Planning* 12(7) (1994), 11–16; Crask, M.R. and Perrault Jr, W.D., 'Validation of discriminant analysis in marketing research', *Journal of Marketing Research* 14 (February 1977) 60–8.

6 Strictly speaking, before testing for the equality of group means, the equality of group covariance matrices should be tested. Box's M test can be used for this purpose. If the equality of group covariance means is rejected, the results of discriminant analysis should be interpreted with caution. In this case, the power of the test for the equality of group means decreases.

7 See Fok, L., Angelidis, J.P., Ibrahim, N.A. and Fok, W.M., 'The utilization and interpretation of multivariate statistical techniques in strategic management', *International Journal of Management* 12(4) (December 1995), 468–81; Morrison, D.G., 'On the interpretation of discriminant analysis', *Journal of Marketing Research* 6 (May 1969), 156–63. For the use of other techniques in conjunction with discriminant analysis to aid interpretation, see Dant, R.P., Lumpkin, J.R. and Bush, R.P., 'Private physicians or walk-in clinics: do the patients differ?', *Journal of Health Care Marketing* (June 1990) 23–35.

8 Hawes, J.M., Rao, C.P. and Baker, T.L., 'Retail salesperson attributes and the role of dependability in the selection of durable goods', *Journal of Personal Selling and Sales Management* 13(4) (Fall 1993), 61–71; Fern, E.E., Avila, R.A. and Grewal, D., 'Salesforce turnover: those who left and those who stayed', *Industrial Marketing Management* (1989), 1–9.

9 For the validation of discriminant analysis, see Bush, R.P., Ortinau, D.J. and Bush, A.J., 'Personal value structures and AIDS prevention', *Journal of Health Care Marketing* 14 (Spring 1994), 12–20.

10 Hair Jr, J.E., Anderson, R.E., Tatham, R.L. and Black, W.C., *Multivariate Data Analysis with Readings*, 5th edn (Upper Saddle River, NJ: Prentice Hall, 1999).

11 Mitchell, V.W., 'How to identify psychographic segments: Part 2', *Marketing Intelligence and Planning* 12(7) (1994), 11–16; Albaum, G. and Baker, K., 'The sampling problem in validation of multiple discriminant analysis', *Journal of the Market Research Society* 18 (July 1976).

12 Rahtz, D.R., Sirgy, M.J. and Kosenko, R., 'Using demographics and psychographic dimensions to discriminate between mature heavy and light television users: an exploratory analysis', in Bahn, K.D. (ed.), *Developments in Marketing Science*, Vol. 11 (Blacksburg, VA: Academy of Marketing Science, 1988), 2–7.

13 For advanced discussion of multiple discriminant analysis, see Johnson, R.A. and Wichern, D.W., *Applied Multivariate Statistical Analysis*, 4th edn (Upper Saddle River, NJ: Prentice Hall, 1998). For an application, see Dant, R.P. and Schul, P.L., 'Conflict resolution processes in contractual channels of distribution', *Journal of Marketing* 56 (January 1992), 38–54.

14 Loucopoulos, C. and Pavur, R.M., 'Computational characteristics of a new mathematical programming model for the three-group discriminant problem', *Computers and Operations Research* 24(2) (February 1997), 179–91. For an application of multiple discriminant analysis, see O'Connor, S.J., Shewchuk, R.M. and Camey, L.W., 'The great gap', *Journal of Health Care Marketing* 14 (Summer 1994), 32–9.

15 Dansky, K.H. and Brannon, D., 'Discriminant analysis: a technique for adding value to patient satisfaction surveys', *Hospital and Health Services Administration* 41(4) (Winter 1996), 503–13; Lim, J.S. and Zallocco, R., 'Determinant attributes in formulation of attitudes toward four health care systems', *Journal of Health Care Marketing* (June 1988), 25–30.

16 Johnson, R.A. and Wichern, D.A., *Applied Multivariate Statistical Analysis*, 4th edn (Upper Saddle River, NJ: Prentice Hall, 1998); Hair Jr, J.E., Anderson, R.E., Tatham, R.L. and Black, W.C., *Multivariate Data Analysis with Readings*, 5th edn (Englewood Cliffs, NJ: Prentice Hall, 1999), 178–255.

17 Whitelock, J., Roberts, C. and Blakeley, J., 'The reality of the eurobrand: an empirical analysis', *Journal of International Marketing* 3(3) (1995) 77–95; Klopp, C. and Sterlicchi, J., 'Customer satisfaction just catching on in Europe', *Marketing News* (28 May 1990), 5.

Factor analysis

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Objectives

After reading this chapter, you should be able to:

1 describe the concept of factor analysis and explain how it is different from analysis of variance, multiple regression and discriminant analysis;

2 discuss the procedure for conducting factor analysis, including problem formulation, construction of the correlation matrix, selection of an appropriate method, determination of the number of factors, rotation and interpretation of factors;

3 understand the distinction between principal component factor analysis and common factor analysis methods;

4 explain the selection of surrogate variables and their application with emphasis on their use in subsequent analysis;

5 describe the procedure for determining the fit of a factor analysis model using the observed and the reproduced correlations.

Factor analysis allows an examination of the potential interrelationships among a number of variables and the evaluation of the underlying reasons for these relationships.

Overview

In analysis of variance (Chapter 19), regression (Chapter 20) and discriminant analysis (Chapter 21), one of the variables is clearly identified as the dependent variable. We now turn to a procedure, factor analysis, in which variables are not classified as independent or dependent. Instead, the whole set of interdependent relationships among variables is examined. This chapter discusses the basic concept of factor analysis and gives an exposition of the factor model. We describe the steps in factor analysis and illustrate them in the context of principal components analysis. Next, we present an application of common factor analysis. To begin, we provide some examples to illustrate the usefulness of factor analysis.

Factor analysis[1]

In the GlobalCash Project, the respondents' ratings of 11 service quality statements were factor analysed to determine the underlying service quality factors. Four factors emerged: close support, speed of activities, coping with errors and matching efficiency expectations. These factors, along with individual country characteristics, were used to profile market segments formed as a result of clustering. ■

Personal alarms[2]

In a study of personal alarms, women were asked to rate eight personal alarms using the following 15 statements:

1 Feels comfortable in the hand
2 Could be easily kept in the pocket
3 Would fit easily into a handbag
4 Could be easily worn on the person
5 Could be carried to be very handy when needed
6 Could be set off almost as a reflex action
7 Would be difficult for an attacker to take it off me
8 Could keep a very firm grip on it if attacked
9 An attacker might be frightened that I might attack him with it
10 Would be difficult for an attacker to switch off
11 Solidly built
12 Would be difficult to break
13 Looks as if it would give off a very loud noise
14 An attacker might have second thoughts about attacking me if he saw me with it
15 I would be embarrassed to carry it around with me

The question was 'could these 15 variables be reduced to a smaller number of derived variables, known as factors, in such a way that too much information was not lost?' Factor analysis enabled these 15 variables to be reduced to four underlying dimensions or factors that women used to evaluate the alarms. Factor 1 seemed to measure a dimension of '*size*', on a continuum of small to large. Factor 2 tapped into aspects of the '*appearance*' of a personal alarm. Factor 3 revealed '*robustness*' characteristics, with factor 4 related to '*hand feel*'. ■

Basic concept

Factor analysis
A class of procedures primarily used for data reduction and summarisation.

Factor
An underlying dimension that explains the correlations among a set of variables.

Factor analysis is a general name denoting a class of procedures primarily used for data reduction and summarisation. In marketing research, there may be a large number of variables, most of which are correlated and which must be reduced to a manageable level. Relationships among sets of many interrelated variables are examined and represented in terms of a few underlying factors. For example, bank image may be measured by asking respondents to evaluate banks on a series of items on a semantic differential scale or a Likert scale. These item evaluations may then be analysed to determine the **factors** underlying bank image.

In analysis of variance, multiple regression and discriminant analysis, one variable is considered the dependent or criterion variable, and the others are considered independent or predictor variables. But no such distinction is made in factor analysis. Rather, factor analysis is an **interdependence technique** in that an entire set of interdependent relationships is examined.[3]

Factor analysis is used in the following circumstances:

Interdependence technique
A multivariate statistical technique in which the whole set of interdependent relationships is examined.

1 To identify underlying dimensions, or factors, that explain the correlations among a set of variables. For example, a set of lifestyle statements may be used to measure the psychographic profiles of consumers. These statements may then be factor analysed to identify the underlying psychographic factors.[4]

2 To identify a new, smaller, set of uncorrelated variables to replace the original set of correlated variables in subsequent multivariate analysis (regression or discriminant analysis). For example, the psychographic factors identified may be used as independent variables in explaining the differences between loyal and non-loyal consumers.

3 To identify a smaller set of salient variables from a larger set for use in subsequent multivariate analysis. For example, a few of the original lifestyle statements that correlate highly with the identified factors may be used as independent variables to explain the differences between the loyal and non-loyal users.

Factor analysis has numerous applications in marketing research. For example:

■ Factor analysis can be used in market segmentation for identifying the underlying variables on which to group the customers. New car buyers might be grouped based on the relative emphasis they place on economy, convenience, performance, comfort and luxury. This might result in five segments: economy seekers, convenience seekers, performance seekers, comfort seekers and luxury seekers.

■ In product research, factor analysis can be employed to determine the brand attributes that influence consumer choice. Toothpaste brands might be evaluated in terms of protection against cavities, whiteness of teeth, taste, fresh breath and price.

■ In advertising studies, factor analysis can be used to understand the media consumption habits of the target market. The users of frozen foods may be heavy viewers of satellite TV, see a lot of videos, and listen to country music.

■ In pricing studies, factor analysis can be used to identify the characteristics of price-sensitive consumers. For example, these consumers might be methodical, economy minded and home centred.

Factor analysis model

Mathematically, factor analysis is somewhat similar to multiple regression analysis in that each variable is expressed as a linear combination of underlying factors. The amount of variance a variable shares with all other variables included in the analysis is referred to as communality. The covariation among the variables is described in terms of a small number of common factors plus a unique factor for each variable. These factors are not overtly observed. If the variables are standardised, the factor model may be represented as

$$X_i = A_{i1}F_1 + A_{i2}F_2 + A_{i3}F_3 + \ldots + A_{im}F_m + V_iU_i$$

where X_i = ith standardised variable
A_{ij} = standardised multiple regression coefficient of variable i on common factor j
F = common factor
V_i = standardised regression coefficient of variable i on unique factor i
U_i = the unique factor for variable i
m = number of common factors.

The unique factors are correlated with each other and with the common factors.[5] The common factors themselves can be expressed as linear combinations of the observed variables

$$F_i = W_{i1}X_1 + W_{i2}X_2 + W_{i3}X_3 + \ldots + W_{ik}X_k$$

where
F_i = estimate of ith factor
W_i = weight or factor score coefficient
k = number of variables.

It is possible to select weights or factor score coefficients so that the first factor explains the largest portion of the total variance. Then a second set of weights can be selected so that the second factor accounts for most of the residual variance, subject to being uncorrelated with the first factor. This same principle could be applied to selecting additional weights for the additional factors. Thus, the factors can be estimated so that their factor scores, unlike the values of the original variables, are not correlated. Furthermore, the first factor accounts for the highest variance in the data, the second factor the second highest, and so on. A technical treatment of the factor analysis model is presented in Appendix 22A.

The key statistics associated with factor analysis are as follows:

Bartlett's test of sphericity. Bartlett's test of sphericity is a test statistic used to examine the hypothesis that the variables are uncorrelated in the population. In other words, the population correlation matrix is an identity matrix; each variable correlates perfectly with itself ($r = 1$) but has no correlation with the other variables ($r = 0$).

Communality. Communality is the amount of variance a variable shares with all the other variables being considered. This is also the proportion of variance explained by the common factors.

Correlation matrix. A correlation matrix is a lower triangle matrix showing the simple correlations, r, between all possible pairs of variables included in the analysis. The diagonal elements, which are all 1, are usually omitted.

Eigenvalue. The eigenvalue represents the total variance explained by each factor.

Factor loadings. Factor loadings are simple correlations between the variables and the factors.

Factor loading plot. A factor loading plot is a plot of the original variables using the factor loadings as coordinates.

Factor matrix. A factor matrix contains the factor loadings of all the variables on all the factors extracted.

Factor scores. Factor scores are composite scores estimated for each respondent on the derived factors.

Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is an index used to examine the appropriateness of factor analysis. High values (between 0.5 and 1.0) indicate that factor analysis is appropriate. Values below 0.5 imply that factor analysis may not be appropriate.

Percentage of variance. The percentage of the total variance attributed to each factor.

Residuals. Residuals are the differences between the observed correlations, as given in the input correlation matrix, and the reproduced correlations, as estimated from the factor matrix.

Scree plot. A scree plot is a plot of the eigenvalues against the number of factors in order of extraction.

We describe the uses of these statistics in the next section, in the context of the procedure for conducting factor analysis.

Conducting factor analysis

The steps involved in conducting factor analysis are illustrated in Figure 22.1. The first step is to define the factor analysis problem and identify the variables to be factor analysed. Then a correlation matrix of these variables is constructed and a method of factor analysis is selected. The researcher decides on the number of factors to be extracted and the method of rotation. Next, the rotated factors should be interpreted. Depending on the objectives, the factor scores may be calculated, or surrogate variables selected, to represent the factors in subsequent multivariate analysis. Finally, the fit of the factor analysis model is determined. We discuss these steps in more detail below.[6]

Formulate the problem

Formulating the problem includes several tasks. First, the objectives of factor analysis should be identified. The variables to be included in the factor analysis should be specified based on past research (quantitative or qualitative), theory, and judgement of the researcher. It is important that the variables be appropriately measured on an interval or ratio scale. An appropriate sample size should be used. As a rough guideline, there should be at least four or five times as many observations (sample size) as there are variables.[7] In many marketing research situations, the sample size is small, and this ratio is considerably lower. In these cases, the results should be interpreted cautiously.

To illustrate factor analysis, suppose that the researcher wants to determine the underlying benefits consumers seek from the purchase of a toothpaste. A sample of 237 respondents was interviewed using street interviewing. The respondents were asked to indicate their degree of agreement with the following statements using a seven-point scale (1 = strongly disagree, 7 = strongly agree):
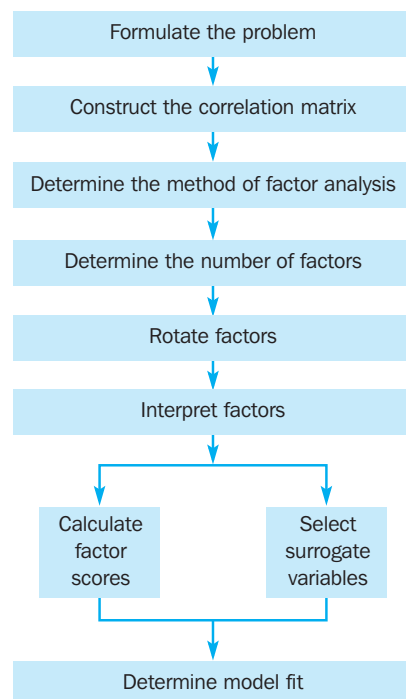
Figure 22.1
Conducting factor analysis

V_1 It is important to buy a toothpaste that prevents cavities.

V_2 I like a toothpaste that gives shiny teeth.

V_3 A toothpaste should strengthen your gums.

V_4 I prefer a toothpaste that freshens breath.

V_5 Prevention of tooth decay should be an important benefit offered by a toothpaste.

V_6 The most important consideration in buying a toothpaste is attractive teeth.

The data obtained are given in Table 22.1. A correlation matrix was constructed based on these ratings data.

Table 22.1 Toothpaste attribute ratings

Respondent number	V_1	V_2	V_3	V_4	V_5	V_6
1	7.00	3.00	6.00	4.00	2.00	4.00
2	1.00	3.00	2.00	4.00	5.00	4.00
3	6.00	2.00	7.00	4.00	1.00	3.00
4	4.00	5.00	4.00	6.00	2.00	5.00
5	1.00	2.00	2.00	3.00	6.00	2.00
6	6.00	3.00	6.00	4.00	2.00	4.00
7	5.00	3.00	6.00	3.00	4.00	3.00
8	6.00	4.00	7.00	4.00	1.00	4.00
9	3.00	4.00	2.00	3.00	6.00	3.00
10	2.00	6.00	2.00	6.00	7.00	6.00
11	6.00	4.00	7.00	3.00	2.00	3.00
12	2.00	3.00	1.00	4.00	5.00	4.00
13	7.00	2.00	6.00	4.00	1.00	3.00
14	4.00	6.00	4.00	5.00	3.00	6.00
15	1.00	3.00	2.00	2.00	6.00	4.00
16	6.00	4.00	6.00	3.00	3.00	4.00
17	5.00	3.00	6.00	3.00	3.00	4.00
18	7.00	3.00	7.00	4.00	1.00	4.00
19	2.00	4.00	3.00	3.00	6.00	3.00
20	3.00	5.00	3.00	6.00	4.00	6.00
21	1.00	3.00	2.00	3.00	5.00	3.00
22	5.00	4.00	5.00	4.00	2.00	4.00
23	2.00	2.00	1.00	5.00	4.00	4.00
24	4.00	6.00	4.00	6.00	4.00	7.00
25	6.00	5.00	4.00	2.00	1.00	4.00
26	3.00	5.00	4.00	6.00	4.00	7.00
27	4.00	4.00	7.00	2.00	2.00	5.00
28	3.00	7.00	2.00	6.00	4.00	3.00
29	4.00	6.00	3.00	7.00	2.00	7.00
30	2.00	3.00	2.00	4.00	7.00	2.00

Construct the correlation matrix

The analytical process is based on a matrix of correlations between the variables. Valuable insights can be gained from an examination of this matrix. For factor analysis to be meaningful, the variables should be correlated. In practice, this is usually the case. If the correlations between all the variables are small, factor analysis may not be appropriate. We would also expect that variables that are highly correlated with each other would also highly correlate with the same factor or factors.

Formal statistics are available for testing the appropriateness of the factor model. Bartlett's test of sphericity can be used to test the null hypothesis that the variables are uncorrelated in the population; in other words, the population correlation matrix is an identity matrix. In an identity matrix, all the diagonal terms are 1, and all off-diagonal terms are 0. The test statistic for sphericity is based on a chi-square transformation of the determinant of the correlation matrix. A large value of the test statistic will favour the rejection of the null hypothesis. If this hypothesis cannot be rejected, then the appropriateness of factor analysis should be questioned. Another useful statistic is the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. This index compares the magnitudes of the observed correlation coefficients with the magnitudes of the partial correlation coefficients. Small values of the KMO statistic indicate that the correlations between pairs of variables cannot be explained by other variables and that factor analysis may not be appropriate.

The correlation matrix, constructed from the data obtained to understand toothpaste benefits, is shown in Table 22.2. There are relatively high correlations among V_1 (prevention of cavities), V_3 (strong gums) and V_5 (prevention of tooth decay). We would expect these variables to correlate with the same set of factors. Likewise, there are relatively high correlations among V_2 (shiny teeth), V_4 (fresh breath) and V_6 (attractive teeth). These variables may also be expected to correlate with the same factors.[8]

Table 22.2 Correlation matrix

Variables	V_1	V_2	V_3	V_4	V_5	V_6
V_1	1.00					
V_2	−0.053	1.00				
V_3	0.873	−0.155	1.00			
V_4	−0.086	0.572	−0.248	1.00		
V_5	−0.858	0.020	−0.778	−0.007	1.00	
V_6	0.004	0.640	−0.018	0.640	−0.136	1.00

The results of factor analysis are given in Table 22.3. The null hypothesis, that the population correlation matrix is an identity matrix, is rejected by Bartlett's test of sphericity. The approximate chi-square statistic is 111.314 with 15 degrees of freedom which is significant at the 0.05 level. The value of the KMO statistic (0.660) is also large (> 0.5). Thus factor analysis may be considered an appropriate technique for analysing the correlation matrix of Table 22.2.

Determine the method of factor analysis

Once it has been determined that factor analysis is an appropriate technique for analysing the data, an appropriate method must be selected. The approach used to derive the weights or factor score coefficients differentiates the various methods of factor analysis. The two basic approaches are principal components analysis and common factor analysis. In principal components analysis, the total variance in the

Common factor analysis
An approach to factor analysis that estimates the factors based only on the common variance. Also called principal axis factoring.

data is considered. The diagonal of the correlation matrix consists of unities, and full variance is brought into the factor matrix. **Principal components analysis** is recommended when the primary concern is to determine the minimum number of factors that will account for maximum variance in the data for use in subsequent multivariate analysis. The factors are called *principal components.*

In **common factor analysis**, the factors are estimated based only on the common variance. Communalities are inserted in the diagonal of the correlation matrix. This method is appropriate when the primary concern is to identify the underlying dimensions and the common variance is of interest. This method is also known as *principal axis factoring.*

Other approaches for estimating the common factors are also available. These include the methods of unweighted least squares, generalised least squares, maximum likelihood, alpha method and image factoring. These methods are complex and are not recommended for inexperienced users.[9]

Table 22.3 shows the application of principal components analysis to the toothpaste example.

Table 22.3 Results of principal components analysis

Bartlett test of sphericity
Approximate chi-square = 111.314, *df* = 15, significance = 0.00000
Kaiser-Meyer-Olkin measure of sampling adequacy = 0.660

Communalities

Variable	Initial	Extraction
V_1	1.000	0.926
V_2	1.000	0.723
V_3	1.000	0.894
V_4	1.000	0.739
V_5	1.000	0.878
V_6	1.000	0.790

Initial eigenvalues

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	2.731	45.520	45.520
2	2.218	36.969	82.488
3	0.442	7.360	89.848
4	0.341	5.688	95.536
5	0.183	3.044	98.580
6	0.085	1.420	100.000

Extraction sums of squared loadings

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	2.731	45.520	45.520
2	2.218	36.969	82.488

Table 22.3 Continued

Factor matrix

	Factor 1	Factor 2
V_1	0.928	0.253
V_2	−0.301	0.795
V_3	0.936	0.131
V_4	−0.342	0.789
V_5	−0.869	−0.351
V_6	−0.177	0.871

Rotation sums of squared loadings

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	2.688	44.802	44.802
2	2.261	37.687	82.488

Rotated factor matrix

	Factor 1	Factor 2
V_1	0.962	−0.027
V_2	−0.057	0.848
V_3	0.934	−0.146
V_4	−0.098	0.854
V_5	−0.933	−0.084
V_6	0.083	0.885

Factor score coefficient matrix

	Factor 1	Factor 2
V_1	0.358	0.011
V_2	−0.001	0.375
V_3	0.345	−0.043
V_4	−0.017	0.377
V_5	−0.350	−0.059
V_6	0.052	0.395

Reproduced correlation matrix

Variables	V_1	V_2	V_3	V_4	V_5	V_6
V_1	0.926*	0.024	−0.029	0.031	0.038	−0.053
V_2	−0.078	0.723*	0.022	−0.158	0.038	−0.105
V_3	0.902	−0.177	0.894*	−0.031	0.081	0.033
V_4	−0.117	0.730	−0.217	0.739*	−0.027	−0.107
V_5	−0.895	−0.018	0.859	0.020	0.878*	0.016
V_6	0.057	−0.746	−0.051	0.748	−0.152	0.790*

The lower left triangle contains the reproduced correlation matrix; the diagonal, the communalities; and the upper right triangle, the residuals between the observed correlations and the reproduced correlations.

Under initial statistics, it can be seen that the communality for each variable, V_1 to V_6, is 1.0 as unities were inserted in the diagonal of the correlation matrix. The table labelled 'Initial eigenvalues' gives the eigenvalues. The eigenvalues for the factors are, as expected, in decreasing order of magnitude as we go from factor 1 to factor 6. The eigenvalue for a factor indicates the total variance attributed to that factor. The total variance accounted for by all the six factors is 6.00, which is equal to the number of variables. Factor 1 accounts for a variance of 2.731, which is (2.731/6) or 45.52% of the total variance. Likewise, the second factor accounts for (2.218/6) or 36.97% of the total variance, and the first two factors combined account for 82.49% of the total variance. Several considerations are involved in determining the number of factors that should be used in the analysis.

Determine the number of factors

It is possible to compute as many principal components as there are variables, but in doing so, no parsimony is gained, i.e. we would not have summarised the information nor revealed any underlying structure. To summarise the information contained in the original variables, a smaller number of factors should be extracted. The question is, how many? Several procedures have been suggested for determining the number of factors. These included *a priori* determination and approaches based on eigenvalues, scree plot, percentage of variance accounted for, split-half reliability, and significance tests.

A priori **determination**. Sometimes, because of prior knowledge, the researcher knows how many factors to expect and thus can specify the number of factors to be extracted beforehand. The extraction of factors ceases when the desired number of factors have been extracted. Most computer programs allow the user to specify the number of factors, allowing for an easy implementation of this approach.

Determination based on eigenvalues. In this approach, only factors with eigenvalues greater than 1.0 are retained; the other factors are not included in the model. An eigenvalue represents the amount of variance associated with the factor. Hence, only factors with a variance greater than 1.0 are included. Factors with variance less than 1.0 are no better than a single variable because, due to standardisation, each variable has a variance of 1.0. If the number of variables is less than 20, this approach will result in a conservative number of factors.

Determination based on scree plot. A scree plot is a plot of the eigenvalues against the number of factors in order of extraction. The shape of the plot is used to determine the number of factors. Typically, the plot has a distinct break between the steep slope of factors, with large eigenvalues and a gradual trailing off associated with the rest of the factors. This gradual trailing off is referred to as the scree. Experimental evidence indicates that the point at which the scree begins denotes the true number of factors. Generally, the number of factors determined by a scree plot will be one or a few more than that determined by the eigenvalue criterion.

Determination based on percentage of variance. In this approach, the number of factors extracted is determined so that the cumulative percentage of variance extracted by the factors reaches a satisfactory level. What level of variance is satisfactory depends upon the problem. It is recommended that the factors extracted should account for at least 60% of the variance.

Determination based on split-half reliability. The sample is split in half, and factor analysis is performed on each half. Only factors with high correspondence of factor loadings across the two sub-samples are retained.

Determination based on significance tests. It is possible to determine the statistical significance of the separate eigenvalues and retain only those factors that are statistically significant. A drawback is that with large samples (size greater than 200) many factors are likely to be statistically significant, although from a practical viewpoint many of these account for only a small proportion of the total variance.

In Table 22.3, we see that the eigenvalue greater than 1.0 (default option) results in two factors being extracted. Our *a priori* knowledge tells us that toothpaste is bought for two major reasons. The scree plot associated with this analysis is given in Figure 22.2. From the scree plot, a distinct break occurs at three factors. Finally, from the cumulative percentage of variance accounted for, we see that the first two factors account for 82.49% of the variance and that the gain achieved in going to three factors is marginal. Furthermore, split-half reliability also indicates that two factors are appropriate. Thus, two factors appear to be reasonable in this situation.

The second column under the 'Communalities' heading in Table 22.3 gives relevant information after the desired number of factors have been extracted. The communalities for the variances under 'Extraction' are different from those under 'Initial' because all of the variances associated with the variables are not explained unless all the factors are retained. The 'Extraction sums of squared loadings' table gives the variances associated with the factors that are retained. Note that these are the same as those under 'Initial eigenvalues'. This is always the case in principal components analysis. The percentage variance accounted for by a factor is determined by dividing the associated eigenvalue by the total number of factors (or variables) and multiplying by 100. Thus, the first factor accounts for $(2.731/6) \times 100$ or 45.52% of the variance of the six variables. Likewise, the second factor accounts for $(2.218/6) \times 100$ or 36.967% of the variance. Interpretation of the solution is often enhanced by a rotation of the factors.

Rotate factors

An important output from factor analysis is the factor matrix, also called the *factor pattern matrix*. The factor matrix contains the coefficients used to express the standardised variables in terms of the factors. These coefficients, the factor loadings, represent the correlations between the factors and the variables. A coefficient with a large absolute value indicates that the factor and the variable are closely related. The coefficients of the factor matrix can be used to interpret the factors.

Although the initial or unrotated factor matrix indicates the relationship between the factors and individual variables, it seldom results in factors that can be interpreted, because the factors are correlated with many variables. For example, in Table
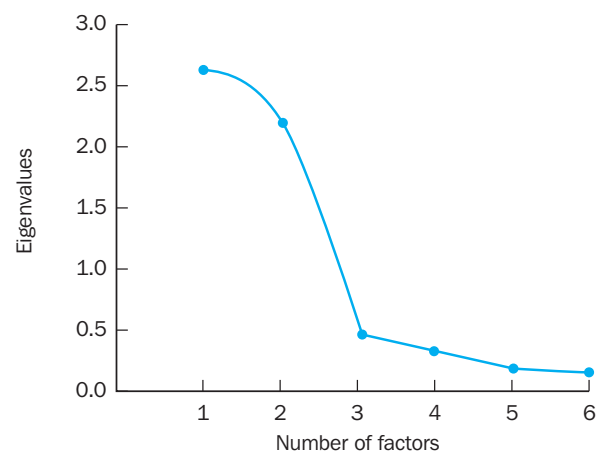
Figure 22.2
Scree plot

Rotating factors allows facets of the dataset to be viewed from different perspectives.

22.3, factor 1 is at least somewhat correlated with five of the six variables (absolute value of factor loading greater than 0.3). How should this factor be interpreted? In such a complex matrix, it is difficult to interpret the factors. Therefore, through rotation, the factor matrix is transformed into a simpler one that is easier to interpret.

In rotating the factors, we would like each factor to have non-zero, or significant, loadings or coefficients for only some of the variables. Likewise, we would like each variable to have non-zero or significant loadings with only a few factors, and if possible with only one. If several factors have high loadings with the same variable, it is difficult to interpret them. Rotation does not affect the communalities and the percentage of total variance explained. The percentage of variance accounted for by each factor does change, however. This is seen in Table 22.3. The variance explained by the individual factors is redistributed by rotation. Hence, different methods of rotation may result in the identification of different factors.

The rotation is called **orthogonal rotation** if the axes are maintained at right angles. The most commonly used method for rotation is the **varimax procedure**. This is an orthogonal method of rotation that minimises the number of variables with high loadings on a factor, thereby enhancing the interpretability of the factors.[10] Orthogonal rotation results in factors that are uncorrelated. The rotation is called **oblique rotation** when the axes are not maintained at right angles, and the factors are correlated. Sometimes, allowing for correlations among factors can simplify the factor pattern matrix. Oblique rotation should be used when factors in the population are likely to be strongly correlated.

In Table 22.3, by comparing the varimax rotated factor matrix with the unrotated matrix (entitled factor matrix), we can see how rotation achieves simplicity and enhances interpretability. Whereas five variables correlated with factor 1 in the unrotated matrix, only variables V_1, V_3 and V_5 correlate highly with factor 1 after rotation. The remaining variables – V_2, V_4 and V_6 – correlate highly with factor 2. Furthermore, no variable correlates highly with both the factors. The rotated factor matrix forms the basis for interpretation of the factors.

Orthogonal rotation
Rotation of factors in which the axes are maintained at right angles.

Varimax procedure
An orthogonal method of factor rotation that minimises the number of variables with high loadings on a factor, thereby enhancing the interpretability of the factors.

Oblique rotation
Rotation of factors when the axes are not maintained at right angles.

Interpret factors

Interpretation is facilitated by identifying the variables that have large loadings on the same factor. That factor can then be interpreted in terms of the variables that load high on it. Another useful aid in interpretation is to plot the variables, using the factor loadings as coordinates. Variables at the end of an axis are those that have high loadings on only that factor and hence describe the factor. Variables near the origin have small loadings on both the factors. Variables that are not near any of the axes are related to both the factors. If a factor cannot be clearly defined in terms of the original variables, it should be labelled as an undefined or a general factor.

In the rotated factor matrix of Table 22.3, factor 1 has high coefficients for variables V_1 (prevention of cavities) and V_3 (strong gums), and a negative coefficient for V_5 (prevention of tooth decay is not important). Therefore, this factor may be labelled a health benefit factor. Note that a negative coefficient for a negative variable (V_5) leads to a positive interpretation that prevention of tooth decay is important. Factor 2 is highly related with variables V_2 (shiny teeth), V_4 (fresh breath) and V_6 (attractive teeth). Thus factor 2 may be labelled a social benefit factor. A plot of the factor loadings, given in Figure 22.3, confirms this interpretation. Variables V_1, V_3, and V_5 (denoted 1, 3, and 5, respectively) are at the end of the horizontal axis (factor 1), with V_5 at the end opposite to V_1 and V_3, whereas variables V_2, V_4 and V_6 (denoted 2, 4 and 6) are at the end of the vertical axis (factor 2). One could summarise the data by stating that consumers appear to seek two major kinds of benefits from a toothpaste: health benefits and social benefits.

Calculate factor scores

Factor scores
Composite scores estimated for each respondent on the derived factors.

Following interpretation, **factor scores** can be calculated, if necessary. Factor analysis has its own stand-alone value. If the goal of factor analysis is to reduce the original set of variables to a smaller set of composite variables (factors) for use in subsequent multivariate analysis, however, it is useful to compute factor scores for each respondent. A factor is simply a linear combination of the original variables. The factor scores for the ith factor may be estimated as follows:

$$F_i = W_{i1}X_1 + W_{i2}X_2 + W_{i3}X_3 + \ldots + W_{ik}X_k$$

where the symbols are as defined earlier in the chapter.

The weights or factor score coefficients used to combine the standardised variables are obtained from the factor score coefficient matrix. Most computer programs allow you to request factor scores. Only in the case of principal components analysis is it possible to compute exact factor scores. Moreover, in principal components analysis, these scores are uncorrelated. In common factor analysis, estimates of these scores are obtained, and there is no guarantee that the factors will be uncorrelated with each other. Factor scores can be used instead of the original variables in subsequent multi-
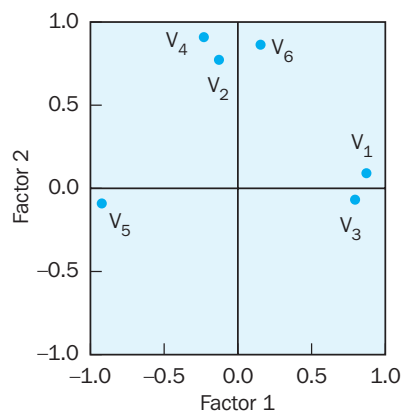
Figure 22.3
Factor loading plot

variate analysis. For example, using the factor score coefficient matrix in Table 22.3, one could compute two factor scores for each respondent. The standardised variable values would be multiplied by the corresponding factor score coefficients to obtain the factor scores.

Select surrogate variables

Surrogate variables
A subset of original variables selected for use in subsequent analysis.

Sometimes, instead of computing factor scores, the researcher wishes to select surrogate variables. Selection of substitute or **surrogate variables** involves singling out some of the original variables for use in subsequent analysis. This allows the researcher to conduct subsequent analysis and to interpret the results in terms of original variables rather than factor scores. By examining the factor matrix, one could select for each factor the variables rather than factor scores. By examining the factor matrix, one could select for each factor the variable with the highest loading on that factor. That variable could then be used as a surrogate variable for the associated factor. This process works well if one factor loading for a variable is clearly higher than all other factor loadings. The choice is not as easy, however, if two or more variables have similarly high loadings. In such a case, the choice between these variables should be based on theoretical and measurement considerations. For example, theory may suggest that a variable with a slightly lower loading is more important than one with a slightly higher loading. Likewise, if a variable has a slightly lower loading but has been measured more precisely, it should be selected as the surrogate variable. In Table 22.3, the variables V_1, V_3 and V_5 all have high loadings on factor 1, and all are fairly close in magnitude, although V_1 has relatively the highest loading and would therefore be a likely candidate. However, if prior knowledge suggests that prevention of tooth decay is a very important benefit, V_5 would be selected as the surrogate for factor 1. Also, the choice of a surrogate for factor 2 is not straightforward. Variables V_2, V_4 and V_6 all have comparable high loadings on this factor. If prior knowledge suggests that attractive teeth are the most important social benefit sought from a toothpaste, the researcher would select V_6.

Determine the model fit

The final step in factor analysis involves the determination of model fit. A basic assumption underlying factor analysis is that the observed correlation between variables can be attributed to common factors. Hence, the correlations between the variables can be deduced or reproduced from the estimated correlations between the variables and the factors. The differences between the observed correlations (as given in the input correlation matrix) and the reproduced correlations (as estimated from the factor matrix) can be examined to determine model fit. These differences are called residuals. If there are many large residuals, the factor model does not provide a good fit to the data and the model should be reconsidered. In Table 22.3, we see that only five residuals are larger than 0.05, indicating an acceptable model fit.

The following example further illustrates principal components factoring in the context of trade promotion.

e x a m p l e

Manufacturing promotion components[11]

The objective of this study was to develop a comprehensive inventory of manufacturer-controlled trade promotion variables and to demonstrate that an association exists between these variables and the retailer's promotion support decision. Retailer or trade support was defined operationally as the trade buyer's attitude towards the promotion.

Factor analysis was performed on the explanatory variables with the primary goal of data reduction. The principal components method, using varimax rotation, reduced the 30 explanatory variables to eight factors having eigenvalues greater than 1.0. For the purpose of

interpretation, each factor was composed of variables that loaded 0.40 or higher on that factor. In two instances, where variables loaded 0.40 or above on two factors, each variable was assigned to the factor where it had the highest loading. Only one variable, ease of handling/stocking at retail, did not load at least 0.40 on any factor. In all, the eight factors explained 62% of the total variance. Interpretation of the factor-loading matrix was straightforward. Table 1 lists the factors in the order in which they were extracted.

Table 1 Factors influencing trade promotional support

Factor	Factor interpretation (% variance explained)	Loading	Variables included in the factor
F1	Item importance (16.3%)	0.77	Item is significant enough to warrant promotion
		0.75	Category responds well to promotion
		0.66	Closest trade competitor is likely to promote item
		0.64	Importance of promoted product category
		0.59	Item regular (non-deal) sales volume
		0.57	Deal meshes with trade promotional requirements
			Buyer's estimate of sales increase on the basis of:
F2	Promotion elasticity (9.3%)	0.86	Price reduction and display
		0.82	Display only
		0.80	Price reduction only
		0.70	Price reduction, display and advertising
			Manufacturer's brand support in the form of:
F3	Manufacturer brand support (8.2%)	0.85	Coupons
		0.81	Radio and television advertising
		0.80	Newspaper advertising
		0.75	Point of purchase promotion (e.g. display)
F4	Manufacturer reputation (7.3%)	0.72	Manufacturer's overall reputation
		0.72	Manufacturer's cooperation in meeting trade's promotional needs
		0.64	Manufacturer's cooperation on emergency orders
		0.55	Quality of sales presentation
		0.51	Manufacturer's overall product quality
F5	Promotion wearout (6.4%)	0.93	Product category is over-promoted
		0.93	Item is over-promoted
F6	Sales velocity (5.4%)	−0.81	Brand market share rank[a]
		0.69	Item regular sales volume[a]
		0.46	Item regular sales volume
F7	Item profitability (4.5%)	0.79	Item regular gross margin
		0.72	Item regular gross margin[a]
		0.49	Reasonableness of deal performance requirements
F8	Incentive amount (4.2%)	0.83	Absolute amount of deal allowances
		0.81	Deal allowances as per cent of regular trade cost[a]
		0.49	Absolute amount of deal allowances[a]

[a] Denotes objectives (archival) measure.

585

Stepwise discriminant analysis was conducted to determine which, if any, of the eight factors predicted trade support to a statistically significant degree. The factor scores for the eight factors were the explanatory variables. The dependent variable consisted of the retail buyer's overall rating of the deal (rating), which was collapsed into a three-group (low, medium and high) measure of trade support. The results of the discriminant analyses are shown in Table 2. All eight entered the discriminant functions. Goodness-of-fit measures indicated that, as a group, the eight factors discriminated between high, medium and low levels of trade support. Multivariate F ratios, indicating the degree of discrimination between each pair of groups, were significant at $p < 0.001$. Correct classification into high, medium and low categories was achieved for 65% of the cases. The order of entry into discriminant analysis was used to determine the relative importance of factors as trade support influencers, as shown in Table 3. ■

Table 2 Discriminant analysis results: analysis on rating and performance ($n = 564$)

Factor		Standardised discriminant coefficients Analysis of rating	
		Function 1	Function 2
F1	Item importance	0.861	−0.253
F2	Promotion elasticity	0.081	0.398
F3	Manufacturer brand support	0.127	−0.036
F4	Manufacturer reputation	0.394	0.014
F5	Promotion wearout	−0.207	0.380
F6	Sales velocity	0.033	−0.665
F7	Item profitability	0.614	0.357
F8	Incentive amount	0.461	0.254
Wilks' λ (for each factor)		All significant at $p < 0.001$	
Multivariate F ratios		All significant at $p < 0.001$	
Percentage of cases correctly classified		65% correct ($t = 14.4$, $p < 0.001$)	

Table 3 Relative importance of trade support influencers (as indicated by order of entry into the discriminant analysis)

Analysis of rating	
Order of entry	Factor name
1	Item importance
2	Item profitability
3	Incentive amount
4	Manufacturer reputation
5	Promotion wearout
6	Sales velocity
7	Promotion elasticity
8	Manufacturer brand support

Applications of common factor analysis

The data of Table 22.1 were analysed using the common factor analysis model. Instead of using unities in the diagonal, the communalities were inserted. The output, shown in Table 22.4, is similar to the output from principal components analysis presented in Table 22.3. Beneath the 'Communalities' heading, below the 'Initial' column,

the communalities for the variables are no longer 1.0. Based on the eigenvalue criterion, again two factors are extracted. The variances, after extracting the factors, are different from the initial eigenvalues. The first factor accounts for 42.84% of the variance, whereas the second accounts for 31.13%, in each case a little less than what was observed in principal components analysis.

Table 22.4 Results of common factor analysis

Bartlett test of sphericity
Approximate chi-square = 111.314, *df* = 15, significance = 0.00000
Kaiser-Meyer-Olkin measure of sampling adequacy = 0.660

Communalities

Variable	Initial	Extraction
V_1	0.859	0.928
V_2	0.480	0.562
V_3	0.814	0.836
V_4	0.543	0.600
V_5	0.763	0.789
V_6	0.587	0.723

Initial eigenvalues

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	2.731	45.520	45.520
2	2.218	36.969	82.488
3	0.442	7.360	89.848
4	0.341	5.688	95.536
5	0.183	3.044	98.580
6	0.085	1.420	100.000

Extraction sums of squared loadings

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	2.570	42.837	42.837
2	1.868	31.126	73.964

Factor matrix

	Factor 1	Factor 2
V_1	0.949	0.168
V_2	−0.206	0.720
V_3	0.914	0.038
V_4	−0.246	0.734
V_5	−0.850	−0.259
V_6	−0.101	0.844

Rotation sums of squared loadings

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	2.541	42.343	42.343
2	1.897	31.621	73.964

Table 22.4 Continued

Rotated factor matrix

	Factor 1	Factor 2
V_1	0.963	−0.030
V_2	−0.054	0.747
V_3	0.902	−0.150
V_4	−0.090	0.769
V_5	−0.885	−0.079
V_6	0.075	0.847

Factor score coefficient matrix

	Factor 1	Factor 2
V_1	0.628	0.101
V_2	−0.024	0.253
V_3	0.217	−0.169
V_4	−0.023	0.271
V_5	−0.166	−0.059
V_6	0.083	0.500

Reproduced correlation matrix

Variables	V_1	V_2	V_3	V_4	V_5	V_6
V_1	0.928*	0.022	−0.000	0.024	−0.008	−0.042
V_2	−0.075	0.562*	0.006	−0.008	0.031	0.012
V_3	0.873	−0.161	0.836*	−0.051	0.008	0.042
V_4	−0.110	0.580	−0.197	0.600*	−0.025	−0.004
V_5	−0.850	−0.012	0.786	0.019	0.789*	−0.003
V_6	0.046	0.629	0.060	0.645	−0.133	0.723*

The lower left triangle contains the reproduced correlation matrix; the diagonal, the communalities; and the upper right triangle, the residuals between the observed correlations and the reproduced correlations.

The values in the unrotated factor pattern matrix of Table 22.4 are a little different from those in Table 22.3, although the pattern of the coefficients is similar. Sometimes, however, the pattern of loadings for common factor analysis is different from that for principal components analysis, with some variables loading on different factors. The rotated factors matrix has the same pattern as that in Table 22.3, leading to a similar interpretation of the factors.

We end with another application of common factor analysis, in the context of consumer perception of rebates.

e x a m p l e

'Common' rebate perceptions[12]

Rebates are effective in obtaining new users, brand switching and repeat purchases among current users. A study was undertaken to determine the factors underlying consumer perception of rebates. A set of 24 items measuring consumer perceptions of rebates was constructed. Respondents were asked to express their degree of agreement with these items on five-point Likert scales. The data were collected by a one-stage area telephone survey conducted in the Memphis metropolitan area. A total of 303 usable questionnaires was obtained.

The 24 items measuring perceptions of rebates were analysed using common factor analysis. The initial factor solution did not reveal a simple structure of underlying rebate

perceptions. Therefore, items that had low loadings were deleted from the scale, and the factor analysis was performed on the remaining items. This second solution yielded three interpretable factors. The factor loadings and the reliability coefficients are presented in the table below. The three factors contained four, four and three items, respectively. Factor 1 seemed to capture the consumers' perceptions of the efforts and difficulties associated with rebate redemption (efforts). Factor 2 was defined as a representation of consumers' faith in the rebate system (faith). Factor 3 represented consumers' perceptions of the manufacturers' motives for offering rebates (motives). The loadings of items on their respective factor ranged from 0.527 to 0.744. ■

Factor analysis of perceptions of rebates

Scale items[a]	Factor loading		
	Factor 1	Factor 2	Factor 3
Manufacturers make the rebate process too complicated	0.194	**0.671**	–0.127
Postal rebates are not worth the trouble involved	–0.031	**0.612**	0.352
It takes too long to receive the rebate cheque from the manufacturer	0.013	**0.718**	0.051
Manufacturers could do more to make rebates easier to use	0.205	**0.616**	0.173
Manufacturers offer rebates because consumers want them[b]	**0.660**	0.172	0.101
Today's manufacturers take real interest in consumer welfare[b]	**0.569**	0.203	0.334
Consumer benefit is usually the primary consideration in rebate offers[b]	**0.660**	0.002	0.318
In general, manufacturers are sincere in their rebate offers to consumers[b]	**0.716**	0.047	–0.033
Manufacturers offer rebates to get consumers to buy something they do not really need	0.099	0.156	**0.744**
Manufacturers use rebate offers to induce consumers to buy slow-moving items	0.090	0.027	**0.702**
Rebate offers require you to buy more of a product than you need	0.230	0.066	**0.527**
Eigenvalues	2.030	1.344	1.062
Percentage of explained variance	27.500	12.200	9.700

a The response categories for all items were strongly agree (I), agree (2), neither agree nor disagree (3), disagree (4), strongly disagree (5) and don't know (6). 'Don't know' responses were excluded from data analysis.
b The scores of these items were reversed.

In this example, when the initial factor solution was not interpretable, items which had low loadings were deleted and the factor analysis was performed on the remaining items. If the number of variables is large (greater than 15), principal components analysis and common factor analysis result in similar solutions. Principal components analysis is less prone to misinterpretation, however, and is recommended for the non-expert user.

Internet and computer applications

SPSS[13]

The program FACTOR may be used for principal components analysis as well as for common factor analysis. Some other methods of factor analysis are also available and factor scores are available.

SAS

The program PRINCOMP performs principal components analysis and calculates principal components scores. To perform common factor analysis, the program FACTOR can be used. The FACTOR program also performs principal components analysis.

Minitab

Factor analysis can be accessed using Multivariate>Factor analysis. Principal components or maximum likelihood can be used to determine the initial factor extraction. If maximum likelihood is used, specify the number of factors to extract. If a number is not specified with a principal component extraction, the program will set it equal to a number of variables in the data set.

Excel

At the time of writing, factor analysis was not available.

Summary

Factor analysis is a class of procedures used for reducing and summarising data. Each variable is expressed as a linear combination of the underlying factors. Likewise, the factors themselves can be expressed as linear combinations of the observed variables. The factors are extracted in such a way that the first factor accounts for the highest variance in the data, the second the next highest, and so on. Additionally, it is possible to extract the factors so that the factors are uncorrelated, as in principal components analysis.

In formulating the factor analysis problem, the variables to be included in the analysis should be specified based on past research, theory, and the judgement of the researcher. These variables should be measured on an interval or ratio scale. Factor analysis is based on a matrix of correlation between the variables. The appropriateness of the correlation matrix for factor analysis can be statistically tested.

The two basic approaches to factor analysis are principal components analysis and common factor analysis. In principal components analysis, the total variance in the data is considered. Principal components analysis is recommended when the researcher's primary concern is to determine the minimum number of factors that will account for maximum variance in the data for use in subsequent multivariate analysis. In common factor analysis, the factors are estimated based only on the common variance. This method is appropriate when the primary concern is to identify the underlying dimensions and when the common variance is of interest. This method is also known as principal axis factoring.

The number of factors that should be extracted can be determined *a priori* or based on eigenvalues, scree plots, percentage of variance, split-half reliability or significance tests. Although the initial or unrotated factor matrix indicates the relationships between the factors and individual variables, it seldom results in factors that can be interpreted, because the factors are correlated with many variables. Therefore, rotation is used to transform the factor matrix into a simpler one that is easier to

interpret. The most commonly used method of rotation is the varimax procedure, which results in orthogonal factors. If the factors are highly correlated in the population, oblique rotation can be used. The rotated factor matrix forms the basis for interpreting the factors.

Factor scores can be computed for each respondent. Alternatively, surrogate variables may be selected by examining the factor matrix and selecting a variable with the highest or near highest loading for each factor. The differences between the observed correlations and the reproduced correlations, as estimated from the factor matrix, can be examined to determine model fit.

Questions

1 How is factor analysis different from multiple regression and discriminant analysis?

2 What are the major uses of factor analysis?

3 Describe the factor analysis model.

4 What hypothesis is examined by Bartlett's test of sphericity? For what purpose is this test used?

5 What is meant by the term communality of a variable?

6 Briefly define the following: eigenvalue, factor loadings, factor matrix and factor scores.

7 For what purpose is the Kaiser-Meyer-Olkin measure of sampling adequacy used?

8 What is the major difference between principal components analysis and common factor analysis?

9 Explain how eigenvalues are used to determine the number of factors.

10 What is a scree plot? For what purpose is it used?

11 Why is it useful to rotate the factors? Which is the most common method of rotation?

12 What guidelines are available for interpreting the factors?

13 When is it useful to calculate factor scores?

14 What are surrogate variables? How are they determined?

15 How is the fit of the factor analysis model examined?

Appendix: Fundamental equations of factor analysis[14]

In the factor analysis model, hypothetical components are derived that account for the linear relationship between observed variables. The factor analysis model requires that the relationships between observed variables be linear and that the variables have non-zero correlations between them. The derived hypothetical components have the following properties:

1 They form a linearly independent set of variables. No hypothetical component is derivable from the other hypothetical components as a linear combination of them.

2 The hypothetical components' variables can be divided into two basic kinds of components: common factors and unique factors. These two components can be distinguished in terms of the patterns of weights in the linear equations that derive

the observed variables from the hypothetical components' variables. A common factor has more than one variable with a non-zero weight or factor loading associated with the factor. A unique factor has only one variable with a non-zero weight associated with the factor. Hence, only one variable depends on a unique factor.

3 Common factors are always assumed to be uncorrelated with the unique factors. Unique factors are also usually assumed to be mutually uncorrelated, but common factors may or may not be correlated with each other.

4 Generally, it is assumed that there are fewer common factors than observed variables. The number of unique factors is usually assumed to be equal to the number of observed variables, however.

The following notations are used.

X = an $n \times 1$ random vector of observed random variables $X_1, X_2, X_3, \ldots, X_n$

It is assumed that

$E(X) = 0$
$E(XX') = R_{xx}$, a correlation matrix with unities in the main diagonal
F = an $m \times 1$ vector of m common factors $F_1, F_2, \ldots, F_m$

It is assumed that

$E(F) = 0$
$E(FF') = R_{ff}$, a correlation matrix
U = an $n \times 1$ random vector of the n unique factor variables, $U_1, U_2, \ldots, U_n$

It is assumed that

$$E(U) = O$$
$$E(UU') = I$$

The unique factors are normalised to have unit variances and are mutually uncorrelated.

A = an $n \times m$ matrix of coefficients called the factor pattern matrix
V = an $n \times n$ diagonal matrix of coefficients for the unique factors

The observed variables, which are the coordinates of X, are weighted combinations of the common factors and the unique factors. The fundamental equation of factor analysis can then be written as

$$X = AF + VU$$

The correlations between variables in terms of the factors may be derived as follows:

$$\begin{aligned} R_{xx} &= E(XX') \\ &= E\{(AF + VU)(AF + VU)'\} \\ &= E\{(AF + VU)(F'A' + U'V')\} \\ &= E(AFF'A' + AFU'V' + VUF'A' + VUU'V') \\ &= AR_{ff}A' + AR_{fu}V' + VR_{uf}A' + V^2 \end{aligned}$$

Given that the common factors are uncorrelated with the unique factors, we have

$$R_{fu} = R_{uf}' = 0$$

Hence,

$$R_{xx} = AR_{ff}A' + V^2$$

Suppose that we subtract the matrix of unique factor variance, V^2, from both sides. We then obtain

$$R_{xx} - V^2 = AR_{ff}A'$$

R_{xx} is dependent only on the common factor variables, and the correlations among the variables are related only to the common factors. Let $R_c = R_{xx} - V^2$ be the reduced correlation matrix.

We have already defined the factor pattern matrix A. The coefficients of the factor pattern matrix are weights assigned to the common factors when the observed variables are expressed as linear combinations of the common and unique factors. We now define the factor structure matrix. The coefficients of the factor structure matrix are the covariances between the observed variables and the factors. The factor structure matrix is helpful in the interpretation of factors as it shows which variables are similar to a common factor variable. The factor structure matrix, A_s, is defined as

$$
\begin{aligned}
A_s &= E(XF') \\
&= E[(AF + VU)F'] \\
&= AR_{ff} + VR_{uf} \\
&= AR_{ff}
\end{aligned}
$$

Thus, the factor structure matrix is equivalent to the factor pattern matrix A multiplied by the matrix of covariances among the factors R_{ff}. Substituting A_s for AR_{ff}, the reduced correlation matrix becomes the product of factor structure and the factor pattern matrix:

$$
\begin{aligned}
R_c &= AR_{ff}A' \\
&= A_sA'
\end{aligned}
$$

Notes

1 Birks, D.F. and Birts, A.N., 'Service quality in domestic cash management banks', in Birks, D.F. (ed.), *Global Cash Management in Europe* (Basingstoke: Macmillan, 1998), 175–205. See also Mels, G., Boshof, C. and Nel, D., 'The dimensions of service quality: the original European perspective revisited', *Service Industries Journal* (January 1997), 173–89.

2 Alt, M., *Exploring Hyperspace* (New York: McGraw Hill, 1990), 74.

3 For a detailed discussion of factor analysis, see Tacq, J., *Multivariate Analysis Techniques in Social Science Research* (Thousand Oaks, CA: Sage, 1996); Dunteman, G.H., *Principal Components Analysis* (Newbury Park, CA: Sage, 1989). For an application, see Aaker, J.L., 'Dimensions of brand personality', *Journal of Marketing Research* 34 (August 1997), 347–56.

4 See, for example, Bo Edvardsson, S., Larsson, G. and Setterlind, S., 'Internal service quality and the psychosocial work environment: an empirical analysis of conceptual interrelatedness', *Services Industries Journal* 17(2) (April 1997), 252–63; Taylor, S., 'Waiting for service: the relationship between delays and evaluations of service', *Journal of Marketing* 58 (April 1994), 56–69.

5 See Gaur, S., 'Adelman and Morris factor analysis of developing countries', *Journal of Policy Modeling* 19(4) (August 1997), 407–15; Lastovicka, J.L. and Thamodaran, K., 'Common factor score estimates in multiple regression problems', *Journal of Marketing Research* 28 (February 1991), 105–12; Dillon, W.R. and Goldstein, M., *Multivariate Analysis: Methods and Applications* (New York: Wiley, 1984), 23–99.

6 For applications of factor analysis, see Ittner, C.D. and Larcker, D.F., 'Product development cycle time and organizational performance', *Journal of Marketing Research* 34 (February 1997), 13–23; Ganesan, S., 'Negotiation strategies and the nature of channel relationships', *Journal of Marketing Research* 30 (May 1993), 183–203.

7 Basilevsky, A., *Statistical Factor Analysis and Related Methods: Theory and Applications* (New York: Wiley, 1994); Hair Jr, J.E., Anderson, R.E., Tatham, R.L. and Black, W.C., *Multivariate Data Analysis with Readings*, 5th edn (Englewood Cliffs, NJ: Prentice Hall, 1999), 364–419.

8 Factor analysis is influenced by the relative size of the correlations rather than the absolute size.

9 See Roberts, J.A. and Beacon, D.R., 'Exploring the subtle relationships between environmental concern and ecologically conscious behavior', *Journal of Business Research* 40(1) (September 1997), 79–89; Chatterjee, S., Jamieson, L. and Wiseman, F., 'Identifying most influential observations in factor analysis', *Marketing Science* (Spring 1991), 145–60; Acito, F. and Anderson, R.D., 'A Monte Carlo comparison of factor analytic methods', *Journal of Marketing Research* 17 (May 1980), 228–36.

10 Other methods of orthogonal rotation are also available. The quartimax method minimises the number of factors needed to explain a variable. The equimax method is a combination of varimax and quartimax.

11 Zemanek Jr, J.E., 'Manufacturer influence versus manufacturer salesperson influence over the industrial distributor', *Industrial Marketing Management* 26(1) (January 1997), 59–66; Curhan, R.C. and Kopp, R.J., 'Obtaining retailer support for trade deals: key success factors', *Journal of Advertising Research* (December 1987–January 1988), 51–60.

12 Bulkeley, W.M., 'Rebates' secret appeal to manufacturers: few consumers actually redeem them', *Wall Street Journal* (10 February 1998), B1–B2; Lichtenstein, D.R., Ridgway, N.M. and Netemeyer, R.G., 'Price perceptions and consumer shopping

behaviour: a field study', *Journal of Marketing Research* 30(2) (May 1993), 234–45; Tat, P., Cunningham III, W.A. and Babakus, E., 'Consumer perceptions of rebates', *Journal of Advertising Research* (August–September 1988), 45–50.

13 Einspruch, E.L., *An Introductory Guide to SPSS for Windows* (Thousand Oaks, CA: Sage, 1998); Spector, P.E., *SAS Programming for Researchers and Social Scientists*, (Thousand

Oaks, CA: Sage, 1993); Norat, M.A., Software reviews, *Economic Journal: The Journal of the Royal Economic Society* 107 (May 1997), 857–82; Seiter, C., 'The statistical difference', *Macworld* 10(10) (October 1993), 116–21.

14 The material in this Appendix 22A has been drawn from Stanley A. Muliak, *The Foundations of Factor Analysis* (New York: McGraw-Hill, 1972).

Cluster analysis

Objectives

After reading this chapter, you should be able to:

1 describe the basic concept and scope of cluster analysis and its importance in marketing research;

2 discuss the statistics associated with cluster analysis;

3 explain the procedure for conducting cluster analysis, including formulating the problem, selecting a distance measure, selecting a clustering procedure, deciding on the number of clusters, interpreting clusters and profiling clusters;

4 describe the purpose and methods for evaluating the quality of clustering results and assessing reliability and validity;

5 discuss the applications of non-hierarchical clustering and clustering of variables.

Cluster analysis aims to identify and classify similar entities, based upon the characteristics they possess. It helps the researcher to understand patterns of similarity and difference that reveal naturally occurring groups.

Overview

Like factor analysis (Chapter 22), cluster analysis examines an entire set of interdependent relationships. Cluster analysis makes no distinction between dependent and independent variables. Rather, interdependent relationships between the whole set of variables are examined. The primary objective of cluster analysis is to classify objects into relatively homogeneous groups based on the set of variables considered. Objects in a group are relatively similar in terms of these variables and different from objects in other groups. When used in this manner, cluster analysis is the obverse of factor analysis in that it reduces the number of objects, not the number of variables, by grouping them into a much smaller number of clusters.

This chapter describes the basic concept of cluster analysis. The steps involved in conducting cluster analysis are discussed and illustrated in the context of hierarchical clustering by using a popular computer program. Then an application of non-hierarchical clustering is presented, followed by a discussion of clustering of variables. We begin with two examples.

e x a m p l e

GlobalCash Project

Cluster analysis of European companies' plans over the next two years[1]

In the GlobalCash Project, respondents were clustered on the basis of the changes that respondents said their companies would be making over the next two years. The results indicated that respondents could be clustered into 20 segments. Differences among the segments were statistically tested. Thus, each segment contained respondents who were relatively homogeneous with respect to their plans. The following descriptions encapsulate four of the distinct segments.

- *Restructure through new electronic systems* represented companies whose distinctive plans involved 'making greater use of electronic banking', 'automating the treasury function', 'installing a new treasury system' and 'restructuring cash management along pan-European lines'.
- *Quality focus* represented those companies whose only significant planned change was for 'bank service quality to become a major issue'.
- *All change* represented those companies that planned changes in most areas, perhaps the most volatile of all groups. Most of this group planned to use a pan-European bank and with that change would come a restructuring of cash management on pan-European lines and installation of a new treasury system.
- *Status quo* represented the companies with the lowest amounts of planned changes. None of this group plan to have more automation in their treasury function or put their domestic banking out to tender. ■

e x a m p l e

Ice cream 'hot spots'[2]

In order to achieve an expanded customer base, Häagen-Dazs identified potential consumer segments that could generate additional sales. They used geodemographic techniques (as discussed in Chapter 5), which are based upon clustering consumers, using geographic, demographic and lifestyle data. Additional primary data was collected to develop an understanding of the demographic, lifestyle and behavioural characteristics of Häagen-Dazs Café users, that included frequency of purchase, time of day to visit café, day of the week and a range of other product variables. The postcodes or zip codes of respondents were also obtained. With a postcode or zip code, respondents can be assigned to one of the array of established geodemographic classifications. Häagen-Dazs compared their profile of customers to the profile of geodemographic classifications to develop a clearer picture of the types of consumer they were attracting. From this they decided which profiles of consumer or target markets they believed to hold the most potential for additional sales. ■

Both of the above examples illustrate the use of clustering to arrive at homogeneous segments for the purpose of formulating specific marketing strategies.

Basic concept

Cluster analysis is a class of techniques used to classify objects or cases into relatively homogeneous groups called clusters. Objects in each cluster tend to be similar to each other and dissimilar to objects in the other clusters. Cluster analysis is also called classification analysis or numerical taxonomy.[3] We are concerned with clustering procedures that assign each object to one and only one cluster.[4] Figure 23.1 shows an ideal clustering situation in which the clusters are distinctly separated on two variables: quality consciousness (variable 1) and price sensitivity (variable 2). Note that each consumer falls into one cluster and there are no overlapping areas. Figure 23.2, on the other hand, presents a clustering situation more likely to be encountered in practice. In Figure 23.2, the boundaries for some of the clusters are not clear cut, and the classification of some consumers is not obvious, because many of them could be grouped into one cluster or another.

Both cluster analysis and discriminant analysis are concerned with classification. Discriminant analysis, however, requires prior knowledge of the cluster or group membership for each object or case included, to develop the classification rule. In contrast, in cluster analysis there is no *a priori* information about the group or cluster membership for any of the objects. Groups or clusters are suggested by the data, not defined *a priori*.[5] Cluster analysis has been used in marketing for a variety of purposes, including the following:[6]

■ *Segmenting the market.* For example, consumers may be clustered on the basis of benefits sought from the purchase of a product. Each cluster would consist of consumers who are relatively homogeneous in terms of the benefits they seek.[7] This approach is called benefit segmentation.
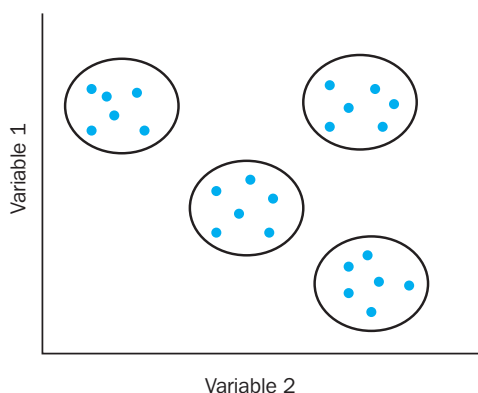
Figure 23.1
An ideal clustering solution
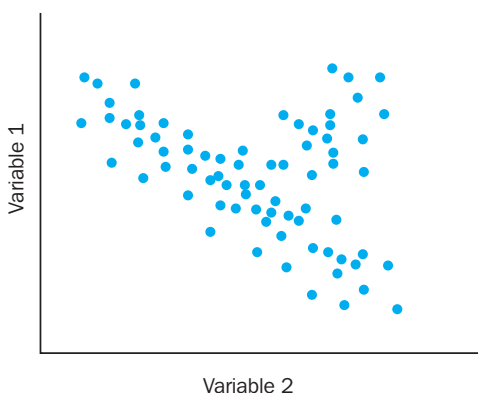
Figure 23.2
A practical clustering solution

- *Understanding buyer behaviours.* Cluster analysis can be used to identify homogeneous groups of buyers. Then the buying behaviour of each group may be examined separately, as happened in another area of the GlobalCash Project. Respondents were clustered on the basis of choice criteria used in selecting a bank.
- *Identifying new product opportunities.* By clustering brands and products, competitive sets within the market can be determined. Brands in the same cluster compete more fiercely with each other than with brands in other clusters. A firm can examine its current offerings compared with those of its competitors to identify potential new product opportunities.
- *Selecting test markets.* By grouping cities into homogeneous clusters, it is possible to select comparable cities to test various marketing strategies.
- *Reducing data.* Cluster analysis can be used as a general data reduction tool to develop clusters or subgroups of data that are more manageable than individual observations. Subsequent multivariate analysis is conducted on the clusters rather than on the individual observations. For example, to describe differences in consumers' product usage behaviour, the consumers may first be clustered into groups. The differences among the groups may then be examined using multiple discriminant analysis.

Before discussing the statistics associated with cluster analysis, it should be mentioned that most clustering methods are relatively simple procedures that are not supported by an extensive body of statistical reasoning. Rather, most clustering methods are heuristics, which are based on algorithms. Thus, cluster analysis contrasts sharply with analysis of variance, regression, discriminant analysis and factor analysis, which are based upon an extensive body of statistical reasoning. Although many clustering methods have important statistical properties, the fundamental simplicity of these methods needs to be recognised.[8] The following statistics and concepts are associated with cluster analysis.

Agglomeration schedule. An agglomeration schedule gives information on the objects or cases being combined at each stage of a hierarchical clustering process.

Cluster centroid. The cluster centroid is the mean values of the variables for all the cases or objects in a particular cluster.

Cluster centres. The cluster centres are the initial starting points in non-hierarchical clustering. Clusters are built around these centres or seeds.

Cluster membership. Cluster membership indicates the cluster to which each object or case belongs.

Dendrogram. A dendrogram, or tree graph, is a graphical device for displaying clustering results. Vertical lines represent clusters that are joined together. The position of the line on the scale indicates the distances at which clusters were joined. The dendrogram is read from left to right. Figure 23.8 later in this chapter is a dendrogram.

Distances between cluster centres. These distances indicate how separated the individual pairs of clusters are. Clusters that are widely separated are distinct and therefore desirable.

Icicle diagram. An icicle diagram is a graphical display of clustering results, so called because it resembles a row of icicles hanging from the eaves of a house. The columns correspond to the objects being clustered, and the rows correspond to the number of clusters. An icicle diagram is read from bottom to top. Figure 23.7 later in this chapter is an icicle diagram.

Similarity/distance coefficient matrix. A similarity/distance coefficient matrix is a lower-triangle matrix containing pairwise distances between objects or cases.

Conducting cluster analysis

The steps involved in conducting cluster analysis are listed in Figure 23.3. The first step is to formulate the clustering problem by defining the variables on which the clustering will be based. Then, an appropriate distance measure must be selected. The distance measure determines how similar or dissimilar the objects being clustered are. Several clustering procedures have been developed, and the researcher should select one that is appropriate for the problem at hand. Deciding on the number of clusters requires judgement on the part of the researcher. The derived clusters should be interpreted in terms of the variables used to cluster them and profiled in terms of additional salient variables. Finally, the researcher must assess the validity of the clustering process.
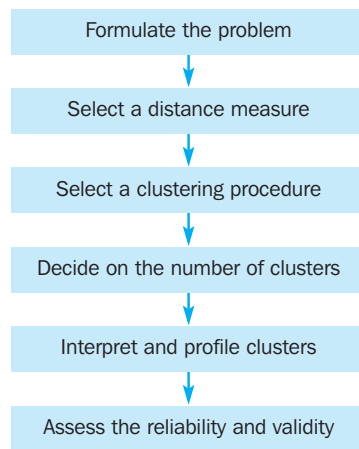
Figure 23.3
Conducting cluster analysis

| Formulate the problem |
| Select a distance measure |
| Select a clustering procedure |
| Decide on the number of clusters |
| Interpret and profile clusters |
| Assess the reliability and validity |

Formulate the problem

Perhaps the most important part of formulating the clustering problem is selecting the variables on which the clustering is based. Inclusion of even one or two irrelevant variables may distort an otherwise useful clustering solution. Basically, the set of variables selected should describe the similarity between objects in terms that are relevant to the marketing research problem. The variables should be selected based on past research, theory or a consideration of the hypotheses being developed or tested. If cluster analysis is used as an exploratory approach, the researcher naturally exercises their judgement and intuition.

To illustrate, we consider a clustering of consumers based on attitudes towards shopping. Based on past research, six attitudinal variables were identified as being the most relevant to the marketing research problem. Consumers were asked to express their degree of agreement with the following statements on a seven-point scale ($1 =$ disagree, $7 =$ agree):

V_1 Shopping is fun.

V_2 Shopping is bad for your budget.

V_3 I combine shopping with eating out.

V_4 I try to get the best buys while shopping.

V_5 I don't care about shopping.

V_6 You can save a lot of money by comparing prices.

Data obtained from a pre-test sample of 20 respondents are shown in Table 23.1. Note that, in practice, clustering is done on much larger samples of 100 or more. A small sample size has been used to illustrate the clustering process.

Table 23.1 Attitudinal data for clustering

Case number	V_1	V_2	V_3	V_4	V_5	V_6
1	6	4	7	3	2	3
2	2	3	1	4	5	4
3	7	2	6	4	1	3
4	4	6	4	5	3	6
5	1	3	2	2	6	4
6	6	4	6	3	3	4
7	5	3	6	3	3	4
8	7	3	7	4	1	4
9	2	4	3	3	6	3
10	3	5	3	6	4	6
11	1	3	2	3	5	3
12	5	4	5	4	2	4
13	2	2	1	5	4	4
14	4	6	4	6	4	7
15	6	5	4	2	1	4
16	3	5	4	6	4	7
17	4	4	7	2	2	5
18	3	7	2	6	4	3
19	4	6	3	7	2	7
20	2	3	2	4	7	2

Select a distance measure

Because the objective of clustering is to group similar objects together, some measure is needed to assess how similar or different the objects are. The most common approach is to measure similarity in terms of distance between pairs of objects. Objects with smaller distances between them are more similar to each other than are those at larger distances. There are several ways to compute the distance between two objects.[9]

Euclidean distance
The square root of the sum of the squared differences in values for each variable.

The most commonly used measure of similarity is the **euclidean distance** or its square.[10] The euclidean distance is the square root of the sum of the squared differences in values for each variable. Other distance measures are also available. The city-block or Manhattan distance between two objects is the sum of the absolute differences in values for each variable. The Chebychev distance between two objects is the maximum absolute difference in values for any variable. For our example, we use the squared euclidean distance.

If the variables are measured in vastly different units, the clustering solution will be influenced by the units of measurement. In a supermarket shopping study, attitudinal variables may be measured on a nine-point Likert-type scale; patronage, in terms of frequency of visits per month and the amount spent; and brand loyalty, in terms of percentage of grocery shopping expenditure allocated to the favourite supermarket. In these cases, before clustering respondents, we must standardise the data by rescal-

ing each variable to have a mean of zero and a standard deviation of unity. Although standardisation can remove the influence of the unit of measurement, it can also reduce the differences between groups on variables that may best discriminate groups or clusters. It is also desirable to eliminate outliers (cases with atypical values).[11]

Use of different distance measures may lead to different clustering results. Hence, it is advisable to use different measures and to compare the results. Having selected a distance or similarity measure, we can next select a clustering procedure.

Select a clustering procedure

Figure 23.4 is a classification of clustering procedures.

Clustering procedures can be hierarchical or non-hierarchical. **Hierarchical clustering** is characterised by the development of a hierarchy or treelike structure. Hierarchical methods can be agglomerative or divisive. **Agglomerative clustering** starts with each object in a separate cluster. Clusters are formed by grouping objects into bigger and bigger clusters. This process is continued until all objects are members of a single cluster. **Divisive clustering** starts with all the objects grouped in a single cluster. Clusters are divided or split until each object is in a separate cluster.

Agglomerative methods are commonly used in marketing research. They consist of linkage methods, error sums of squares or variance methods, and centroid methods. **Linkage methods** include single linkage, complete linkage and average linkage. The **single linkage** method is based on minimum distance or the nearest neighbour rule. The first two objects clustered are those that have the smallest distance between them. The next shortest distance is identified, and either the third object is clustered with the first two or a new two-object cluster is formed. At every stage, the distance between two clusters is the distance between their two closest points (see Figure 23.5). Two clusters are merged at any stage by the single shortest link between them. This process is continued until all objects are in one cluster. The single linkage method does not work well when the clusters are poorly defined. The **complete linkage** method is similar to single linkage, except that it is based on the maximum distance or the farthest neighbour approach. In complete linkage, the distance between two clusters is calculated as the distance between their two farthest points (see Figure 23.5). The average linkage

Hierarchical clustering
A clustering procedure characterised by the development of a hierarchy or treelike structure.

Agglomerative clustering
A hierarchical clustering procedure where each object starts out in a separate cluster. Clusters are formed by grouping objects into bigger and bigger clusters.

Divisive clustering
A hierarchical clustering procedure where all objects start out in one giant cluster. Clusters are formed by dividing this cluster into smaller and smaller clusters.

Linkage methods
Agglomerative methods of hierarchical clustering that cluster objects based on a computation of the distance between them.

Single linkage
A linkage method based on minimum distance or the nearest neighbour rule.

Complete linkage
A linkage method that is based on maximum distance or the farthest neighbour approach.

Figure 23.4
A classification of clustering procedures
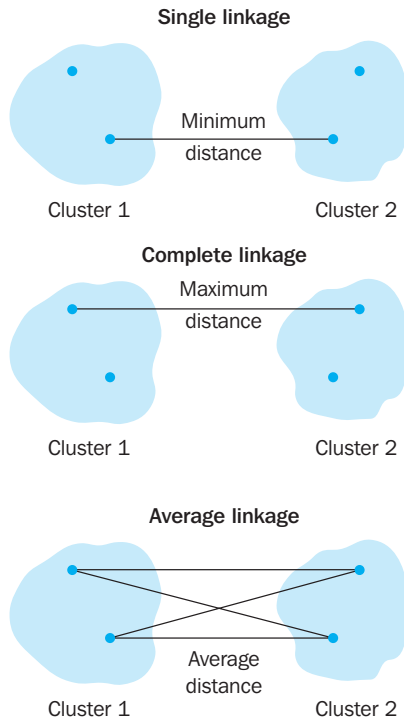
Figure 23.5
Linkage methods of clustering

Variance method
An agglomerative method of hierarchical clustering in which clusters are generated to minimise the within-cluster variance.

Ward's procedure
A variance method in which the squared euclidean distance to the cluster means is minimised.

Centroid method
A variance method of hierarchical clustering in which the distance between two clusters is the distance between their centroids (means for all the variables).

method works similarly. In this method, however, the distance between two clusters is defined as the average of the distances between all pairs of objects, where one member of the pair is from each of the clusters (Figure 23.5). As can be seen, the average linkage method uses information on all pairs of distances, not merely the minimum or maximum distances. For this reason, it is usually preferred to the single and complete linkage methods.

The **variance methods** attempt to generate clusters to minimise the within-cluster variance. A commonly used variance method is **Ward's procedure**. For each cluster, the means for all the variables are computed. Then, for each object, the squared euclidean distance to the cluster means is calculated (Figure 23.6), and these distances are summed for all the objects. At each stage, the two clusters with the smallest increase in the overall sum of squares within cluster distances are combined. In the **centroid method**, the distance between two clusters is the distance between their centroids (means for all the variables), as shown in Figure 23.6. Every time objects are
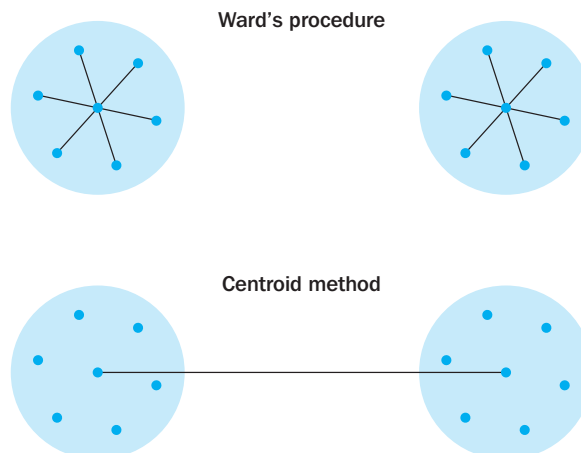
Figure 23.6
Other agglomerative clustering methods

grouped, a new centroid is computed. Of the hierarchical methods, the average linkage method and Ward's procedure have been shown to perform better than the other procedures.[12]

The second type of clustering procedures, the **non-hierarchical clustering** methods, are frequently referred to as *k*-means clustering. These methods include sequential threshold, parallel threshold and optimising partitioning. In the sequential threshold method, a cluster centre is selected and all objects within a prespecified threshold value from the centre are grouped together. A new cluster centre or seed is then selected, and the process is repeated for the unclustered points. Once an object is clustered with a seed, it is no longer considered for clustering with subsequent seeds. The **parallel threshold method** operates similarly except that several cluster centres are selected simultaneously and objects within the threshold level are grouped with the nearest centre. The **optimising partitioning method** differs from the two threshold procedures in that objects can later be reassigned to clusters to optimise an overall criterion, such as average within-cluster distance for a given number of clusters.

Two major disadvantages of the non-hierarchical procedures are that the number of clusters must be prespecified and that the selection of cluster centres is arbitrary. Furthermore, the clustering results may depend on how the centres are selected. Many non-hierarchical programs select the first *k* cases (*k* = number of clusters) without missing values as initial cluster centres. Thus, the clustering results may depend on the order of observations in the data. Yet non-hierarchical clustering is faster than hierarchical methods and has merit when the number of objects or observations is large. It has been suggested that the hierarchical and non-hierarchical methods be used in tandem. First, an initial clustering solution is obtained using a hierarchical procedure, such as average linkage or Ward's. The number of clusters and cluster centroids so obtained are used as inputs to the optimising partitioning method.[13]

The choice of a clustering method and the choice of a distance measure are interrelated. For example, squared euclidean distances should be used with the Ward's and centroid methods. Several non-hierarchical procedures also use squared euclidean distances.

We will use Ward's procedure to illustrate hierarchical clustering. The output obtained by clustering the data of Table 23.1 is given in Table 23.2. Useful information is contained in the agglomeration schedule, which shows the number of cases or clusters being combined at each stage. The first line represents stage 1, with 19 clusters. Respondents 14 and 16 are combined at this stage, as shown in the columns labelled 'Clusters combined'. The squared euclidean distance between these two respondents is given under the column labelled 'Coefficient'. The column entitled 'Stage cluster first appears' indicates the stage at which a cluster is first formed. To illustrate, an entry of 1 at stage 7 indicates that respondent 14 was first grouped at stage 1. The last column, 'Next stage', indicates the stage at which another case (respondent) or cluster is combined with this one. Because the number in the first line of the last column is 7, we see that, at stage 7, respondent 10 is combined with 14 and 16 to form a single cluster. Similarly, the second line represents stage 2 with 18 clusters. In stage 2, respondents 2 and 13 are grouped together.

Another important part of the output is contained in the icicle plot given in Figure 23.7. The columns correspond to the objects being clustered; in this case, they are the respondents labelled 1 to 20. The rows correspond to the number of clusters. This figure is read from bottom to top. At first, all cases are considered as individual clusters. Since there are 20 respondents, there are 20 initial clusters. At the first step, the two closest objects are combined, resulting in 19 clusters. The last line of Figure 23.7 shows these 19 clusters. The two cases, respondents 14 and 16, that have been combined at this stage have no blank space separating them. Row number 18 corresponds

Non-hierarchical clustering
A procedure that first assigns or determines a cluster centre and then groups all objects within a pre-specified threshold value from the centre.

Parallel threshold method
A non-hierarchical clustering method that specifies several cluster centres at once. All objects that are within a pre-specified threshold value from the centre are grouped together.

Optimising partitioning method
A non-hierarchical clustering method that allows for later reassignment of objects to clusters to optimise an overall criterion.

Table 23.2 Results of hierarchical clustering

Agglomeration schedule using Ward's procedure

	Clusters combined			Stage cluster first appears		
Stage	Cluster 1	Cluster 2	Coefficient	Cluster 1	Cluster 2	Next stage
1	14	16	1.000000	0	0	7
2	2	13	2.500000	0	0	15
3	7	12	4.000000	0	0	10
4	5	11	5.500000	0	0	11
5	3	8	7.000000	0	0	16
6	1	6	8.500000	0	0	10
7	10	14	10.166667	0	1	9
8	9	20	12.666667	0	0	11
9	4	10	15.250000	0	7	12
10	1	7	18.250000	6	3	13
11	5	9	22.750000	4	8	15
12	4	19	27.500000	9	0	17
13	1	17	32.700001	10	0	14
14	1	15	40.500000	13	0	16
15	2	5	51.000000	2	11	18
16	1	3	63.125000	14	5	19
17	4	18	78.291664	12	0	18
18	2	4	171.291656	15	17	19
19	1	2	330.450012	16	18	0

Cluster membership of cases using Ward's procedure

	Number of clusters		
Label case	4	3	2
1	1	1	1
2	2	2	2
3	1	1	1
4	3	3	2
5	2	2	2
6	1	1	1
7	1	1	1
8	1	1	1
9	2	2	2
10	3	3	2
11	2	2	2
12	1	1	1
13	2	2	2
14	3	3	2
15	1	1	1
16	3	3	2
17	1	1	1
18	4	3	2
19	3	3	2
20	2	2	2
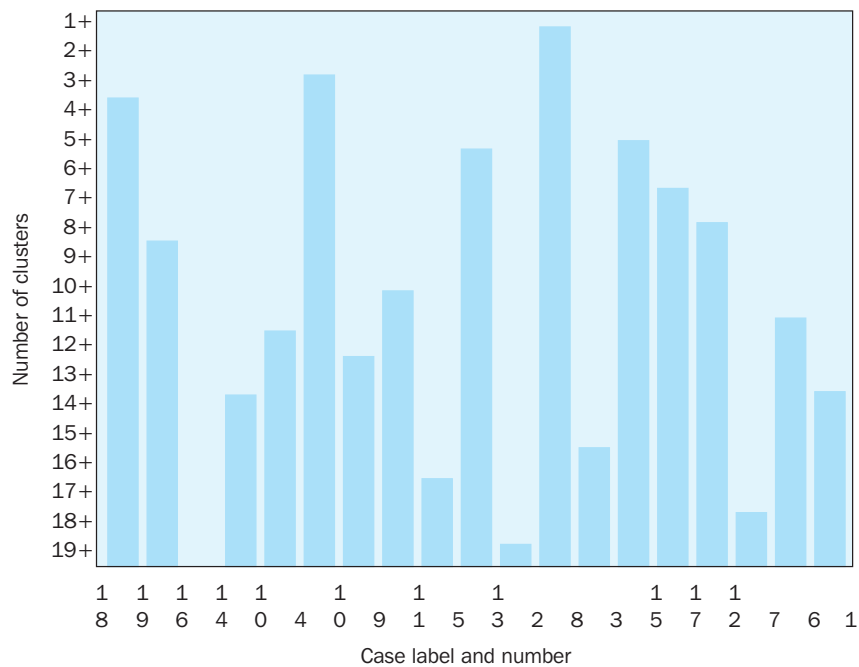
Figure 23.7
Vertical icicle plot using Ward's procedure

to the next stage, with 18 clusters. At this stage, respondents 2 and 13 are grouped together. Thus, at this stage there are 18 clusters; 16 of them consist of individual respondents, and two contain two respondents each. Each subsequent step leads to the formation of a new cluster in one of three ways: (1) two individual cases are grouped together, (2) a case is joined to an already existing cluster, or (3) two clusters are grouped together.

Another graphic device that is useful in displaying clustering results is the dendrogram (see Figure 23.8). The dendrogram is read from left to right. Vertical lines represent clusters that are joined together. The position of the line on the scale indicates the distances at which clusters were joined. Because many distances in the early stages are of similar magnitude, it is difficult to tell the sequence in which some of the early clusters are formed. It is clear, however, that in the last two stages, the distances
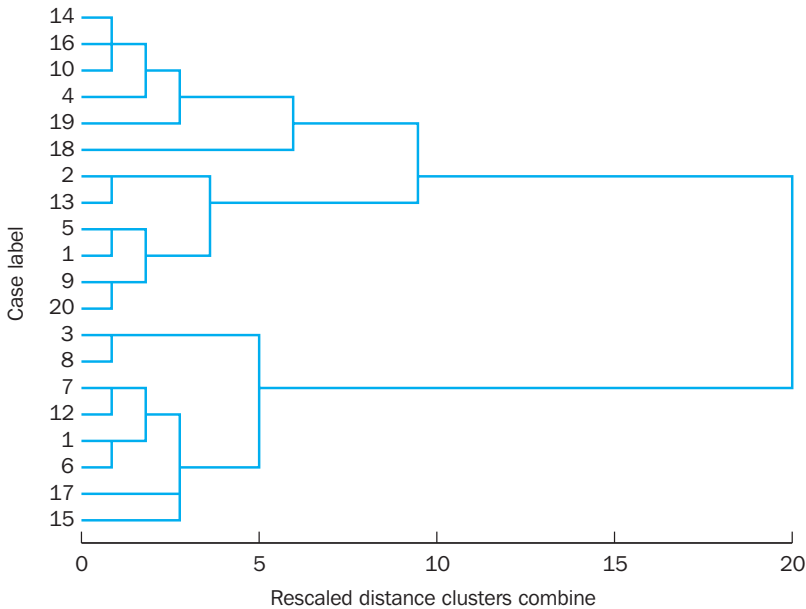
Figure 23.8
Dendrogram using Ward's procedure

at which the clusters are being combined are large. This information is useful in deciding on the number of clusters.

It is also possible to obtain information on cluster membership of cases if the number of clusters is specified. Although this information can be discerned from the icicle plot, a tabular display is helpful. Table 23.2 contains the cluster membership for the cases, depending on whether the final solution contains two, three or four clusters. Information of this type can be obtained for any number of clusters and is useful for deciding on the number of clusters.

Decide on the number of clusters

A major issue in cluster analysis is deciding on the number of clusters. Although there are no hard and fast rules, some guidelines are available.

1 Theoretical, conceptual or practical considerations may suggest a certain number of clusters. For example, if the purpose of clustering is to identify market segments, management may want a particular number of clusters.
2 In hierarchical clustering, the distances at which clusters are combined can be used as criteria. This information can be obtained from the agglomeration schedule or from the dendrogram. In our case, we see from the agglomeration schedule in Table 23.2 that the value in the 'Coefficient' column suddenly more than doubles between stages 17 and 18. Likewise, at the last two stages of the dendrogram in Figure 23.8, the clusters are being combined at large distances. Therefore, it appears that a three-cluster solution is appropriate.
3 In non-hierarchical clustering, the ratio of total within-group variance to between-group variance can be plotted against the number of clusters. The point at which an elbow or a sharp bend occurs indicates an appropriate number of clusters. Increasing the number of clusters beyond this point is usually not worthwhile.
4 The relative sizes of the clusters should be meaningful. In Table 23.2, by making a simple frequency count of cluster membership, we see that a three-cluster solution results in clusters with eight, six and six elements. If we go to a four-cluster solution, however, the sizes of the clusters are eight, six, five and one. It is not meaningful to have a cluster with only one case, so a three-cluster solution is preferable in this situation.

Interpret and profile clusters

Interpreting and profiling clusters involves examining the cluster centroids. The centroids represent the mean values of the objects contained in the cluster on each of the variables. The centroids enable us to describe each cluster by assigning it a name or label. If the clustering program does not print this information, it may be obtained through discriminant analysis. Table 23.3 gives the centroids or mean values for each cluster in our example. Cluster 1 has relatively high values on variables V_1 (Shopping is fun) and V_3 (I combine shopping with eating out). It also has a low value on V_5 (I don't care about shopping). Hence cluster 1 could be labelled 'fun-loving and concerned shoppers'. This cluster consists of cases 1, 3, 6, 7, 8, 12, 15 and 17. Cluster 2 is just the opposite, with low values on V_1 and V_3 and a high value on V_5 and this cluster could be labelled 'apathetic shoppers'. Members of cluster 2 are cases 2, 5, 9, 11, 13 and 20. Cluster 3 has high values on V_2 (Shopping is bad for your budget), V_4 (I try to get the best buys while shopping) and V_6 (You can save a lot of money by comparing prices). Thus, this cluster could be labelled 'economical shoppers'. Cluster 3 is composed of cases 4, 10, 14, 16, 18 and 19. It is often helpful to profile the clusters in terms of variables that were not used for clustering, such as demographic, psychographic, product usage, media usage or other variables. For example, the clusters may have been derived based on benefits sought. Further profiling may be done in terms

Interpreting and profiling clusters involves examining the cluster centroids.
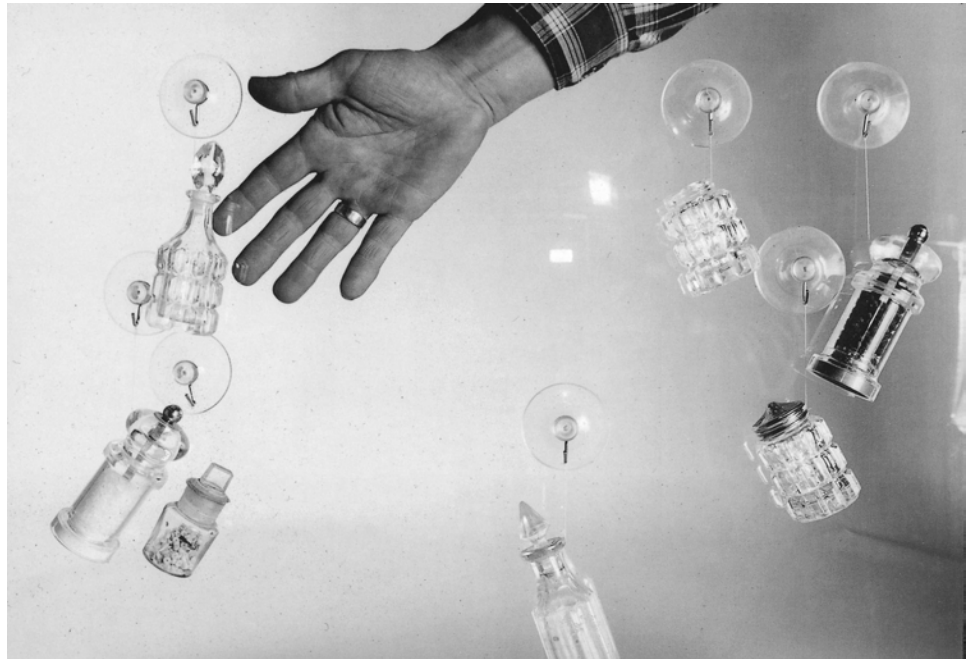
of demographic and psychographic variables to target marketing efforts for each cluster. The variables that significantly differentiate between clusters can be identified via discriminant analysis and one-way analysis of variance.

Table 23.3 Cluster centroids

Cluster number	Means of variables					
	V_1	V_2	V_3	V_4	V_5	V_6
1	5.750	3.625	6.000	3.125	1.750	3.875
2	1.667	3.000	1.833	3.500	5.500	3.333
3	3.500	5.833	3.333	6.000	3.500	6.000

Assess the reliability and validity

Given the several judgements entailed in cluster analysis, no clustering solution should be accepted without some assessment of its reliability and validity. Formal procedures for assessing the reliability and validity of clustering solutions are complex and not fully defensible.[14] Hence, we omit them here. The following procedures, however, provide adequate checks on the quality of clustering results.

1 Perform cluster analysis on the same data using different distance measures. Compare the results across measures to determine the stability of the solutions.
2 Use different methods of clustering and compare the results.
3 Split the data randomly into halves. Perform clustering separately on each half. Compare cluster centroids across the two sub-samples.
4 Delete variables randomly. Perform clustering based on the reduced set of variables. Compare the results with those obtained by clustering based on the entire set of variables.
5 In non-hierarchical clustering, the solution may depend on the order of cases in the data set. Make multiple runs using different order of cases until the solution stabilises.

We further illustrate hierarchical clustering with a study of differences in marketing strategy among British, Japanese and US firms.

It is a small world[15]

Data for a study of British, Japanese and US competitors were obtained from detailed personal interviews with chief executives and top marketing decision makers for defined product groups in 90 companies. To control for market differences, the methodology was based upon matching 30 British companies with their major Japanese and American competitors in the British market. The study involved 30 triads of companies, each composed of a British, US and Japanese business that competed directly with one another.

Most of the data on the characteristics of the companies' performance, strategy and organisation were collected on five-point semantic differential scales. The first stage of the analysis involved factor analysis of variables describing the firms' strategies and marketing activities. The factor scores were used to identify groups of similar companies using Ward's hierarchical clustering routine. A six-cluster solution was developed.

Strategic clusters

Cluster	I	II	III	IV	V	VI
Name	Innovators	Quality marketeers	Price promoters	Product marketeers	Mature marketeers	Aggressive pushers
Size	22	11	14	13	13	17
Successful (%)	55	100	36	38	77	41
Nationality (%)						
British	23	18	64	38	31	29
Japanese	59	46	22	31	15	18
American	18	36	14	31	54	53

Membership in the six clusters was then interpreted against the original performance, strategy and organisational variables. All the clusters contained some successful companies, although some contained significantly more than others. The clusters lent support to the hypothesis that successful companies were similar irrespective of nationality, since British, Japanese and US companies were found in all the clusters. There was, however, a preponderance of Japanese companies in the more successful clusters and a predominance of British companies in the two least successful clusters. Apparently, Japanese companies do not deploy strategies that are unique to them; rather, more of them pursue strategies that work effectively in the British market.

The findings indicate that there are generic strategies that describe successful companies irrespective of their industry. Three successful strategies can be identified. The first is the Quality Marketing strategy. These companies have strengths in marketing and research and development. They concentrate their technical developments on achieving high quality rather than pure innovation. These companies are characterised by entrepreneurial organisations, long-range planning and a well-communicated sense of mission. The second generic strategy is that of the Innovators who are weaker on advanced research and development but are entrepreneurial and driven by a quest for innovation. The last successful group are the Mature Marketers, who are highly profit oriented and have in-depth marketing skills. All three appear to consist of highly marketing-oriented businesses. ■

Applications of non-hierarchical clustering

We illustrate the non-hierarchical procedure using the data in Table 23.1 and an optimising partitioning method. Based on the results of hierarchical clustering, a three-cluster solution was prespecified. The results are presented in Table 23.4.

The initial cluster centres are the values of the first three cases. The classification cluster centres are interim centres used for the assignment of cases. Each case is assigned to the nearest classification cluster centre. The classification centres are

Table 23.4 Results of non-hierarchical clustering

Initial cluster centres

Cluster	V_1	V_2	V_3	V_4	V_5	V_6
1	4.0000	6.0000	3.0000	7.0000	2.0000	7.0000
2	2.0000	3.0000	2.0000	4.0000	7.0000	2.0000
3	7.0000	2.0000	6.0000	4.0000	1.0000	3.0000

Classification cluster centres

Cluster	V_1	V_2	V_3	V_4	V_5	V_6
1	3.8135	5.8992	3.2522	6.4891	2.5149	6.6957
2	1.8507	3.0234	1.8327	3.7864	6.4436	2.5056
3	6.3558	2.8356	6.1576	3.6736	1.3047	3.2010

Case listing of cluster membership

Case ID	Cluster	Distance
1	3	1.780
2	2	2.254
3	3	1.174
4	1	1.882
5	2	2.525
6	3	2.340
7	3	1.862
8	3	1.410
9	2	1.843
10	1	2.112
11	2	1.923
12	3	2.400
13	2	3.382
14	1	1.772
15	3	3.605
16	1	2.137
17	3	3.760
18	1	4.421
19	1	0.853
20	2	0.813

Final cluster centres

Cluster	V_1	V_2	V_3	V_4	V_5	V_6
1	3.5000	5.8333	3.3333	6.0000	3.5000	6.0000
2	1.6667	3.0000	1.8333	3.5000	5.5000	3.3333
3	5.7500	3.6250	6.0000	3.1250	1.7500	3.8750

Table 23.4 Continued

Distances between final cluster centres

Cluster	1	2	3
1	0.0000		
2	5.5678	0.0000	
3	5.7353	6.9944	0.0000

Analysis of variance

Variable	Cluster MS	df	Error MS	df	F	p
V_1	29.1083	2	0.6078	17.0	47.8879	0.000
V_2	13.5458	2	0.6299	17.0	21.5047	0.000
V_3	31.3917	2	0.8333	17.0	37.6700	0.000
V_4	15.7125	2	0.7279	17.0	21.5848	0.000
V_5	24.1500	2	0.7353	17.0	32.8440	0.000
V_6	12.1708	2	1.0711	17.0	11.3632	0.001

Number of cases in each cluster

Cluster	Unweighted cases	Weighted cases
1	6.0	6.0
2	6.0	6.0
3	8.0	8.0
Missing	0.0	
Total	20.0	20.0

updated until the stopping criteria are reached. The final cluster centres represent the variable means for the cases in the final clusters.

Table 23.4 also displays cluster membership and the distance between each case and its classification centre. Note that the cluster memberships given in Table 23.2 (hierarchical clustering) and Table 23.4 (non-hierarchical clustering) are identical. (Cluster 1 of Table 23.2 is labelled cluster 3 in Table 23.4, and cluster 3 of Table 23.2 is labelled cluster 1 in Table 23.4.) The distances between the final cluster centres indicate that the pairs of clusters are well separated. The univariate *F* test for each clustering variable is presented. These *F* tests are only descriptive. Because the cases or objects are systematically assigned to clusters to maximise differences on the clustering variables, the resulting probabilities should not be interpreted as testing the null hypothesis of no differences among clusters.

The following example of hospital choice further illustrates non-hierarchical clustering.

example

Segmentation with surgical precision[16]

Cluster analysis was used to classify and segment respondents, based upon their preferences for hospitals that provide in-patient care. The clustering was based on the reasons respondents gave for preferring a particular hospital. The demographic profiles of the grouped respondents were compared to learn whether the segments could be identified more efficiently.

Because different individuals perceive scales of importance differently, each individual's ratings were normalised before clustering. The results indicated that the respondents could be best classified into four clusters. The cross-validation procedure for cluster analysis was run twice, on halves of the total sample.

As expected, the four groups differed substantially by their distributions and average responses to the reasons for their hospital preferences. The names assigned to the four groups reflected the demographic characteristics and reasons for hospital preferences: '*old-fashioned*', '*affluent*', '*value conscious*', and '*professional want-it-alls*'. ■

Clustering variables

Sometimes cluster analysis is also used for clustering variables to identify homogeneous groups. In this instance, the units used for analysis are the variables, and the distance measures are computed for all pairs of variables. For example, the correlation coefficient, either the absolute value or with the sign, can be used as a measure of similarity (the opposite of distance) between variables.

Hierarchical clustering of variables can aid in the identification of unique variables, or variables that make a unique contribution to the data. Clustering can also be used to reduce the number of variables. Associated with each cluster is a linear combination of the variables in the cluster, called the cluster component. A large set of variables can often be replaced by the set of cluster components with little loss of information. A given number of cluster components does not generally explain as much variance as the same number of principal components, however. Why, then, should the clustering of variables be used? Cluster components are usually easier to interpret than the principal components, even if the latter are rotated.[17] We illustrate the clustering of variables with an example from advertising research.

e x a m p l e

Feelings – nothing more than feelings[18]

A study was conducted to identify feelings that are precipitated by advertising. A total of 655 feelings were reduced to a set of 180 that were judged by respondents to be most likely to be stimulated by advertising. This group was clustered on the basis of judgements of similarity between feelings resulting in 31 feelings clusters. These were divided into 16 positive and 15 negative clusters, as shown in the table.

Positive feelings		Negative feelings	
1	Playful/childish	1	Affraid
2	Friendly	2	Bad/sick
3	Humorous	3	Confused
4	Delighted	4	Indifferent
5	Interested	5	Bored
6	Strong/confident	6	Sad
7	Warm/tender	7	Anxious
8	Relaxed	8	Helpless/timid
9	Energetic/impulsive	9	Ugly/stupid
10	Eager/excited	10	Pity/deceived
11	Contemplative	11	Mad
12	Proud	12	Disagreeable
13	Persuaded/expectant	13	Disgusted
14	Vigorous/challenged	14	Irritated
15	Amazed	15	Moody/frustrated
16	Set/informed		

Thus, 655 feelings responses to advertising were reduced to a core set of 31 feelings. In this way, advertisers now have a manageable set of feelings for understanding and measuring responses to advertising. When measured, these feelings can provide information on a commercial's ability to persuade target consumers ■

Internet and computer applications

SPSS[19]

The main program for hierarchical clustering of objects or cases is CLUSTER. Different distance measures can be computed, and all the hierarchical clustering procedures discussed here are available. For non-hierarchical clustering, the QUICK CLUSTER program can be used. This program is particularly helpful for clustering a large number of cases. All the default options will result in a *k*-means clustering. To cluster variables, the distance measures should be computed across variables using the PROXIMITIES program. This proximity matrix can be read into CLUSTER to obtain a grouping of the variables.

SAS

The CLUSTER program can be used for the hierarchical clustering of cases or objects. All the clustering procedures discussed here are available, as well as some additional ones. Non-hierarchical clustering of cases or objects can be accomplished using FASTCLUS. For clustering of variables, the VARCLUS program can be used. Dendrograms are not automatically computed but can be obtained using the TREE program.

Minitab

Cluster analysis can be accessed in the Multivariate>Cluster observation function. Also available are Clustering of Variables and Cluster *K*-Means.

Excel

At the time of writing, cluster analysis was not available.

Summary

Cluster analysis is used for classifying objects or cases, and sometimes variables, into relatively homogeneous groups. The groups or clusters are suggested by the data and are not defined *a priori*.

The variables on which the clustering is based should be selected based on past research, theory, the hypotheses being tested, or the judgement of the researcher. An appropriate measure of distance or similarity should be selected. The most commonly used measure is the euclidean distance or its square.

Clustering procedures may be hierarchical or non-hierarchical. Hierarchical clustering is characterised by the development of a hierarchy or treelike structure. Hierarchical methods can be agglomerative or divisive. Agglomerative methods consist of linkage methods, variance methods and centroid methods. Linkage methods are composed of single linkage, complete linkage and average linkage. A commonly used variance method is the Ward's procedure. The non-hierarchical methods are frequently referred to as *k*-means clustering. These methods can be classified as

sequential threshold, parallel threshold and optimising partitioning. Hierarchical and non-hierarchical methods can be used in tandem. The choice of a clustering procedure and the choice of a distance measure are interrelated.

The number of clusters may be based on theoretical, conceptual or practical considerations. In hierarchical clustering, the distances at which the clusters are being combined is an important criterion. The relative sizes of the clusters should be meaningful. The clusters should be interpreted in terms of cluster centroids. It is often helpful to profile the clusters in terms of variables that were not used for clustering. The reliability and validity of the clustering solutions may be assessed in different ways.

Questions

1 Discuss the similarity and difference between cluster analysis and discriminant analysis.

2 What is a 'cluster'?

3 What are some of the uses of cluster analysis in marketing?

4 Briefly define the following terms: dendrogram, icicle plot, agglomeration schedule, and cluster membership.

5 What is the most commonly used measure of similarity in cluster analysis?

6 Present a classification of clustering procedures.

7 Upon what basis may a researcher decide which variables should be selected to formulate a clustering problem?

8 Why is the average linkage method usually preferred to single linkage and complete linkage?

9 What are the two major disadvantages of non-hierarchical clustering procedures?

10 What guidelines are available for deciding the number of clusters?

11 What is involved in the interpretation of clusters?

12 What role may qualitative methods play in the interpretation of clusters?

13 What are some of the additional variables used for profiling the clusters?

14 Describe some procedures available for assessing the quality of clustering solutions.

15 How is cluster analysis used to group variables?

Notes

1 Birks, D.F. and Birts, A.N., 'Cash management market segmentation', in Birks, D.F. (ed.), *Global Cash Management in Europe* (Basingstoke, Macmillan, 1998), 83–109.

2 Stuart, L., 'Häagen-Dazs aims to scoop a larger share', *Marketing Week* 19(46/2) (21 February 1997), 26.

3 For applications of cluster analysis, see Kale, S.H., 'Grouping euroconsumers: a culture-based clustering approach', *Journal of International Marketing* 3(3) (1995), 35–48; Day, G.S. and Nedungali, P., 'Managerial representation of competitive advantage', *Journal of Marketing* 58 (April 1994), 31–44.

4 Overlapping clustering methods that permit an object to be grouped into more than one cluster are also available. See Chaturvedi, A., Carroll, J.D., Green, P.E. and Rotondo, J.A., 'A feature based approach to market segmentation via overlapping k-centroids clustering', *Journal of Marketing Research* 34 (August 1997), 370–7.

5 Excellent discussions on the various aspects of cluster analysis may be found in Aldenderfer, M.S. and Blashfield, R.K., *Cluster Analysis* (Beverly Hills. CA: Sage, 1984); Everitt, B., *Cluster Analysis*, 3rd edn (New York: Halsted Press, 1993); and Romsburg, H.C., *Cluster Analysis for Researchers* (Melbourne: Krieger Publishing, 1990).

6 Douglas, V., 'Questionnaires too long? Try variable clustering', *Marketing News* 29(5) (27 February 1995), 38; Punj, G. and Stewart, D., 'Cluster analysis in marketing research: review and suggestions for application', *Journal of Marketing Research* 20 (May 1983), 134–48.

7 For use of cluster analysis for segmentation, see Peterson, M. and Malhotra, N.K., 'Comparative marketing measures of societal quality of life: substantive dimensions in 186 countries', *Journal of Macromarketing* 17(1) (Spring 1997), 25–38; Chang, T.-Z. and Chen, S.-J., 'Benefit segmentation: a useful tool for financial investment services', *Journal of Professional Services Marketing* 12(2) (1995), 69–80; 'Using cluster analysis for segmentation', *Sawtooth News* 10 (Winter 1994–95), 6–7.

8 Everitt, B., *Cluster Analysis*, 3rd edn (New York: Halstead Press, 1993).

9 For a detailed discussion on the different measures of similarity, and formulas for computing them, see Chepoi, V. and Dragan, F., 'Computing a median point of a simple rectilinear polygon', *Information Processing Letters* 49(6) (22 March 1994), 281–5; Romsburg, H.C., *Cluster Analysis for Researchers* (Belmont, CA: Lifetime Learning Publications, 1984).

10 Hirata, T., 'A unified linear-time algorithm for computing distance maps', *Information Processing Letters* 58(3) (13 May 1996) 129–33; Hair Jr, J.E., Anderson, R.E., Tatham, R.L. and Black, W.C., *Multivariate Data Analysis with Readings*, 4th edn (Englewood Cliffs, NJ: Prentice Hall, 1995), 420–83.

11 For further discussion of the issues involved in standardisation, see Romsburg, H.C., *Cluster Analysis for Researchers* (Melbourne: Krieger Publishing, 1990).

12 Johnson, R.A. and Wichern, D.A., *Applied Multivariate Statistical Analysis*, 4th edn (Upper Saddle River, NJ: Prentice Hall, 1998); Milligan, G., 'An examination of the effect of six types of error perturbation on fifteen clustering algorithms', *Psychometrika* 45 (September 1980), 325–42.

13 Everitt, B., *Cluster Analysis*, 3rd edn (New York: Halstead Press, 1993); Punj, G. and Stewart, D., 'Cluster analysis in marketing research: reviews and suggestions for application', *Journal of Marketing Research* 20 (May 1983), 134–48.

14 For a formal discussion of reliability, validity and significance testing in cluster analysis, see Dibbs, S. and Stern, P., 'Questioning the reliability of market segmentation techniques', *Omega* 23(6) (December 1995), 625–36; Funkhouser, G.R., 'A note on the reliability of certain clustering algorithms', *Journal of Marketing Research* 30 (February 1983),

99–102; Klastorin, T.D., 'Assessing cluster analysis results', *Journal of Marketing Research* 20 (February 1983), 92–8; and Arnold, S.J., 'A test for clusters', *Journal of Marketing Research* 16 (November 1979), 545–51.

15 Saunders, J., Wong, V. and Doyle, P., 'The congruence of successful international competitors: a study of successful international competitors: a study of the marketing strategies and organisations of Japanese and US competitors in the UK', *Journal of Global Marketing* 7(3) (1994), 41–59; Doyle, P., Saunders, J. and Wong, V., 'International marketing strategies and organisations: a study of U.S., Japanese, and British competitors', in Bloom, P., Winer, R., Kassarjian, H.H., Scammon, D.L., Weitz, B., Spekman, R.E., Mahajan, V. and Levy, M. (eds), *Enhancing Knowledge Development in Marketing*, Series no. 55 (Chicago, IL: American Marketing Association, 1989), 100–4.

16 Holohean Jr, E.J., Banks, S.M. and Maddy, B.A., 'System impact and methodological issues in the development of an empirical typology of psychiatric hospital residents', *Journal of Mental Health Administration* 22(2) (Spring 1995), 177–88; Woodside, A.G., Nielsen, R.L., Walters, F. and Muller, G.D., 'Preference segmentation of health care services: the old-fashioneds, value conscious, affluents, and professional want-it-alls', *Journal of Health Care Marketing* (June 1988), 14–24.

17 Douglas, V., 'Questionnaire too long? Try variable clustering', *Marketing News* 29(5) (27 February 1995), 38.

18 Helgesen, T., 'The power of advertising – myths and realities', *Marketing and Research Today* 24(2) (May 1996), 63–71; Aaker, D.A., Stayman, D.M. and Vezina, R., 'Identifying feelings elicited by advertising', *Psychology and Marketing* (Spring 1988), 1–16.

19 Einspruch, E.L., *An Introductory Guide to SPSS for Windows* (Thousand Oaks, CA: Sage, 1998); Spector, P.E., *SAS Programming for Researchers and Social Scientists* (Thousand Oaks, CA: Sage, 1993); Norat, M.A., Software reviews, *Economic Journal: The Journal of the Royal Economic Society* 107 (May 1997), 857–82; Seiter, C., 'The statistical difference', *Macworld* 10(10) (October 1993), 116–21.

Multidimensional scaling and conjoint analysis

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

**Stage 5
Data preparation and analysis**

Stage 6
Report preparation and presentation

Objectives

After reading this chapter, you should be able to:

1 discuss the basic concept and scope of multidimensional scaling (MDS) in marketing research and describe its various applications;

2 describe the steps involved in multidimensional scaling of perception data, including formulating the problem, obtaining input data, selecting an MDS procedure, deciding on the number of dimensions, labelling the dimensions and interpreting the configuration, and assessing reliability and validity;

3 explain the multidimensional scaling of preference data and distinguish between internal and external analysis of preferences;

4 explain correspondence analysis and discuss its advantages and disadvantages;

5 understand the relationship between MDS discriminant analysis and factor analysis;

6 discuss the basic concepts of conjoint analysis, contrast it with MDS and discuss its various applications;

7 describe the procedure for conducting conjoint analysis, including formulating the problem, constructing the stimuli, deciding the form of input data, selecting a conjoint analysis procedure, interpreting the results, and assessing reliability and validity;

8 define the concept of hybrid conjoint analysis and explain how it simplifies the data collection task.

Multidimensional scaling allows the perceptions and preferences of consumers to be clearly represented in a spatial map. Conjoint analysis helps to determine the relative importance of attributes that consumers use in choosing products.

Overview

This final chapter on quantitative data analysis presents two related techniques for analysing consumer perceptions and preferences: multidimensional scaling (MDS) and conjoint analysis. We outline and illustrate the steps involved in conducting MDS and discuss the relationships among MDS, factor analysis and discriminant analysis. Then we describe conjoint analysis and present a step-by-step procedure for conducting it. We also provide brief coverage of hybrid conjoint models.

We begin with examples illustrating MDS and conjoint analysis.

e x a m p l e

Colas collide[1]

In a survey, respondents were asked to rank-order all the possible pairs of nine brands of soft drinks in terms of their similarity. These data were analysed via multidimensional scaling and resulted in the following spatial representation of soft drinks.
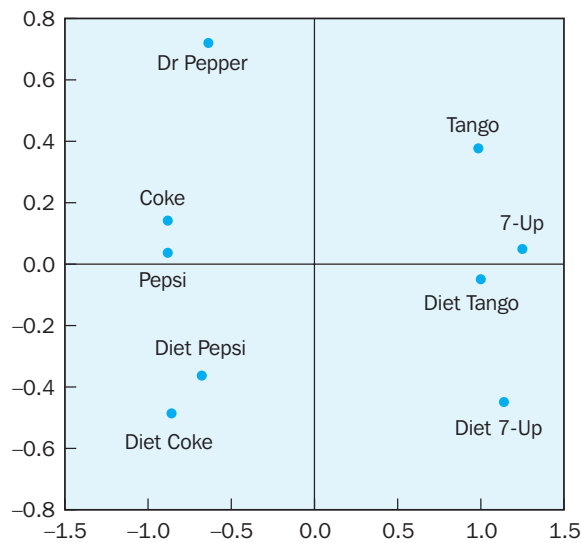
From other information obtained in the questionnaire, the horizontal axis was labelled 'cola flavour'. Diet Coke was perceived to be the most cola-flavoured and 7-Up the least cola-flavoured. The vertical axis was labelled 'dietness', with Diet Coke being perceived to be the most dietetic and Dr Pepper the least dietetic. Note that Coke and Pepsi were perceived to be very similar as indicated by their closeness in the perceptual map. Close similarity was also perceived between 7-Up and Tango, Diet 7-Up and Diet Tango, and Diet Coke and Diet Pepsi. Notice that Dr Pepper is perceived to be relatively dissimilar to the other brands. Such MDS maps are very useful in understanding the competitive structure of the soft drink market. ■

e x a m p l e

The conjoint path over the cultural divide[2]

Boots the Chemist was considering whether to open new stores in the Netherlands, Japan and Thailand. Research was conducted to help decide whether to enter these markets and also to decide which element of Boots' product and service offering to prioritise.

The key research objectives were to:

■ Understand the key drivers of store choice
■ Assess the performance of main competitors already in the market
■ Estimate the proportion of shoppers likely to visit new Boots stores.

Conjoint analysis was used to understand the key drivers of store choice, the impact of features such as range, price, quality, service and convenience, and the trade-offs made in prioritising these features.

To understand the strengths and weaknesses of existing retailers, respondents stated for each of the attributes under review what the named competitors offered. To enable take-up of the new stores to be forecast, respondents were first shown a video of the Boots concept store. The concept store was then assessed on the same series of attributes used for the existing competitors. Over 1,000 interviews were conducted in each country. The research results found:

■ The characteristics of the target market in terms of age, sex, income and lifestage, frequency of and attitudes to shopping.
■ The key success factors in each product area, which influenced store design, merchandising, staff training and marketing decisions.
■ Which existing players posed the greatest threat, in terms of being differentiated from current competitors and having possible areas of leverage against Boots. ■

The first example illustrates the derivation and use of perceptual maps, which lie at the heart of MDS. The Boots example involves the trade-offs that respondents make while evaluating alternatives in choosing stores and desirable features within those stores. The conjoint analysis procedure is based on these trade-offs.

Basic concepts in multidimensional scaling (MDS)

Multidimensional scaling (MDS)
A class of procedures for representing perceptions and preferences of respondents spatially by means of a visual display.

Multidimensional scaling (MDS) is a class of procedures for representing perceptions and preferences of respondents spatially by means of a visual display. Perceived or psychological relationships among stimuli are represented as geometric relationships among points in a multidimensional space. These geometric representations are often called spatial maps. The axes of the spatial map are assumed to denote the psychological bases or underlying dimensions respondents use to form perceptions and preferences for stimuli.[3] MDS has been used in marketing to identify the following:

1 The number and nature of dimensions consumers use to perceive different brands
2 The positioning of brands on these dimensions
3 The positioning of consumers' ideal brand on these dimensions.

Information provided by MDS has been used for a variety of marketing applications, including:

■ *Image measurement.* Comparing the customers' and non-customers' perceptions of the firm with the firm's perceptions of itself.
■ *Market segmentation.* Brands and consumers can be positioned in the same space and thus groups of consumers with relatively homogeneous perceptions can be identified.
■ *New product development.* Gaps in a spatial map indicate potential opportunities for positioning new products. MDS can be used to evaluate new product concepts and existing brands on a test basis to determine how consumers perceive the new concepts. The proportion of preferences for each new product is one indicator of its success.
■ *Assessing advertising effectiveness.* Spatial maps can be used to determine whether advertising has been successful in achieving the desired brand positioning.
■ *Pricing analysis.* Spatial maps developed with and without pricing information can be compared to determine the impact of pricing.
■ *Channel decisions.* Judgements on compatibility of brands with different retail outlets could lead to spatial maps useful for making channel decisions.
■ *Attitude scale construction.* MDS techniques can be used to develop the appropriate dimensionality and configuration of the attitude space.

617

The important statistics and terms associated with MDS include the following:

Similarity judgements. Similarity judgements are ratings on all possible pairs of brands or other stimuli in terms of their similarity using a Likert-type scale.

Preference rankings. Preference rankings are rank orderings of the brands or other stimuli from the most preferred to the least preferred. They are normally obtained from respondents.

Stress. Stress is a lack-of-fit measure; higher values of stress indicate poorer fits.

***R*-square.** *R*-square is a squared correlation index that indicates the proportion of variance of the optimally scaled data that can be accounted for by the MDS procedure. This is a goodness-of-fit measure.

Spatial map. Perceived relationships among brands or other stimuli are represented as geometric relationships among points in a multidimensional space.

Coordinates. Coordinates indicate the positioning of a brand or a stimulus in a spatial map.

Unfolding. The representation of both brands and respondents as points in the same space.

Conducting multidimensional scaling

Figure 24.1 shows the steps in MDS. The researcher must formulate the MDS problem carefully because a variety of data may be used as input into MDS. The researcher must also determine an appropriate form in which data should be obtained and select an MDS procedure for analysing the data. An important aspect of the solution involves determining the number of dimensions for the spatial map. Also, the axes of the map should be labelled and the derived configuration interpreted. Finally, the researcher must assess the quality of the results obtained.[4] We describe each of these steps, beginning with problem formulation.

Formulate the problem

Formulating the problem requires that the researcher specify the purpose for which the MDS results would be used and select the brands or other stimuli to be included in the analysis. The number of brands or stimuli selected and the specific brands included determine the nature of the resulting dimensions and configurations. At a minimum, eight brands or stimuli should be included to obtain a well-defined spatial map. Including more than 25 brands is likely to be cumbersome and may result in respondent fatigue.
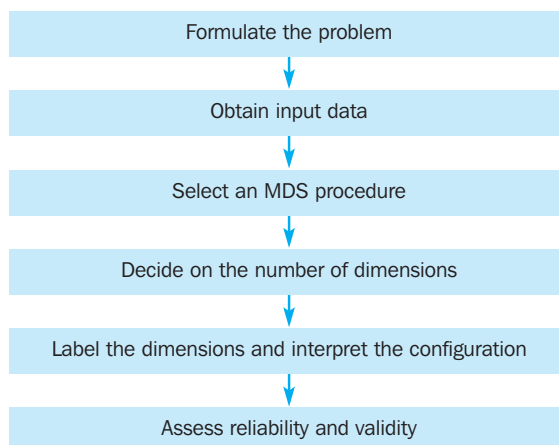
Figure 24.1
Conducting multidimensional scaling

Formulate the problem

Obtain input data

Select an MDS procedure

Decide on the number of dimensions

Label the dimensions and interpret the configuration

Assess reliability and validity

The decision regarding which specific brands or stimuli to include should be made carefully. Suppose that a researcher is interested in obtaining consumer perceptions of cars. If luxury cars are not included in the stimulus set, this dimension may not emerge in the results. The choice of the number and specific brands or stimuli to be included should be based on the statement of the marketing research problem, theory and the judgement of the researcher.

Multidimensional scaling will be illustrated in the context of obtaining a spatial map for 10 brands of beer. These brands are Becks, Budvar, Budweiser, Carlsberg, Corona, Grolsch, Harp, Holsten, San Miguel and Stella Artois. Given the list of brands, the next question is: how should we obtain data on these 10 brands?

Obtain input data

As shown in Figure 24.2, input data obtained from the respondents may be related to perceptions or preferences. Perception data, which may be direct or derived, is discussed first.

Perception data: direct approaches. In direct approaches to gathering perception data, respondents are asked to judge how similar or dissimilar various brands or stimuli are, using their own criteria. Respondents are often required to rate all possible pairs of brands or stimuli in terms of similarity on a Likert scale. These data are referred to as similarity judgements. For example, similarity judgements on all the possible pairs of bottled beer brands may be obtained in the following manner:

	Very dissimilar						Very similar
Becks versus Budweiser	1	2	3	4	5	6	7
Budweiser versus Carlsberg	1	2	3	4	5	6	7
Carlsberg versus Corona	1	2	3	4	5	6	7
. . .							
Becks versus Stella Artois	1	2	3	4	5	6	7

The number of pairs to be evaluated is $n(n-1)/2$, where n is the number of stimuli. Other procedures are also available. Respondents could be asked to rank-order all the possible pairs from the most similar to the least similar. In another method, the respondent rank-orders the brands in terms of their similarity to an anchor brand. Each brand, in turn, serves as the anchor.

In our example, the direct approach was adopted. Subjects were asked to provide similarity judgements for all 45 ($10 \times 9/2$) pairs of bottled beer brands, using a seven-point scale. The data obtained from one respondent are given in Table 24.1. [5]
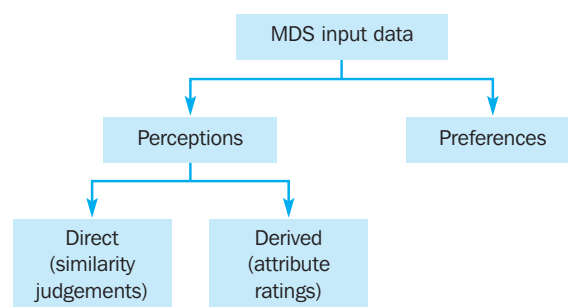
Figure 24.2
Input data for multidimensional scaling

Table 24.1 Similarity ratings of bottled beer brands

	Becks	Budvar	Budweiser	Carlsberg	Corona	Grolsch	Harp	Holsten	San Miguel	Stella Artois
Becks										
Budvar	5									
Budweiser	6	7								
Carlsberg	4	6	6							
Corona	2	3	4	5						
Grolsch	3	3	4	4	5					
Harp	2	2	2	3	5	5				
Holsten	2	2	2	2	6	5	6			
San Miguel	2	2	2	2	6	6	7	6		
Stella Artois	1	2	4	2	4	3	3	4	3	

Perception data: derived approaches. Derived approaches to collecting perception data are attribute-based approaches requiring the respondents to rate the brands or stimuli on the identified attributes using semantic differential or Likert scales. For example, the different brands of bottled beer may be rated on attributes like these:

Best drunk with food	———————————	Best drunk on its own
Bottle feels good to hold	———————————	Bottle does not feel good to hold
Has a strong smell of hops	———————————	No smell of hops

Sometimes an ideal brand is also included in the stimulus set. The respondents are asked to evaluate their hypothetical ideal brand on the same set of attributes. If attribute ratings are obtained, a similarity measure (such as euclidean distance) is derived for each pair of brands.

Direct vs. derived approaches. Direct approaches have the advantage that the researcher does not have to identify a set of salient attributes. Respondents make similarity judgements using their own criteria, as they would under normal circumstances. The disadvantages are that the criteria are influenced by the brands or stimuli being evaluated. If the various brands of cars being evaluated are in the same price range, then price will not emerge as an important factor. It may be difficult to determine before analysis if and how the individual respondent's judgements should be combined. Furthermore, it may be difficult to label the dimensions of the spatial map.

The advantage of the attribute-based approach is that it is easy to identify respondents with homogeneous perceptions. The respondents can be clustered based on the attribute ratings. It is also easier to label the dimensions. A disadvantage is that the researcher must identify all the salient attributes, a difficult task. The spatial map obtained depends on the attributes identified.

The direct approaches are more frequently used than the attribute-based approaches. It may, however, be best to use both these approaches in a complementary way. Direct similarity judgements may be used for obtaining the spatial map, and attribute ratings may be used as an aid to interpreting the dimensions of the perceptual map.

Preference data. Preference data order the brands or stimuli in terms of respondents' preference for some property. A common way in which such data are obtained is preference rankings. Respondents are required to rank the brands from the most preferred to the least preferred. Alternatively, respondents may be required to make

paired comparisons and indicate which brand in a pair they prefer. Another method is to obtain preference ratings for the various brands. (The rank-order, paired comparison and rating scales were discussed in Chapter 12 on scaling techniques.) When spatial maps are based on preference data, distance implies differences in preference. The configuration derived from preference data may differ greatly from that obtained from similarity data. Two brands may be perceived as different in a similarity map yet similar in a preference map, and vice versa. For example, Becks and Harp may be perceived by a group of respondents as very different brands and thus appear far apart on a perception map. But these two brands may be about equally preferred and may appear close together on a preference map.

We continue using the perception data obtained in the bottled beer example to illustrate the MDS procedure and then consider the scaling of preference data.

Select an MDS procedure

Non-metric MDS
A type of multidimensional scaling which assumes that the input data are ordinal.

Metric MDS
A multidimensional scaling method that assumes that input data are metric.

Selecting a specific MDS procedure depends on whether perception or preference data are being scaled or whether the analysis requires both kinds of data. The nature of the input data is also a determining factor. **Non-metric MDS** procedures assume that the input data are ordinal, but they result in metric output. The distances in the resulting spatial map may be assumed to be interval scaled. These procedures find, in a given dimensionality, a spatial map whose rank orders of estimated distances between brands or stimuli best preserve or reproduce the input rank orders. In contrast, **metric MDS** methods assume that input data are metric. Since the output is also metric, a stronger relationship between the output and input data is maintained, and the metric (interval or ratio) qualities of the input data are preserved. The metric and non-metric methods produce similar results.[6]

Another factor influencing the selection of a procedure is whether the MDS analysis will be conducted at the individual respondent level or at an aggregate level. In individual-level analysis, the data are analysed separately for each respondent, resulting in a spatial map for each respondent. Although individual-level analysis is useful from a research perspective, it is not appealing from a managerial standpoint. Marketing strategies are typically formulated at the segment or aggregate level, rather than at the individual level. If aggregate-level analysis is conducted, some assumptions must be made in aggregating individual data. Typically, it is assumed that all respondents use the same dimensions to evaluate the brands or stimuli, but that different respondents weight these common dimensions differentially.

The data of Table 24.1 were treated as rank-ordered and scaled using a non-metric procedure. Because these data were provided by one respondent, an individual-level analysis was conducted. Spatial maps were obtained in one to four dimensions, and then a decision on an appropriate number of dimensions was made. This decision is central to all MDS analyses; therefore, it is explored in greater detail in the following section.

Decide on the number of dimensions

The objective in MDS is to obtain a spatial map that best fits the input data in the smallest number of dimensions. However, spatial maps are computed in such a way that the fit improves as the number of dimensions increases, which means that a compromise has to be made. The fit of an MDS solution is commonly assessed by the stress measure. Stress is a lack-of-fit measure; higher values of stress indicate poorer fits. The following guidelines are suggested for determining the number of dimensions.

1 A priori *knowledge*. Theory or past research may suggest a particular number of dimensions.
2 *Interpretability of the spatial map*. Generally, it is difficult to interpret configurations or maps derived in more than three dimensions.
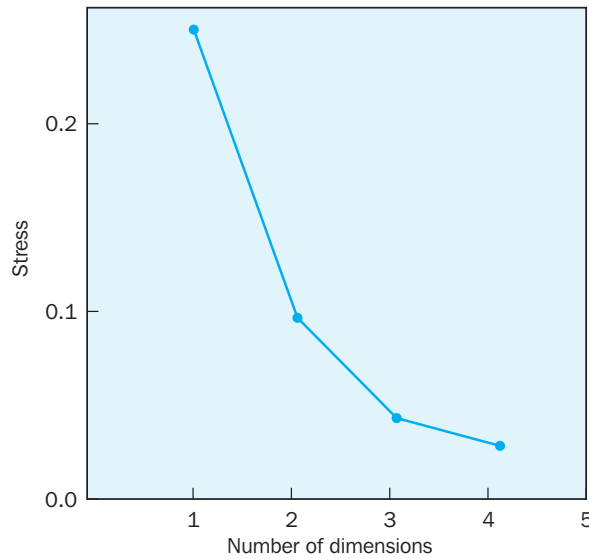
Figure 24.3
Plot of stress versus dimensionality

3 *Elbow criterion.* A plot of stress versus dimensionality should be examined. The points in this plot usually form a convex pattern, as shown in Figure 24.3. The point at which an elbow or a sharp bend occurs indicates an appropriate number of dimensions. Increasing the number of dimensions beyond this point is usually not worth the improvement in fit. This criterion for determining the number of dimensions is called the **elbow criterion**.

Elbow criterion
A plot of stress versus dimensionality used in MDS. The point at which an elbow or a sharp bend occurs indicates an appropriate dimensionality.

4 *Ease of use.* It is generally easier to work with two-dimensional maps or configurations than with those involving more dimensions.

5 *Statistical approaches.* For the sophisticated user, statistical approaches are also available for determining the dimensionality.[7]

Based on the plot of stress versus dimensionality (Figure 24.3), interpretability of the spatial map and ease-of-use criteria, it was decided to retain a two-dimensional solution. This is shown in Figure 24.4.
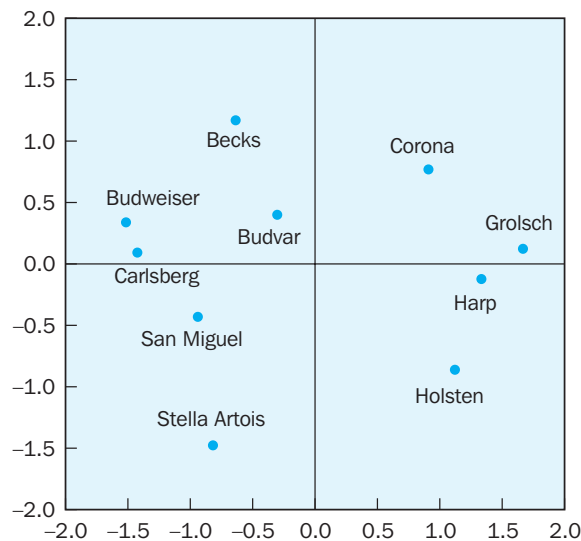
Figure 24.4
A spatial map of beer brands

Label the dimensions and interpret the configuration

Once a spatial map is developed, the dimensions must be labelled and the configuration interpreted. Labelling the dimensions requires subjective judgement on the part of the researcher. The following guidelines can assist in this task:

1 Even if direct similarity judgements are obtained, ratings of the brands on researcher-supplied attributes may still be collected. Using statistical methods such as regression these attribute vectors may be fitted in the spatial map (see Figure 24.5). The axes may then be labelled for the attributes with which they are most closely aligned.
2 After providing direct similarity or preference data, the respondents may be asked to indicate the criteria they used in making their evaluations. These criteria may then be subjectively related to the spatial map to label the dimensions.
3 If possible, the respondents can be shown their spatial maps and asked to label the dimensions by inspecting the configurations.
4 If objective characteristics of the brands are available (e.g. horsepower or kilometres per litre for cars), these could be used as an aid in interpreting the subjective dimensions of the spatial maps.

Often, the dimensions represent more than one attribute. The configuration or the spatial map may be interpreted by examining the coordinates and relative positions of the brands. For example, brands located near each other compete more fiercely than brands far apart. An isolated brand has a unique image. Brands that are farther along in the direction of a descriptor are stronger on that characteristic than others. Thus, the strengths and weaknesses of each product can be understood. Gaps in the spatial map may indicate potential opportunities for introducing new products.

In Figure 24.5, the vertical axis may be labelled as 'strength', representing the power of particular flavours and smells when the beer is first tasted. Brands with high positive values on this axis include Grolsch, Harp, Holsten and Corona. The horizontal axis may be labelled as 'aftertaste', representing the flavour of the beer that lingers on the palate after the beer has been drunk. Brands with large negative values on this dimension include Stella Artois, Holsten and San Miguel. Note that negative scores on the map do not necessarily represent negative characteristics for certain consumers. Thus, the strength of flavour from initial smell and taste through to a strong aftertaste in a brand such as Stella Artois may be seen as desirable characteristics for many beer drinkers.
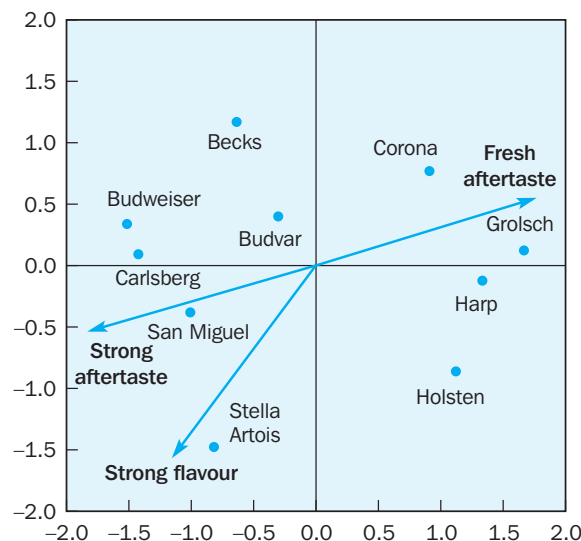
Figure 24.5
Using attribute vectors to label dimensions

623

The gaps in the spatial map indicate potential opportunities for new brands, for example, one that has a strong initial taste but does not have a strong lingering aftertaste.

Assess reliability and validity

The input data, and consequently the MDS solutions, are invariably subject to substantial random variability. Hence, it is necessary that some assessment be made of the reliability and validity of MDS solutions. The following guidelines are suggested.

1 The index of fit, or R-square, should be examined. This is a squared correlation index that indicates the proportion of variance of the optimally scaled data that can be accounted for by the MDS procedure. Thus, it indicates how well the MDS model fits the input data. Although higher values of R-square are desirable, values of 0.60 or better are considered acceptable.

2 Stress values are also indicative of the quality of MDS solutions. Whereas R-square is a measure of goodness-of-fit, stress measures badness-of-fit, or the proportion of variance of the optimally scaled data that is not accounted for by the MDS model. Stress values vary with the type of MDS procedure and the data being analysed. For Kruskal's stress formula 1, the recommendations for evaluating stress values are as follows.[8]

Stress (%)	Goodness of fit
20	Poor
10	Fair
5	Good
2.5	Excellent
0	Perfect

3 If an aggregate-level analysis has been done, the original data should be split into two or more parts. MDS analysis should be conducted separately on each part and the results compared.

4 Stimuli can be selectively eliminated from the input data and the solutions determined for the remaining stimuli.

5 A random error term could be added to the input data. The resulting data are subjected to MDS analysis and the solutions compared.

6 The input data could be collected at two different points in time and the test–retest reliability determined.

Formal procedures are available for assessing the validity of MDS.[9] In the case of our illustrative example, the stress value of 0.095 indicates a fair fit. One brand, namely Stella Artois, is different from the others. Would the elimination of Stella Artois from the stimulus set appreciably alter the relative configuration of the other brands? The spatial map obtained by deleting Stella Artois is shown in Figure 24.6. There is some change in the relative positions of the brands, particularly Corona and Holsten. Yet the changes are modest, indicating fair stability.[10]

Assumptions and limitations of MDS

It is worthwhile to point out some assumptions and limitations of MDS. It is assumed that the similarity of stimulus A to B is the same as the similarity of stimulus B to A. There are some instances where this assumption may be violated. For example, New Zealand is perceived as more similar to Australia than Australia is to New Zealand.
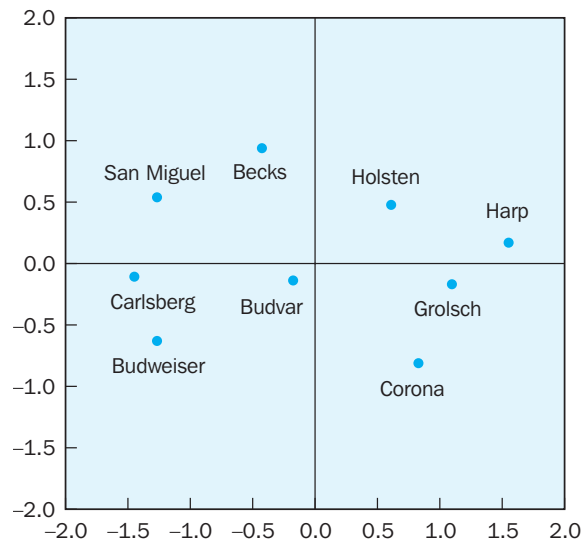
Figure 24.6
Assessment of stability by deleting one brand

MDS assumes that the distance (similarity) between two stimuli is some function of their partial similarities on each of several perceptual dimensions. Not much research has been done to test this assumption. When a spatial map is obtained, it is assumed that inter-point distances are ratio scaled and that the axes of the map are multidimensional interval scaled. A limitation of MDS is that dimension interpretation relating physical changes in brands or stimuli to changes in the perceptual map is difficult at best. These limitations also apply to the scaling of preference data.

Scaling preference data

Internal analysis of preferences
A method of configuring a spatial map such that the spatial map represents both brands or stimuli and respondent points or vectors and is derived solely from the preference data.

External analysis of preferences
A method of configuring a spatial map such that the ideal points or vectors based on preference data are fitted in a spatial map derived from the perception data.

Analysis of preference data can be internal or external. In **internal analysis of preferences**, a spatial map representing both brands or stimuli and respondent points or vectors is derived solely from the preference data. Thus, by collecting preference data, both brands and respondents can be represented in the same spatial map. In **external analysis of preferences**, the ideal points or vectors based on preference data are fitted in a spatial map derived from perception (e.g. similarities) data. To perform external analysis, both preference and perception data must be obtained. The representation of both brands and respondents as points in the same space, by using internal or external analysis, is referred to as *unfolding*.

External analysis is preferred in most situations.[11] In internal analysis, the differences in perceptions are confounded with differences in preferences. It is possible that the nature and relative importance of dimensions may vary between the perceptual space and the preference space. Two brands may be perceived to be similar (located closely to each other in the perceptual space), yet one brand may be distinctly preferred over the other (i.e. the brands may be located apart in the preference space). These situations cannot be accounted for in internal analysis. In addition, internal analysis procedures are beset with computational difficulties.[12]

We illustrate external analysis by scaling the preferences of our respondent into his spatial map. The respondent ranked the brands in the following order of preference (most preferred first): Stella Artois, Holsten, Harp, San Miguel, Carlsberg, Grolsch, Budvar, Budweiser, Corona and Becks. These preference rankings, along with the coordinates of the spatial map (Figure 24.5), were used as input into a preference scaling program to derive Figure 24.7. Notice the location of the ideal point. It is close to Stella Artois, Holsten, Harp and San Miguel, the four most preferred brands, and far
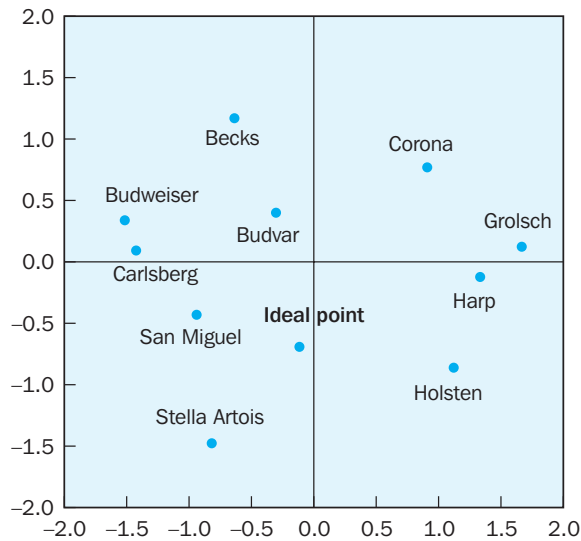
from Corona and Becks, the two least preferred brands. If a new brand were to be located in this space, its distance from the ideal point, relative to the distances of other brands from the ideal point, would determine the degree of preference for this brand.

Although we have considered only quantitative data so far, qualitative data can also be mapped using procedures such as correspondence analysis.

Correspondence analysis

Correspondence analysis
An MDS technique for scaling qualitative data that scales the rows and columns of the input contingency table in corresponding units so that each can be displayed in the same low-dimensional space.

Correspondence analysis is an MDS technique for scaling qualitative data in marketing research. The input data are in the form of a contingency table indicating a qualitative association between the rows and columns. Correspondence analysis scales the rows and columns in corresponding units so that each can be displayed graphically in the same low-dimensional space. These spatial maps provide insights into:

1 Similarities and differences within the rows with respect to a given column category
2 Similarities and differences within the column categories with respect to a given row category
3 Relationships among the rows and columns.[13]

The interpretation of results in correspondence analysis is similar to that in principal components analysis (Chapter 22), given the similarity of the algorithms. Correspondence analysis results in the grouping of categories (activities, brands or other stimuli) found within the contingency table, just as principal components analysis involves the grouping of the independent variables. The results are interpreted in terms of proximities among the rows and columns of the contingency table. Categories that are closer together than others are more similar in underlying structure.[14]

Compared with other multidimensional scaling techniques, the advantage of correspondence analysis is that it reduces the data collection demands imposed on the respondents, since only binary or categorical data are obtained. The respondents are merely asked to tick which attributes apply to each of several brands, or in the GlobalCash study, tick which events they plan to undertake over the next two years. The input data are the number of yes responses for each brand on each attribute. The brands and the attributes are then displayed in the same multidimensional space. The disadvantage is that between-set (i.e. between column and row) distances cannot be

meaningfully interpreted. Other users have criticised the technique as causing confusion when interpreting attribute-brand relationships and complications in the tracking of perceptual changes.[15] Ultimately, it must be remembered that correspondence analysis is an exploratory data analysis technique that is not suitable for hypothesis testing.[16]

Relationship among MDS, factor analysis and discriminant analysis

MDS, including correspondence analysis, is not the only procedure available for obtaining perceptual maps. Two other techniques that we have discussed before, discriminant analysis (Chapter 21) and factor analysis (Chapter 22), can also be used for this purpose.

If the attribute-based approaches are used to obtain input data, spatial maps can also be obtained by using factor or discriminant analysis. In this approach, each respondent rates n brands on m attributes. By factor analysing the data, one could derive for each respondent n factor scores for each factor, one for each brand. By plotting brand scores on the factors, a spatial map could be obtained for each respondent. If an aggregate map is desired, the factor score for each brand for each factor can be averaged across respondents. The dimensions would be labelled by examining the factor loadings, which are estimates of the correlations between attribute ratings and underlying factors.[17]

The goal of discriminant analysis is to select the linear combinations of attributes that best discriminate between the brands or stimuli. To develop spatial maps by means of discriminant analysis, the dependent variable is the brand rated and the independent or predictor variables are the attribute ratings. A spatial map can be obtained by plotting the discriminant scores for the brands. The discriminant scores are the ratings on the perceptual dimensions, based on the attributes which best distinguish the brands. The dimensions can be labelled by examining the discriminant weights, or the weightings of attributes that make up a discriminant function or dimension.[18]

Basic concepts in conjoint analysis

Conjoint analysis
A technique that attempts to determine the relative importance consumers attach to salient attributes and the utilities they attach to the levels of attributes.

Conjoint analysis attempts to determine the relative importance consumers attach to salient attributes and the utilities they attach to the levels of attributes.[19] This information is derived from consumers' evaluations of brands or from brand profiles composed of these attributes and their levels. The respondents are presented with stimuli that consist of combinations of attribute levels. They are asked to evaluate these stimuli in terms of their desirability. Conjoint procedures attempt to assign values to the levels of each attribute so that the resulting values or utilities attached to the stimuli match, as closely as possible, the input evaluations provided by the respondents. The underlying assumption is that any set of stimuli – such as products, brands or banks – are evaluated as a bundle of attributes.[20]

Like multidimensional scaling, conjoint analysis relies on respondents' subjective evaluations. In MDS, however, the stimuli are products or brands. In conjoint analysis, the stimuli are combinations of attribute levels determined by the researcher. The goal in MDS is to develop a spatial map depicting the stimuli in a multidimensional perceptual or preference space. Conjoint analysis, on the other hand, seeks to develop the part-worth or utility functions describing the utility consumers attach to the levels of each attribute. The two techniques are complementary.[21]

Conjoint analysis has been used in marketing for a variety of purposes, including the following:

■ *Determining the relative importance of attributes in the consumer choice process.* A standard output from conjoint analysis consists of derived relative importance weights. The relative importance weights indicate which attributes are important in influencing consumer choice.
■ *Estimating market share of brands that differ in attribute levels.* The utilities derived from conjoint analysis can be used as input into a choice simulator to determine the share of choices, and hence the market share, of different brands.
■ *Determining the composition of the most preferred brand.* Brand features can be varied in terms of attribute levels and the corresponding utilities determined. The brand features that yield the highest utility indicate the composition of the most preferred brand.
■ *Segmenting the market based on similarity of preferences for attribute levels.* The part-worth functions derived for the attributes may be used as a basis for clustering respondents to arrive at homogeneous preference segments.[22]

Applications of conjoint analysis have been made in consumer goods, industrial goods and financial and other services. Moreover, these applications have spanned all areas of marketing. A recent survey of conjoint analysis reported applications in the areas of new product and concept identification, competitive analysis, pricing, market segmentation, advertising and distribution.[23]

The important statistics and terms associated with conjoint analysis include the following:

Part-worth functions. The part-worth or *utility functions* describe the utility consumers attach to the levels of each attribute.

Relative importance weights. The relative importance weights are estimated and indicate which attributes are important in influencing consumer choice.

Attribute levels. Denote the values assumed by the attributes.

Full profiles. Full profiles or complete profiles of brands are constructed in terms of all the attributes by using the attribute levels specified by the design.

Pairwise tables. Respondents evaluate two attributes at a time until all the required pairs of attributes have been evaluated.

Cyclical designs. Designs employed to reduce the number of paired comparisons.

Fractional factorial designs. Designs employed to reduce the number of stimulus profiles to be evaluated in the full-profile approach.

Orthogonal arrays. A special class of fractional designs that enable the efficient estimation of all main effects.

Internal validity. This involves correlations of the predicted evaluations for the hold-out or validation stimuli with those obtained from the respondents.

Conducting conjoint analysis

Figure 24.8 lists the steps in conjoint analysis. Formulating the problem involves identifying the salient attributes and their levels. These attributes and levels are used for constructing the stimuli to be used in a conjoint evaluation task. The respondents rate or rank the stimuli using a suitable scale, and the data obtained are analysed. The results are interpreted and their reliability and validity assessed. We now describe each of the steps of conjoint analysis in detail.
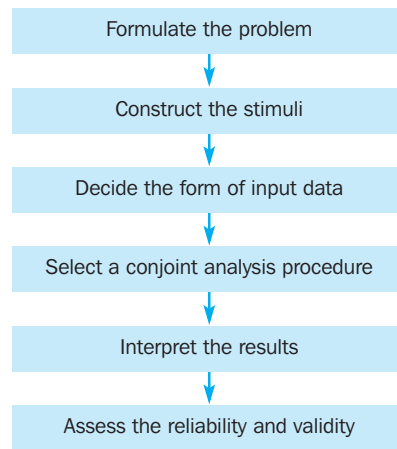
Figure 24.8
Conducting conjoint analysis

Formulate the problem

In formulating the conjoint analysis problem, the researcher must identify the attributes and attribute levels to be used in constructing the stimuli. Attribute levels denote the values assumed by the attributes. From a theoretical standpoint, the attributes selected should be salient in influencing consumer preference and choice. For example, in the choice of a car, price, fuel efficiency, interior space and so forth should be included. From a managerial perspective, the attributes and their levels should be actionable. To tell a manager that consumers prefer a sporty car to one that is conservative-looking is not helpful, unless sportiness and conservativeness are defined in terms of attributes over which a manager has control. The attributes can be identified through discussions with management and industry experts, analysis of secondary data, qualitative research and pilot surveys. A typical conjoint analysis study may involve six or seven attributes.

Once the salient attributes have been identified, their appropriate levels should be selected. The number of attribute levels determines the number of parameters that will be estimated and also influences the number of stimuli that will be evaluated by the respondents. To minimise the respondent evaluation task and yet estimate the parameters with reasonable accuracy, it is desirable to restrict the number of attribute levels. The utility or part-worth function for the levels of an attribute may be non-linear. For example, a consumer may prefer a medium-sized car to either a small or a large one. Likewise, the utility for price may be non-linear. The loss of utility in going from a low price to a medium price may be much smaller than the loss in utility in going from a medium price to a high price. In these cases, at least three levels should be used. Some attributes, though, may naturally occur in binary form (two levels): a car does or does not have a sunroof.

The attribute levels selected will affect the consumer evaluations. If the price of a car brand is varied at €14,000, €16,000 and €18,000, price will be relatively unimportant. On the other hand, if the price is varied at €20,000, €30,000 and €40,000, it will be an important factor. Hence, the researcher should take into account the attribute levels prevalent in the marketplace and the objectives of the study. Using attribute levels that are beyond the range reflected in the marketplace will decrease the believability of the evaluation task, but it will increase the accuracy with which the parameters are estimated. The general guideline is to select attribute levels so that the ranges are somewhat greater than those prevalent in the marketplace but not so large as to impact the believability of the evaluation task adversely.

We illustrate the conjoint methodology by considering the problem of how students evaluate boots, for example brands such as Dr Martens, Timberland, Bally and

Caterpillar. Qualitative research identified three attributes as salient: the material used for the upper, the country or region in which they were designed and manufactured and the price. Each was defined in terms of three levels, as shown in Table 24.2. These attributes and their levels were used for constructing the conjoint analysis stimuli. It has been argued that pictorial stimuli should be used when consumers' marketplace choices are strongly guided by the product's styling, such that the choices are heavily based on an inspection of actual products or pictures of products.[24]

Table 24.2 Boot attributes and levels

Attribute	Level	
	Number	Description
Uppers	3	Leather
	2	Suede
	1	Imitation leather
Country	3	Italy
	2	America
	1	Far East
Price	3	€50
	2	€125
	1	€200

Construct the stimuli

Two broad approaches are available for constructing conjoint analysis stimuli: the pairwise approach and the full-profile procedure.

In the pairwise approach, also called two-factor evaluations, respondents evaluate two attributes at a time until all the possible pairs of attributes have been evaluated.

Qualitative research revealed that students do not evaluate boots based on the sole patterns they leave on cars.

You will be presented with information on Boots in terms of pairs of features described in the form of a matrix. For each matrix, please rank the nine feature combinations in terms of your preference. A rank of 1 should be assigned to the most preferred combination and 9 to least preferred.
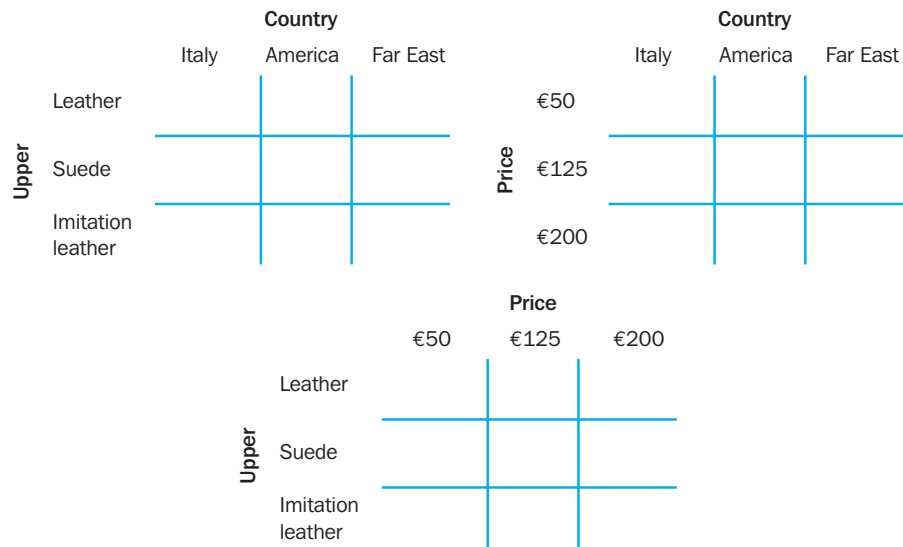
Figure 24.9
Pairwise approach to collecting conjoint data

This approach is illustrated in the context of the boots example in Figure 24.9. For each pair, respondents evaluate all the combinations of levels of both the attributes, which are presented in a matrix.

In the full-profile approach, also called multiple-factor evaluations, full or complete profiles of brands are constructed for all the attributes. Typically, each profile is described on a separate index card. This approach is illustrated in the context of the boots example in Table 24.3.

Table 24.3 Full-profile approach to collecting conjoint data

Example of boot product profile	
Upper	Made of leather
Country	Designed and made in Italy
Price	Costing €200

It is not necessary to evaluate all the possible combinations, nor is it feasible in all cases. In the pairwise approach, it is possible to reduce the number of paired comparisons by using cyclical designs. Likewise, in the full-profile approach, the number of stimulus profiles can be greatly reduced by means of fractional factorial designs. A special class of fractional designs, orthogonal arrays, allows for the efficient estimation of all main effects. Orthogonal arrays permit the measurement of all main effects of interest on an uncorrelated basis. These designs assume that all interactions are negligible. Orthogonal arrays are constructed from basic full factorial designs by substituting a new factor for selected interaction effects that are presumed to be negligible.[25] Generally, two sets of data are obtained. One, the estimation set, is used to calculate the part-worth functions for the attribute levels. The other, the holdout set, is used to assess reliability and validity.

The advantage of the pairwise approach is that it is easier for the respondents to provide these judgements. Its relative disadvantage, however, is that it requires more evaluations than the full-profile approach. Also, the evaluation task may be unrealistic

when only two attributes are being evaluated simultaneously. Studies comparing the two approaches indicate that both methods yield comparable utilities, yet the full-profile approach is more commonly used.

The boots example follows the full-profile approach. Given three attributes, defined at three levels each, a total of $3 \times 3 \times 3 = 27$ profiles can be constructed. To reduce the respondent evaluation task, a fractional factorial design was employed and a set of nine profiles was constructed to constitute the estimation stimuli set (see Table 24.4). Another set of nine stimuli was constructed for validation purposes. Input data were obtained for both the estimation and validation stimuli. Before the data could be obtained, however, it was necessary to decide on the form of the input data.[26]

Table 24.4 **Boot profiles and their ratings**

Profile number	Attribute levels[a]			
	Upper	Country	Price	Preference rating
1	1	1	1	9
2	1	2	2	7
3	1	3	3	5
4	2	1	2	6
5	2	2	3	5
6	2	3	1	6
7	3	1	3	5
8	3	2	1	7
9	3	3	2	6

[a]The attribute levels correspond to those in Table 24.2.

Decide on the form of input data

As in the case of MDS, conjoint analysis input data can be either non-metric or metric. For non-metric data, respondents are typically required to provide rank-order evaluations. For the pairwise approach, respondents rank all the cells of each matrix in terms of their desirability. For the full-profile approach, they rank all the stimulus profiles. Rankings involve relative evaluations of the attribute levels. Proponents of ranking data believe that such data accurately reflect the behaviour of consumers in the marketplace.

In the metric form, respondents provide ratings, rather than rankings. In this case, the judgements are typically made independently. Advocates of rating data believe they are more convenient for the respondents and easier to analyse than rankings. In recent years, the use of ratings has become increasingly common.

In conjoint analysis, the dependent variable is usually preference or intention to buy. In other words, respondents provide ratings or rankings in terms of their preference or intentions to buy. The conjoint methodology, however, is flexible and can accommodate a range of other dependent variables, including actual purchase or choice.

In evaluating boot profiles, respondents were required to provide preference ratings for the boots described by the nine profiles in the estimation set. These ratings were obtained using a nine-point Likert scale (1 = not preferred, 9 = greatly preferred). Ratings obtained from one respondent are shown in Table 24.4.

Select a conjoint analysis procedure

The basic **conjoint analysis model** may be represented by the following formula:[27]

$$U(X) = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \alpha_{ij} x_{ij}$$

where

$U(X)$ = overall utility of an alternative

α_{ij} = the part-worth contribution or utility associated with the jth level ($j = 1, 2, \ldots, k_j$) of the ith attribute ($i = 1, 2, \ldots, m$)

k_i = number of levels of attribute i

m = number of attributes

The importance of an attribute, I_i, is defined in terms of the range of the part-worths, α_{ij}, across the levels of that attribute:

$$I_i = \{\max(\alpha_{ij}) - \min(\alpha_{ij})\} \text{ for each } i$$

The attribute's importance is normalised to ascertain its importance relative to other attributes, W_i:

$$W_i = \frac{I_i}{\sum_{i=1}^{m} I_i}$$

so that

$$\sum_{i=1}^{m} W_i = 1$$

Several different procedures are available for estimating the basic model. The simplest is dummy variable regression (see Chapter 20). In this case, the predictor variables consist of dummy variables for the attribute levels. If an attribute has k_i levels, it is coded in terms of $k_i - 1$ dummy variables. If metric data are obtained, the ratings, assumed to be interval scaled, form the dependent variable. If the data are non-metric, the rankings may be converted to 0 or 1 by making paired comparisons between brands. In this case, the predictor variables represent the differences in the attribute levels of the brands being compared. Other procedures that are appropriate for non-metric data include LINMAP, MONANOVA and the LOGIT model.[28]

The researcher must also decide whether the data will be analysed at the individual respondent or the aggregate level. At the individual level, the data of each respondent are analysed separately. If an aggregate-level analysis is to be conducted, some procedure for grouping the respondents must be devised. One common approach is to estimate individual-level part-worth or utility functions first. Respondents are then clustered on the basis of the similarity of their part-worth functions. Aggregate analysis is then conducted for each cluster.[29] An appropriate model for estimating the parameters should be specified.[30]

The data reported in Table 24.4 were analysed using ordinary least squares (OLS) regression with dummy variables. The dependent variable was the preference ratings. The independent variables or predictors were six dummy variables, two for each variable. The transformed data are shown in Table 24.5.

Since the data pertain to a single respondent, an individual-level analysis was conducted. The part-worth or utility functions estimated for each attribute, as well as the relative importance of the attributes, are given in Table 24.6.[31]

Table 24.5 Boot data coded for dummy variable regression

Preference ratings	Attributes					
	Upper		Country		Price	
Y	X_1	X_2	X_3	X_4	X_5	X_6
9	1	0	1	0	1	0
7	1	0	0	1	0	1
5	1	0	0	0	0	0
6	0	1	1	0	0	1
5	0	1	0	1	0	0
6	0	1	0	0	1	0
5	0	0	1	0	0	0
7	0	0	0	1	1	0
6	0	0	0	0	0	1

Table 24.6 Boot attributes and levels

Attribute	Level			
	Number	Description	Utility	Importance
Uppers	3	Leather	0.778	
	2	Suede	−0.556	
	1	Imitation leather	−0.222	0.268
Country	3	Italy	0.445	
	2	America	0.111	
	1	Far East	−0.556	0.214
Price	3	€50	1.111	
	2	€125	0.111	
	1	€200	−1.222	0.500

The model estimated may be represented as

$$U = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5 + b_6 X_6$$

where X_1, X_2 = dummy variables representing upper
X_3, X_4 = dummy variables representing country
X_5, X_6 = dummy variables representing price

For upper, the attribute levels were coded as follows:

	X_1	X_2
Level 1	1	0
Level 2	0	1
Level 3	0	0

The levels of the other attributes were coded similarly. The parameters were estimated as follows:

$b_0 = 4.222$
$b_1 = 1.000$
$b_2 = -0.333$
$b_3 = 1.000$
$b_4 = 0.667$
$b_5 = 2.333$
$b_6 = 1.333$

Given the dummy variable coding, in which level 3 is the base level, the coefficients may be related to the part-worths. As explained in Chapter 20, each dummy variable coefficient represents the difference in the part-worth for that level minus the part-worth for the base level. For upper, we have the following:

$$\alpha_{11} - \alpha_{13} = b_1$$
$$\alpha_{12} - \alpha_{13} = b_2$$

To solve for the part-worths, an additional constraint is necessary. The part-worths are estimated on an interval scale, so the origin is arbitrary. Therefore, the additional constraint imposed is of the form

$$\alpha_{11} + \alpha_{12} + \alpha_{13} = 0$$

These equations for the first attribute, upper, are

$$\alpha_{11} - \alpha_{13} = 1.000$$
$$\alpha_{12} - \alpha_{13} = -0.333$$
$$\alpha_{11} + \alpha_{12} + \alpha_{13} = 0$$

Solving these equations, we get

$$\alpha_{11} = 0.778$$
$$\alpha_{12} = -0.556$$
$$\alpha_{13} = -0.222$$

The part-worths for other attributes reported in Table 24.6 can be estimated similarly. For country, we have

$$\alpha_{21} - \alpha_{23} = b_3$$
$$\alpha_{22} - \alpha_{23} = b_4$$
$$\alpha_{21} + \alpha_{22} + \alpha_{23} = 0$$

For the third attribute, price, we have

$$\alpha_{31} - \alpha_{33} = b_5$$
$$\alpha_{32} - \alpha_{33} = b_6$$
$$\alpha_{31} + \alpha_{32} + \alpha_{33} = 0$$

The relative importance weights were calculated based on ranges of part-worths, as follows:

Sum of ranges of part-worths $= [0.778 - (-0.556)] + [0.445 - (-0.556)]$
$+ [1.111 - (-1.222)]$
$= 4.668$

$$\text{Relative importance of upper} \quad = \frac{1.334}{4.668} = 0.286$$

$$\text{Relative importance of country} = \frac{1.001}{4.668} = 0.214$$

$$\text{Relative importance of price} \quad = \frac{2.333}{4.668} = 0.500$$

The estimation of the part-worths and the relative importance weights provides the basis for interpreting the results.

Interpret the results

For interpreting the results, it is helpful to plot the part-worth functions. The part-worth function values for each attribute given in Table 24.6 are graphed in Figure 24.10. As can be seen from Table 24.6 and Figure 24.10, this respondent has the greatest preference for a leather upper when evaluating boots. Second preference is for an imitation leather upper, and a suede upper is least preferred. An Italian boot is most preferred, followed by American boots and boots from the Far East. As may be expected, a price of €50.00 has the highest utility and a price of €200.00 the lowest. The utility values reported in Table 24.6 have only interval scale properties, and their
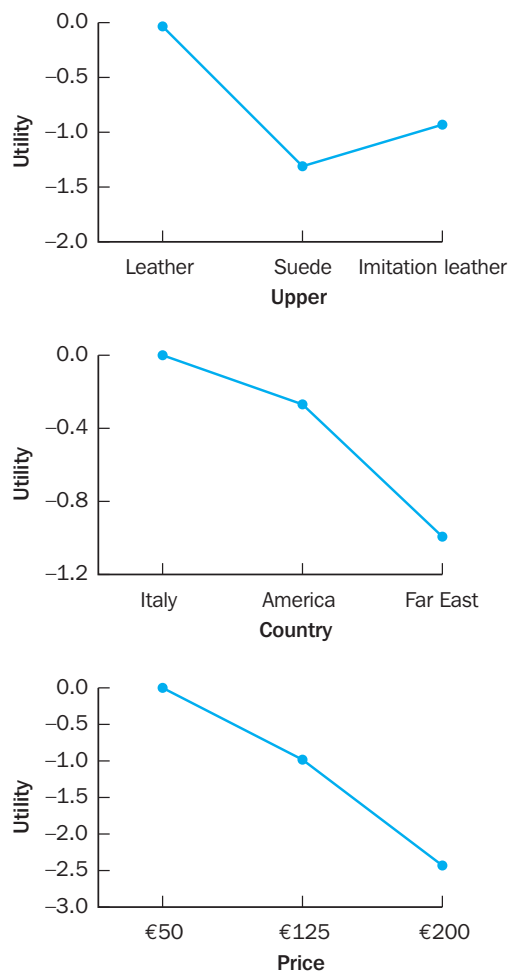
Figure 24.10
Part-worth functions

origin is arbitrary. In terms of relative importance of the attributes, we see that price is number one. Second most important is upper, followed closely by country. Because price is by far the most important attribute for this respondent, this person could be labelled as price sensitive.

Assess the reliability and validity

Several procedures are available for assessing the reliability and validity of conjoint analysis results.[32]

1 The goodness of fit of the estimated model should be evaluated. For example, if dummy variable regression is used, the value of R^2 will indicate the extent to which the model fits the data. Models with poor fit are suspect.
2 Test–retest reliability can be assessed by obtaining a few replicated judgements later in data collection. In other words, at a later stage in the interview, the respondents are asked to evaluate certain selected stimuli again. The two values of these stimuli are then correlated to assess test–retest reliability.
3 The evaluations for the holdout or validation stimuli can be predicted by the estimated part-worth functions. The predicted evaluations can then be correlated with those obtained from the respondents to determine internal validity.
4 If an aggregate-level analysis has been conducted, the estimation sample can be split in several ways and conjoint analysis conducted on each sub-sample. The results can be compared across sub-samples to assess the stability of conjoint analysis solutions.

In running a regression analysis on the data of Table 24.5, an R^2 of 0.934 was obtained, indicating a good fit. The preference ratings for the nine validation profiles were predicted from the utilities reported in Table 24.6. These were correlated with the input ratings for these profiles obtained from the respondent. The correlation coefficient was 0.95, indicating a good predictive ability. This correlation coefficient is significant at $\alpha = 0.05$.

The following example further illustrates an application of conjoint analysis.

e x a m p l e

Fab's fabulous foamy fight[33]

Competition in the clothes detergent market was brewing in Thailand. Superconcentrate detergent was fast becoming the prototype, with a market share of over 26% in the clothes detergent category. Market potential research in Thailand indicated that superconcentrates would continue to grow at around 40% a year. In addition, this category had already dominated other Asian markets such as Taiwan, Hong Kong and Singapore. Consequently, Colgate entered this new line of competition with *Fab Power Plus* with the objective of capturing 4% market share.

The main players in this market were Kao Corporation's *Attack* (14.6%), Lever Brothers' *Breeze Ultra* (2.8%), Lion Corporation's *Pao M. Wash* (1.1%) and Lever's *Omo* (0.4%). Based on qualitative research and secondary data, Colgate assessed the critical factors for the success of superconcentrates. Some of these factors were environmental appeal, hand-wash and machine-wash convenience, superior cleaning abilities, optimum level of suds for hand-wash, and brand name. Research also revealed that no brand had both hand-wash and machine-wash capabilities. *Pao Hand Force* was formulated as the hand-washing brand, and *Pao M. Wash* as the machine-wash version. Lever's *Breezematic* was targeted for machine use.

Therefore, a formula that had both hand- and machine-wash capabilities was desirable. A conjoint study was designed, and these factors varied at either two or three levels. Preference ratings were gathered from respondents, and part-worth functions for the factors were estimated both at the individual and the group level. Results showed that the factor on hand–machine capability had a substantial contribution supporting earlier claims. Based on these findings, *Fab Power Plus* was successfully launched as a brand with both hand- and machine-wash capabilities. ■

Assumptions and limitations of conjoint analysis

Although conjoint analysis is a popular technique, like MDS it carries a number of assumptions and limitations. Conjoint analysis assumes that the important attributes of a product can be identified. Furthermore, it assumes that consumers evaluate the choice alternatives in terms of these attributes and make trade-offs. In situations where image or brand name is important, however, consumers may not evaluate the brands or alternatives in terms of attributes. Even if consumers consider product attributes, the trade-off model may not be a good representation of the choice process. Another limitation is that data collection may be complex, particularly if a large number of attributes are involved and the model must be estimated at the individual level. This problem has been mitigated to some extent by procedures such as interactive or adaptive conjoint analysis and hybrid conjoint analysis. It should also be noted that the part-worth functions are not unique.

Hybrid conjoint analysis

Hybrid conjoint analysis
A form of conjoint analysis that can simplify the data collection task and estimate selected interactions as well as all main effects.

Hybrid conjoint analysis is an attempt to simplify the burdensome data collection task required in traditional conjoint analysis. Each respondent evaluates a large number of profiles, yet usually only simple part-worth functions, without any interaction effects, are estimated. In the simple part-worths or main effects model, the value of a combination is simply the sum of the separate main effects (simple part-worths). In actual practice, two attributes may interact in the sense that the respondent may value the combination more than the average contribution of the separate parts. Hybrid models have been developed to serve two main purposes: (1) to simplify the data collection task by imposing less of a burden on each respondent, and (2) to permit the estimation of selected interactions (at the subgroup level) as well as all main (or simple) effects at the individual level.

In the hybrid approach, the respondents evaluate a limited number, generally no more than nine, conjoint stimuli, such as full profiles. These profiles are drawn from a large master design and different respondents evaluate different sets of profiles so that, over a group of respondents, all the profiles of interest are evaluated. In addition, respondents directly evaluate the relative importance of each attribute and desirability of the levels of each attribute. By combining the direct evaluations with those derived from the evaluations of the conjoint stimuli, it is possible to estimate a model at the aggregate level and still retain some individual differences.[34]

MDS and conjoint analysis are complementary techniques and may be used in combination, as the following example shows.

e x a m p l e

Weeding out the competition[35]

ICI Agricultural Products did not know whether it should lower the price of Fusilade, its herbicide. It knew that it had developed a potent herbicide, but it was not sure that the weedkiller would survive in a price-conscious market. So a survey was designed to assess the relative importance of different attributes in selecting herbicides and to measure and map perceptions of major herbicides on the same attributes. Personal interviews were conducted with 601 soybean and cotton farmers who had at least 120 hectares dedicated to growing these crops and who had used herbicides during the past growing season. First, conjoint analysis was used to determine the relative importance of attributes farmers use when selecting herbicides. Then multidimensional scaling was used to map farmers' perceptions of herbicides. The study showed that price greatly influenced herbicide selections, and respondents were particularly sensitive when costs were more than €10 per hectare. But price was not the only

determinant. Farmers also considered how much weed control the herbicide provided. They were willing to pay higher prices to keep weeds off their land. The study showed that herbicides that failed to control even one of the four most common weeds would have to be very inexpensive to attain a reasonable market share. Fusilade promised good weed control. Furthermore, multidimensional scaling indicated that one of Fusilade's competitors was considered to be expensive. Hence, ICI kept its original pricing plan and did not lower the price of Fusilade. ■

Internet and computer applications

Several computer programs have been developed for conducting MDS analysis using microcomputers and mainframes. The ALSCAL program, available in the mainframe versions of SPSS and SAS, incorporates several different MDS models and can be used for conducting individual- or aggregate-level analysis. Other MDS programs are easily available and widely used. Most are available in both micro-computer and mainframe versions. Visit the following sites for further details, cases and demos.

- **www.newmdsx.com/INDSCAL/indscal.htm** INDSCAL, denoting individual differences scaling, is useful for conducting MDS at the aggregate level. Similarity data are used as input.
- **www.newmdsx.com/MDPREF/mdpref.htm** MDPREF performs internal analysis of preference data. The program develops vector directions for preferences and the configuration of brands or stimuli in a common space.
- **www.newmdsx.com/PREFMAP/prefmap.htm** PREFMAP performs external analysis of preference data. This program uses a known spatial map of brands or stimuli to portray an individual's preference data.
- **www.sawtoothsoftware.com** Sawtooth technologies include choice-based conjoint and multimedia conjoint programs that demonstrate product features rather than just describe them.
- **www.statpac.com** Click on 'Advanced Statistical Analysis' to see an array of multi-variate data analysis programs, including correspondence analysis.
- **www.tigris-software.com** Click on Conjoint Analysis and then TiCon and TiCon Web.

Summary

Multidimensional scaling is used for obtaining spatial representations of respondents' perceptions and preferences. Perceived or psychological relationships among stimuli are represented as geometric relationships among points in a multidimensional space. Formulating the MDS problem requires a specification of the brands or stimuli to be included. The number and nature of brands selected influence the resulting solution. Input data obtained from the respondents can be related to perceptions or preferences. Perception data can be direct or derived. The direct approaches are more common in marketing research.

The selection of an MDS procedure depends on the nature (metric or non-metric) of the input data and whether perceptions or preferences are being scaled. Another determining factor is whether the analysis will be conducted at the individual or aggregate level. The decision about the number of dimensions in which to obtain a

solution should be based on theory, interpretability, elbow criterion and ease-of-use considerations. Labelling of the dimensions is a difficult task that requires subjective judgement. Several guidelines are available for assessing the reliability and validity of MDS solutions. Preference data can be subjected to either internal or external analysis. If the input data are of a qualitative nature, they can be analysed via correspondence analysis. If the attribute-based approaches are used to obtain input data, spatial maps can also be obtained by means of factor or discriminant analysis.

Conjoint analysis is based on the notion that the relative importance that consumers attach to salient attributes, and the utilities they attach to the levels of attributes, can be determined when consumers evaluate brand profiles that are constructed using these attributes and their levels. Formulating the problem requires an identification of the salient attributes and their levels. The pairwise and the full-profile approaches are commonly employed for constructing the stimuli. Statistical designs are available for reducing the number of stimuli in the evaluation task. The input data can be either non-metric (rankings) or metric (ratings). Typically, the dependent variable is preference or intention to buy.

Although other procedures are available for analysing conjoint analysis data, regression using dummy variables is becoming increasingly important. Interpreting the results requires an examination of the part-worth functions and relative importance weights. Several procedures are available for assessing the reliability and validity of conjoint analysis results.

Questions

1 For what purposes are MDS procedures used?

2 Identify two marketing research problems where MDS could be applied. Explain how you would apply MDS in these situations.

3 What is meant by a spatial map?

4 Describe the steps involved in conducting MDS.

5 Describe the direct and derived approaches to obtaining MDS input data.

6 What factors influence the choice of an MDS procedure?

7 What guidelines are used for deciding on the number of dimensions in which to obtain an MDS solution?

8 Describe the ways in which the reliability and validity of MDS solutions can be assessed.

9 What is the difference between internal and external analysis of preference data?

10 What is involved in formulating a conjoint analysis problem?

11 Describe the full profile approach to constructing stimuli in conjoint analysis.

12 Describe the pairwise approach to constructing stimuli in conjoint analysis.

13 How can regression analysis be used for analysing conjoint data?

14 Graphically illustrate what is meant by part-worth functions.

15 What procedures are available for assessing the reliability and validity of conjoint analysis results?

Notes

1 Green, P.E., Carmone Jr, F.J. and Smith, S.M., *Multidimensional Scaling: Concepts and Applications* (Boston, MA: Allyn & Bacon, 1989), 16-17. See also Deogun, N., 'Coke claims dominance in the Mideast and North Africa, but Pepsi disagrees', *Wall Street Journal* (3 March 1998), A4.

2 Barrett, P., Burr, P. and Dolman, N., 'The conjoint path over the cultural divide', *Marketing Week* (14 October 1997).

3 For a review of MDS studies in marketing, see Carroll, J.D. and Green, P.E., 'Psychometric methods in marketing research: Part II, Multidimensional scaling', *Journal of Marketing Research* 34 (February 1997), 193–204; Cooper, L.G., 'A review of multidimensional scaling in marketing research', *Applied Psychological Measurement* 7 (Fall 1983), 427–50.

4 An excellent discussion of the various aspects of MDS may be found in Davison, M.L., *Multidimensional Scaling* (Melbourne: Krieger Publishing, 1992).

5 The data are commonly treated as symmetric. For an asymmetric approach, see Desarbo, W.S. and Manrai, A.K., 'A new multidimensional scaling methodology for the analysis of asymmetric proximity data in marketing research', *Marketing Science* 11(1) (Winter 1992), 1–20. For other approaches to MDS data, see Bijmolt, T.H.A. and Wedel, M., 'The effects of alternative methods of collecting similarity data for multidimensional scaling', *International Journal of Research in Marketing* 12(4) (November 1995), 363–71.

6 See Borg, I. and Groenen, P.J., *Modern Multidimensional Scaling Theory and Applications* (New York: Springer-Verlag, 1996); Malhotra, N.K., Jain, A.K. and Pinson, C., 'The robustness of MDS configurations in the case of incomplete data', *Journal of Marketing Research* 25 (February 1988), 95–102; Steenkamp, J.-B.E.M. and Van Trijp, H.C.M., 'Task experience and validity in perceptual mapping: a comparison of two consumer-adaptive techniques', *International Journal of Research in Marketing* 13(3) (July 1996) 265–76.

7 See Cox, T., *Multidimensional Scaling* (New York: Routledge, Chapman & Hall, 1994).

8 Kruskal's stress is probably the most commonly used measure for lack of fit. See Borg, I. and Groenen, P.J., *Modern Multidimensional Scaling Theory and Applications* (New York: Springer-Verlag, 1996). For the original article, see Kruskal, J.B., 'Multidimensional scaling by optimising goodness of fit to a nonmetric hypothesis', *Psychometrika* 29 (March 1964), 1–27.

9 Carroll, J.D. and Green, P.E., 'Psychometric methods in marketing research: Part II, Multidimensional scaling', *Journal of Marketing Research* 34 (February 1997), 193–204; Malhotra, N.K., 'Validity and structural reliability of multidimensional scaling', *Journal of Marketing Research* 24 (May 1987), 164–73.

10 For an examination of the reliability and validity of MDS solutions, see Steenkamp, J.-B.E.M., Van Trijp, H.C.M. and Ten Berge, J.M.F., 'Perceptual mapping based on idiosyncratic sets of attributes', *Journal of Marketing Research* 31 (February 1994), 15–27.

11 Hair Jr, J.E., Anderson, R.E., Tatham, R.L. and Black, W.C., *Multivariate Data Analysis with Readings*, 5th edn (Englewood Cliffs, NJ: Prentice Hall, 1999), 484–555.

12 See, for example, DeSarbo, W.S., Hoffman, D. and Rangaswamy, A., 'A parametric multidimensional unfolding procedure for incomplete nonmetric preference/choice set data marketing research', *Journal of Marketing Research* 34(4) (November 1997) 499–516; Mackay, D.B., Easley, R.F. and

Zinnes, J.L., 'A single ideal point model for market structure analysis', *Journal of Marketing Research* 32(4) (November 1995), 433–43.

13 For an application of correspondence analysis, see Math, J.J., Candel, M. and Maris, E., 'Perceptual analysis of two-way, two-mode frequency data: probability matrix decomposition and two alternatives', *International Journal of Research in Marketing* 14(4) (October 1997), 321–39; Green, P.E. and Krieger, A.M., 'A simple approach to target market advertising strategy', *Journal of the Market Research Society* 35 (April 1993), 161–70.

14 Kara, A., Kaynak, E. and Kucukemiroglu, O., 'Positioning of fast food outlets in two regions of north America: a comparative study using correspondence analysis', *Journal of Professional Services Marketing* 14(2) (1996), 99–119; O'Brien, T.V., 'Correspondence analysis' *Marketing Research: A Magazine of Management and Applications* 5(4) (Fall 1993): 54–6; Green, P.E., Carmone Jr, F.J. and Smith, S.M., *Multidimensional Scaling: Concepts and Applications* (Boston, MA: Allyn & Bacon, 1989), 16–17.

15 Whitlark, D.B. and Smith, S.M., 'Using correspondence analysis to map relationships', *Marketing Research: A Magazine of Management and Applications* 13(4) (Fall 2001), 22–7.

16 See Blasius, J. and Greenacre, M.J., *Visualization of Categorical Data* (Academic Press, 1998); Greenacre, M.J., *Correspondence Analysis in Practice* (New York: Academic Press, 1993); Greenacre, M.J., 'The Carroll–Green–Schaffer scaling in correspondence analysis: a theoretical and empirical appraisal', *Journal of Marketing Research* 26 (August 1989), 358–65; Greenacre, M.J., *Theory and Applications of Correspondence Analyses* (New York: Academic Press, 1984); Hoffman, D.L. and Franke, G.R., 'Correspondence analysis: graphical representation of categorical data in marketing research', *Journal of Marketing Research* 23 (August 1986), 213–27.

17 For the use of factor analysis in constructing spatial maps, see Hasson, L., 'Monitoring social change', *Journal of the Market Research Society* 37 (January 1995), 69–80.

18 Hauser, J.R. and Koppelman, F.S., 'Alternative perceptual mapping techniques: relative accuracy and usefulness', *Journal of Marketing Research* 16 (November 1979), 495–506. Hauser and Koppelman conclude that factor analysis is superior to discriminant analysis. See also Borg, I. and Groenen, P.J., *Modern Multidimensional Scaling Theory and Applications* (New York: Springer-Verlag, 1996).

19 For applications and issues in conjoint analysis, see Naude, P. and Buttle, F., 'Assessing relationship quality', *Industrial Marketing Management* 29 (2000), 351–61; Srinivasan, V. and Park, C.S., 'Surprising robustness of the self-explicated approach to customer preference structure measurement', *Journal of Marketing Research* 34 (May 1997), 286–91; Green, P.E. and Krieger, A.M., 'Segmenting markets with conjoint analysis', *Journal of Marketing* 55 (October 1991), 20–31.

20 Danaher, F.J., 'Using conjoint analysis to determine the relative importance of service attributes measured in customer satisfaction surveys', *Journal of Retailing* 73(2) (Summer 1997), 235–60.

21 For an overview of conjoint analysis in marketing, see Carroll, J.D. and Green, P.E., 'Psychometric methods in marketing research: Part I, Conjoint analysis', *Journal of Marketing Research* 32 (November 1995), 385–91; Green, P.E. and Srinivasan, V., 'Conjoint analysis in marketing: new developments with implications for research and practice', *Journal of*

Marketing 54 (October 1990), 3–19; Green, P.E. and Srinivasan, V., 'Conjoint analysis in consumer research: issues and outlook', *Journal of Consumer Research* 5 (September 1978), 102–23.

22 Thomas-Miller, J., Ogden, J.R. and Latshaw, C.A., 'Using trade-off analysis to determine value–price sensitivity of custom calling features', *American Business Review* 16(1) (January 1998), 8–13.

23 Wittink, D.R., Vriens, M. and Burhenne, W., 'Commercial uses of conjoint analysis in Europe: results and critical reflections', *International Journal of Research in Marketing* 11(1) (January 1994), 41–52; Wittink, D.R. and Cattin, P., 'Commercial use of conjoint analysis: an update', *Journal of Marketing* 53 (July 1989), 91–7. For using conjoint analysis to measure price sensitivity, see 'Multi-stage conjoint methods to measure price sensitivity', *Sawtooth News* 10 (Winter 1994–1995), 5–6.

24 Loosschilder, G.H., Rosbergen, E., Vriens, M. and Wittink, D.R., 'Pictorial stimuli in conjoint analysis – to support product styling decisions', *Journal of the Market Research Society* 37 (January 1995), 17–34.

25 See Carroll, J.D. and Green, P.E., 'Psychometric methods in marketing research: Part I, Conjoint analysis', *Journal of Marketing Research* 32 (November 1995), 385–91; Kuhfeld, W.F., Tobias, R.D. and Garratt, M., 'Efficient experimental designs with marketing applications', *Journal of Marketing Research* 31 (November 1994), 545–57; Addleman, S., 'Orthogonal main-effect plans for asymmetrical factorial experiments', *Technometrics* 4 (February 1962), 21–36; Green, P.E., 'On the design of choice experiments involving multifactor alternatives', *Journal of Consumer Research* 1 (September 1974), 61–8.

26 More complex conjoint designs are also possible. See Oppewal, H., Louviere, J.J. and Timmermans, H.J.P., 'Modeling hierarchical conjoint processes with integrated choice experiments', *Journal of Marketing Research* 31 (February 1994), 15–27.

27 Carroll, J.D. and Green, P.E., 'Psychometric methods in marketing research: Part I, Conjoint analysis', *Journal of Marketing Research* 32 (November 1995), 385–91; Jain, A.K., Acito, F., Malhotra, N.K. and Mahajan V., 'A comparison of the internal validity of alternative parameter estimation methods in decompositional multiattribute preference models', *Journal of Marketing Research* (August 1979), 313–22.

28 Oppewal, H., Timmermans, H.J. and Louviere, J.J., 'Modeling the effect of shopping center size and store variety in consumer choice behaviour', *Environment and Planning* 29(6) (June 1997), 1073–90.

29 Christen, M., Gupta, S., Porter, J.C., Staelin, R. and Wittink, D.R., 'Using market-level data to understand the promotion effects in a nonlinear model', *Journal of Marketing Research* 34(3) (August 1997), 322–34; Moore, W.L., 'Levels of aggregation in conjoint analysis: an empirical comparison', *Journal of Marketing Research* 17 (November 1980), 516–23. See also Brice, R., 'Conjoint analysis: a review of conjoint paradigms and discussion of the outstanding design issues', *Marketing and Research Today* 25(4) (November 1997), 260–6.

30 Carroll, J.D. and Green, P.E., 'Psychometric methods in marketing research: Part I, Conjoint analysis', *Journal of Marketing Research* 32 (November 1995), 385–91; Carmone, E.J. and Green, P.E., 'Model mis-specification in multiattribute parameter estimation', *Journal of Marketing Research* 18 (February 1981), 87–93.

31 For an application of conjoint analysis using OLS regression see Ostrom, A. and Iacobucci, D., 'Consumer trade-offs and the evaluation of services', *Journal of Marketing* 59 (January 1995), 17–28; Danaher, P.J., 'Using conjoint analysis to determine the relative importance of service attributes measured in customer satisfaction surveys', *Journal of Retailing* 73(2) (Summer 1997), 235–60.

32 Carroll, J.D. and Green, P.E., 'Psychometric methods in marketing research: Part I, Conjoint analysis', *Journal of Marketing Research* 32 (November 1995), 385–91; Malhotra, N.K., 'Structural reliability and stability of nonmetric conjoint analysis', *Journal of Marketing Research* 19 (May 1982), 199–207; Leigh, T.W., MacKay, D.B. and Summers, J.O., 'Reliability and validity of conjoint analysis and self-explicated weights: a comparison', *Journal of Marketing Research* 21 (November 1984), 456–62; and Segal, M.N., 'Reliability of conjoint analysis: contrasting data collection procedures', *Journal of Marketing Research* 19 (February 1982), 139–43.

33 Butler, D., 'Thai superconcentrates foam', *Advertising Age* (18 January 1993).

34 Hu, C. and Hiemstra, S.J., 'Hybrid conjoint analysis as a research technique to measure meeting planners' preferences in hotel selection', *Journal of Travel Research* 35(2) (Fall 1996), 62–9; Green, P.E. and Krieger, A.M., 'Individualized hybrid models for conjoint analysis', *Management Science* 42(6) (June 1996), 850–67; Green, P.E., 'Hybrid models for conjoint analysis: an expository review', *Journal of Marketing Research* 21 (May 1984), 155–69.

35 Pfeifer, S., Gain, B. and Walsh, K., 'Managing specialities: how to grow when prices stall', *Chemical Week* 159(47) (10 December 1997), 30–4; Schneidman, D., 'Research method designed to determine price for new products, line extensions', *Marketing News* (23 October 1987), 11.

Report preparation and presentation

Objectives

After reading this chapter, you should be able to:

1 discuss the basic requirements of report preparation, including report format, report writing, graphs and tables;

2 discuss the nature and scope of the oral presentation;

3 describe the approach to the marketing research report from the client's perspective;

4 explain the reason for follow-up with the client and describe the assistance that should be given to the client and the evaluation of the research project;

5 understand the report preparation and presentation process in international marketing research;

6 identify the ethical issues related to the interpretation and reporting of the research process and findings to the client and the use of these results by the client.

Managers should find reports easy to understand, be confident in the findings, and be clear about the action they should take, based on the researcher's approach, insight and integrity.

Overview

This chapter describes the importance of report preparation and presentation and outlines the process of producing written and oral presentations. We provide guidelines for report preparation, including report writing and preparing tables and graphs, and we discuss oral presentation of the report. Research follow-up, including assisting the client and evaluating the research process, is described. The special considerations for report preparation and presentation in international marketing research are discussed, and relevant ethical issues are identified. We begin with an example of the potential array of reports that can emerge from a marketing research project that is primarily presenting quantitative findings. This is followed with an example of a very creative means to present findings to a study that is primarily presenting qualitative findings. Both examples illustrate the different styles and means to present marketing research reports.

e x a m p l e

GlobalCash Project

Report preparation and presentation

In the GlobalCash Project, funding for the research came from 15 pan-European banks and trade associations throughout Europe, such as the Dutch and the UK Association of Corporate Treasurers. With different sponsoring organisations, a number of formats of report presentation were made. As 19 European countries were surveyed, reports were prepared for each country. There were additional reports for the Nordic/Scandinavia region that combined the results from Denmark, Finland, Norway and Sweden, plus an aggregated report for the whole of Europe. Given the importance of the introduction of European Monetary Union, interim reports (based upon 250 and then 500 responses) were presented as the total dataset was being compiled.

The aggregated European report, compiled from data from 1994, 1996, 1998 and 2000, was presented in a written format. As well as the statistical findings, interpretation of the statistics was presented. For individual countries, statistics and interpretations were presented in electronic formats, available on the Internet in a password-protected location. As well as receiving the written and electronic findings, each of the sponsoring banks made an oral presentation. Some of the banks, such as Citibank and the Chase Manhattan Bank, used video-conferencing to allow executives in the United States to participate in the oral presentation. Each oral presentation was tailored to the individual sponsoring bank. They wished to understand the impact of the findings upon their future strategy in the European market.

For the sponsoring trade associations, a written report for their country was presented. This report was shorter than the bank report in that individual bank ratings and performance were omitted. Again, an oral presentation was made to an association meeting which could number up to 300 delegates from a wide array of companies and banks within individual countries.

Finally a much shorter summary of the statistics was written and sent to all questionnaire respondents. ■

e x a m p l e

Meet Matthias, Stephanie, Seb, Justine and Stan [1]

Allied Domecq Spirits and Wines (ADSW) commissioned an innovative study to understand the factors of brand adoption of young emergent adult drinkers. The innovative approach to gathering data continued into communicating the findings. ADSW managers were invited to spend a day of 'discovery'. The day started by holding breakout sessions to gauge current assumptions about young drinkers. They then 'met' a set of fictional characters that represented the adult emergent drinker generation that had been the focus of the study. The idea behind this was to allow ADSW managers to be able to visualise their consumers when developing new product development or communications strategies. Personalities were created and brought to life using actors. In France, for example, the clients were able to meet Matthias, Stephanie, Seb, Justine and Stan. These five characters symbolised the richness and the diversity of the generation. Each of the characters engaged with the manager audience via dialogue, discussing for example their lifestyle, behaviours, what's 'in' or 'out', values, concerns and expectations for the future, as well as their current attitudes towards alcohol. In

addition, the audience was presented with workshop 'souvenirs', notebooks with pictures and biographies of the character types where they could take notes during the presentation. The effect was immediate; with a bar as a cue, managers were able to step into a new world and easily meet and interact with their consumers. Moreover, their 'consumers' were eager to explain what was and wasn't important to them. This multi-media/multi-layered presentation of findings allowed information to be assimilated both visually, audibly and kinesthetically. ■

Importance of the report and presentation

For the following reasons, the report and its presentation are important parts of the marketing research project:

1 They are the tangible products of the research effort. After the project is complete and management has made the decision, there is little documentary evidence of the project other than the written report. The report serves as a historical record of the project.
2 Management decisions are guided by the report and the presentation. If the first five steps in the project are carefully conducted but inadequate attention is paid to the sixth step, the value of the project to management will be greatly diminished.
3 The involvement of many marketing managers in the project is limited to the written report and the oral presentation. These managers evaluate the quality of the entire project on the quality of the report and presentation.
4 Management's decision to undertake marketing research in the future or to use the particular research supplier again will be influenced by the perceived usefulness of the report and the presentation.

Preparation and presentation process

Figure 25.1 illustrates report preparation and presentation.
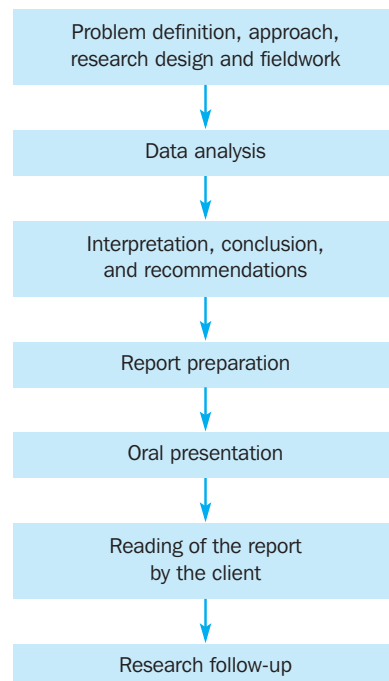
Figure 25.1
The report preparation and presentation process

Problem definition, approach, research design and fieldwork

↓

Data analysis

↓

Interpretation, conclusion, and recommendations

↓

Report preparation

↓

Oral presentation

↓

Reading of the report by the client

↓

Research follow-up

The process begins by interpreting the results of data analysis in the light of the marketing research problem, approach, research design and fieldwork. Instead of merely summarising the quantitative and/or qualitative analyses, the researcher should present the findings in such a way that they can be used directly as input into decision-making. Wherever appropriate, conclusions should be drawn and recommendations made. The researcher should aim to make their recommendations actionable. Before writing the report, the researcher should discuss the major findings, conclusions and recommendations with the key decision-makers. These discussions play a major role in ensuring that the report meets the client's needs and is ultimately accepted. These discussions should confirm specific dates for the delivery of the written report and other data.

The entire marketing research project should be summarised in a single written report or in several reports addressed to different readers. Generally, an oral presentation supplements the written documents. The client should be given an opportunity to read the report. After that, the researcher should take necessary follow-up actions. The researcher should assist the client in understanding the report, help in interpretations of the findings that can affect their implementation, offer to undertake further research and reflect upon the research process to evaluate its overall worth.

Report preparation

Researchers differ in the way they prepare a research report. The personality, background, expertise and responsibility of the researcher, along with the marketing decision-maker to whom the report is addressed, interact to give each report a unique character. Yet there are guidelines for formatting and writing reports and designing tables and graphs.[2]

Report format

Report formats are likely to vary with the researcher or the marketing research firm conducting the project, the client for whom the project is being conducted, and the nature of the project itself. Hence, the following is intended as a guideline from which the researcher can develop a format for the research project at hand. Most research reports include the following elements:

1 Submission letter

2 Title page

3 Table of contents
    (a) Main sections
    (b) List of tables
    (c) List of graphs
    (d) List of appendices
    (e) List of exhibits

4 Executive summary
    (a) Summary of prime objectives
    (b) Major findings
    (c) Conclusions and recommendations

5 Problem definition
    (a) Background to the problem
    (b) Statement of the marketing problem
    (c) Statement of the research objectives – information needs

6 **Approach to the problem and research design**
   (a) Type of research design
   (b) Data collection from secondary sources
   (c) Data collection from primary sources

7 **Data analysis**
   (a) Research design
   (b) Plan of data analysis and means of interpreting results

8 **Results**

9 **Conclusions and recommendations**

10 **Limitations and caveats**

11 **Appendices**
   (a) Letter of authorisation
   (b) Questionnaire development and pre-testing
   (c) Questionnaires, forms, interview guides
   (d) Sampling techniques, including error and confidence levels
   (e) Fieldwork
   (f) Lists including contact individuals and organisations

This format closely follows the earlier steps of the marketing research process. The results may be presented in several chapters of the report. For example, in a national survey, data analysis may be conducted for the overall sample and then the data for each geographic region may be analysed separately. If so, the results from each analysis may be presented in a separate chapter.

Submission letter. A formal report generally contains a letter of submission that delivers the report to the client and summarises the researcher's overall experience with the project, without mentioning the findings. The letter should also identify the need for further action on the part of the client, such as implementation of the findings or further research that should be undertaken.

Title page. The title page should include the title of the report, information (name, address and telephone number) about the researcher or organisation conducting the research, the name of the client for whom the report was prepared, and the date of release. The title should encapsulate the nature of the project with a tone that is meaningful to the target managers, not one of technical 'research-speak'.

Table of contents. The table of contents should list the topics covered and the appropriate page numbers. In most reports, only the major headings and subheadings are included. The table of contents is followed by a list of tables, a list of graphs, a list of appendices and a list of exhibits.

Executive summary. The executive summary is an extremely important part of the report, because this is often the only portion of the report that executives read. The summary should concisely describe the problem, approach and research design that were adopted. A summary section should be devoted to the major results, conclusions and recommendations. The executive summary should be written after the rest of the report.

Problem definition. The problem definition section of the report gives the background to the problem. This part summarises elements of the marketing and research problem diagnosis. Key elements of any discussions with decision makers, industry

experts and initial secondary data analyses are presented. Having set this context for the whole project, a clear statement of the management decision problem(s) and the marketing research problem(s) should be presented.

Approach to the problem and research design. The approach to the problem section should discuss the broad approach that was adopted in addressing the problem. This section should summarise the theoretical foundations that guided the research, any analytical models formulated, research questions, hypotheses, and the factors that influenced the research design. The research design should specify the details of how the research was conducted, preferably with a graphical presentation of the stages undertaken, showing the relationships between stages. This should detail the methods undertaken in the data collection from secondary and primary sources. These topics should be presented in a non-technical, easy-to-understand manner. The technical details should be included in an appendix. This section of the report should justify the specific methods selected.

Data analysis. The section on data analysis, be it quantitative or qualitative, should describe the plan of data analysis and justify the data analysis strategy and techniques used. The techniques used for analysis should be described in simple, non-technical terms, with examples to guide the reader through the interpretations.

Results. The results section is normally the longest part of the report and may entail several chapters. It may be presented in any of the following ways.

1 *Forms of analysis.* For example, in a health care marketing survey of hospitals, the results were presented in four chapters. One chapter presented the overall results, another examined the differences between geographical regions, a third presented the differences between for-profit and non-profit hospitals, and a fourth presented the differences according to bed capacity. Often, results are presented not only at the aggregate level but also at the subgroup (market segment, geographical area, etc.) level.
2 *Forms of data collection.* For example, a study may contain significant elements of secondary data collection and analyses, a series of focus group interviews and a survey. The results in such circumstances may be best presented by drawing conclusions from one method before moving on to another method. The conclusions derived from focus groups, for example, may need to be established to show the link to a sample design and questions used in a survey.
3 *Objectives.* There may be a series of research objectives whose fulfilment may incorporate a variety of data collection methods and levels of analysis. In these circumstances the results combine methods and levels of analyses to show connections and to develop and illustrate emerging issues.

The results should be organised in a coherent and logical way. Choosing whether to present by forms of analysis, forms of data collection or objectives helps to build that coherence and logic. The presentation of the results should be geared directly to the components of the marketing research problem and the information needs that were identified. The nature of the information needs and characteristics of the recipients of the report ultimately determine the best way to present results.

Conclusions and recommendations. Presenting a mere summary of the quantitative or qualitative findings is not enough. The researcher should interpret the results in light of the problem being addressed to arrive at major conclusions. Based on the

results and conclusions, the researcher may make recommendations to the decision-makers. Sometimes, marketing researchers are not asked to make recommendations because they research only one area and do not understand the bigger picture at the client firm. The researcher may not have been fully involved in the diagnosis of the marketing and research problems, in which case their interpretations may not fit into the context that the marketer understands.

In any research project there are many approaches that can be taken to analyse the data. This can result in a potential over-abundance of data (quantitative and/or qualitative) and distilling the 'meaning' from the data and presenting this in a clear report can result in much of the original meaning or richness being lost.[3] To maintain the meaning or richness, the researcher should strive to understand the nature of the decision-making process that is being supported. Only then can sound interpretations of the collected data be made.

Limitations and caveats. All marketing research projects have limitations caused by time, budget and other organisational constraints. Furthermore, the research design adopted may be limited in terms of the various types of errors, and some of these may be serious enough to warrant discussion. This section should be written with great care and a balanced perspective. On the one hand, the researcher must make sure that management does not rely too heavily on the results or use them for unintended purposes, such as projecting them to unintended populations. On the other hand, this section should not erode their confidence in the research or unduly minimise its importance.

Appendices. At the end of the report, documents can be compiled that may be used by different readers to help them to understand characteristics of the research project in more detail. These should include the letter of authorisation to conduct the research; this authorisation could include the agreed research proposal. Details that relate to individual techniques should be included relating to questionnaires, interview guides, sampling and fieldwork activities. The final part of the appendix should include lists of contacts, references used and further sources of reference.

Report writing

Readers. A report should be written for a specific reader or readers: the marketing decision-makers who will use the results. The report should take into account the readers' technical sophistication and interest in the project as well as the circumstances under which they will read the report and how they will use it.[4]

Technical jargon should be avoided. As expressed by one expert, 'The readers of your reports are busy people; and very few of them can balance a research report, a cup of coffee, and a dictionary at one time.'[5] Instead of technical terms such as maximum likelihood, heteroscedasticity, and non-parametric, use descriptive explanations. If some technical terms cannot be avoided, briefly define them in an appendix. When it comes to marketing research, decision-makers would rather live with a problem they cannot solve than accept a solution they cannot understand.

Often the researcher must cater to the needs of several audiences with different levels of technical sophistication and interest in the project. Such conflicting needs may be met by including different sections in the report for different readers or separate reports entirely.

Easy to follow. The report should be easy to follow.[6] It should be structured logically and written clearly. The material, particularly the body of the report, should be struc-

tured in a logical manner so that the reader can easily see the inherent connections and linkages. Headings should be used for different topics and subheadings for subtopics.

A logical organisation also leads to a coherent report. Clarity can be enhanced by using well-constructed sentences that are short and to the point. The words used should express precisely what the researcher wants to communicate. Difficult words, slang and clichés should be avoided. An excellent check on the clarity of a report is to have two or three people who are unfamiliar with the project read it and offer critical comments. Several revisions of the report may be needed before the final document emerges.

Presentable and professional appearance. The look of a report is important. The report should be professionally reproduced with quality paper, typing and binding. The typography should be varied. Variation in type size and skilful use of white space can greatly contribute to the appearance and readability of the report. However, a balance should be sought with styles of variation. Too much variation can lead to confusion; variation is only useful if it aids understanding.

Objective. Objectivity is a virtue that should guide report writing. Researchers can become so fascinated with their project that they overlook their 'scientific' role. The report should accurately present the research design, results and conclusions of the project, without slanting the findings to conform to the expectations of management. Decision-makers are unlikely to receive with enthusiasm a report that reflects unfavourably on their judgement or actions. Yet the researcher must have the courage to present and defend the results objectively.

Reinforce text with tables and graphs. It is important to reinforce key information in the text with tables, graphs, pictures, maps and other visual devices. Visual aids can greatly facilitate communication and add to the clarity and impact of the report. Guidelines for tabular and graphical presentation are discussed later.

Reinforce tables and graphs with text. Conversely it is important to illustrate tables and graphs with verbatim quotes from questionnaires and interviews. Quotes can bring to life the meaning in tables and graphs and, used carefully, can make the reading of the report far more interesting than a solid body of statistics.

Terse. A report should be terse and concise. Anything unnecessary should be omitted. If too much information is included, important points may be lost. Avoid lengthy discussions of common procedures. Yet brevity should not be achieved at the expense of completeness.

Guidelines for tables

Statistical tables are a vital part of the report and deserve special attention. We illustrate the guidelines for tables using data from the GlobalCash study. Table 25.1 presents the findings from two questions. The rows in Table 25.1 show how European companies plan to change their relationships with their existing banks once EMU is introduced. The columns in Table 25.1 summarise the number of relationships companies have with banks in their home countries. This number of relationships is termed the 'sourcing strategy', i.e. a *single* strategy means all transactions go through one bank, a *dual* strategy means that transactions go through two banks and a *multiple* strategy means that transactions go through more than two banks, in some instances up to 70 banks for some Italian companies.

(1a) **Table 25.1** **How the introduction of EMU will affect company relationships with their banks – analysed by sourcing strategy[a]**

(1b)

(2a) (3a) (5a)

How EMU will affect cash management banking relationships	Totals	Sourcing strategy[a]		
		Single	Dual	Multiple
Totals	896	155	179	562
Existing country relationships maintained	46%	41%	45%	48%
Fewer banks used across euro zone	33%	31%	30%	35%
One major bank to co-ordinate	33%	43%	39%	29%
Fewer banks used in each country	22%	15%	17%	25%

(4a)

(5b)

(5c) [a] Sourcing strategy relates to the number of banks used for domestic business

(6a) Source: GlobalCash-Europe98, Statistical Report for Europe, p. 137

The numbers in parentheses in the following paragraphs refer to the numbered sections of the table.

Title and number. Every table should have a number (1a) and title (1b). The title should be brief yet clearly descriptive of the information provided. Arabic numbers are used to identify tables so that they can be referenced in the text.[7]

Arrangement of data items. The arrangement of data items in a table should emphasise the most significant aspect of the data. For example, when the data pertain to time, the items should be arranged by appropriate time period. When order of magnitude is most important, the data items should be arranged in that order (2a). If ease of locating items is critical, an alphabetical arrangement is most appropriate.

Basis of measurement. The basis or unit of measurement should be clearly stated (3a). In Table 25.1, the totals sample size is shown and the sub-sample sizes of the different sourcing strategies. The main body of data is shown in percentages. The % signs would normally be removed, with a note to tell the reader that the main body is based upon column percentages or row percentages, or percentages related to the total sample size.

Leaders, rulings and spaces. The reader's eye should be guided to be able to read across the table clearly. This can be achieved with ruled lines (4a), alternate shaded rows, or white spaces with dotted lines leading from the row headings to the data.

Explanations and comments: headings, stubs and footnotes. Explanations and comments clarifying the table can be provided in the form of captions, stubs and footnotes. Designations placed over the vertical columns are called headings (5a). Designations placed in the left-hand column are called stubs (5b). Information that cannot be incorporated in the table should be explained by footnotes (5c). Letters or symbols should be used for footnotes rather than numbers. The footnotes that are part of the original source should come after the main table, but before the source note.

Sources of the data. If the data contained in the table are secondary, the source of data should be cited (6a).

Guidelines for graphs

As a general rule, graphical aids should be employed whenever practical. Graphical display of information can effectively complement the text and tables to enhance clarity of communication and impact.[8] As the saying goes, a picture is worth a thousand words. The guidelines for preparing graphs are similar to those for tables. Therefore, this section focuses on the different types of graphical aids.[9] We illustrate several of these using the GlobalCash data from Table 25.1.

Geographic and other maps. Geographic and other maps, such as product positioning maps, can communicate relative location and other comparative information. Geographic maps form the bases of presentations in geodemographic analyses as discussed in Chapter 5. The maps used in geodemographic analyses can portray customer locations and types, potential consumers, location of competitors, road networks to show consumer flows and other facilities that may attract consumers to certain locations.

Pie chart
A round chart divided into sections.

Round or pie charts. In a **pie chart**, the area of each section, as a percentage of the total area of the circle, reflects the percentage associated with the value of a specific variable. Pie charts are very useful in presenting simple relative frequencies in numbers or percentages. A pie chart is not useful for displaying relationships over time or relationships among several variables. As a general guideline, a pie chart should not require more than seven sections.[10] Figure 25.2 shows a pie chart for the sourcing strategies of European companies. Great care must be taken with 3D pie charts as the relative sizes of the pie segments become distorted.

Line chart
A chart that connects a series of data points using continuous lines.

Line charts. A **line chart** connects a series of data points using continuous lines. This is an attractive way of illustrating trends and changes over time. Several series can be compared on the same chart, and forecasts, interpolations and extrapolations can be shown. If several series are displayed simultaneously, each line should have a distinctive colour or form (see Figure 25.3).[11]

Bar chart
A chart that displays data in bars positioned horizontally or vertically.

Histograms and bar charts. A **bar chart** displays data in various bars that may be positioned horizontally or vertically. Bar charts can be used to present absolute and

Figure 25.2
Pie chart that shows the percentage of European companies with different sourcing strategies for domestic banks

Single – 17%

Dual – 20%

Multiple – 63%
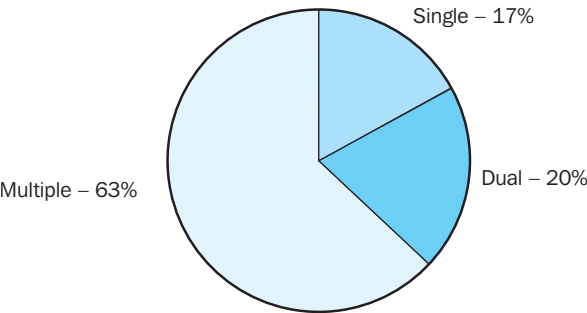
Figure 25.3
Line chart of how the introduction of EMU will affect cash management banking relationships
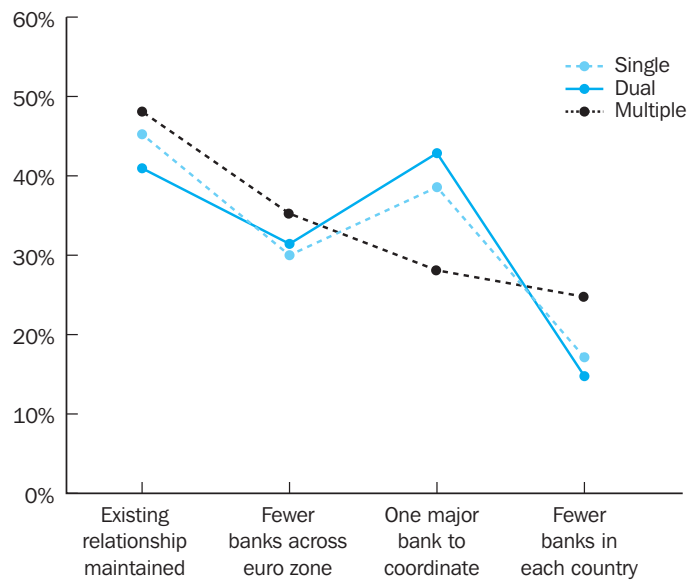
Histogram
A vertical bar chart in which the height of the bars represents the relative or cumulative frequency of occurrence.

relative magnitudes, differences and change. A **histogram** is a vertical bar chart in which the height of the bars represents the relative or cumulative frequency of occurrence of a specific variable (see Figure 25.4). Other variations on the basic bar chart include the stacked bar chart (Figure 25.5) and the cluster bar chart (Figure 25.6). Stacked and cluster bar charts can work well with a few data items presented, to qualitatively represent differences between groups. As noted with pie charts, 3D charts should be used with great caution as they can distort the message and confuse an audience. Most graphics packages have a great array of 3D options; however, there are few circumstances where they can be used to present data in a clear and unbiased manner.

Schematic figures and flowcharts. Schematic figures and flowcharts take on a number of different forms. They can be used to display the steps or components of a process, as in Figure 25.1. They can also be of great value in presenting qualitative

Figure 25.4
Bar chart that shows the percentage of company types that planned to use one major bank to coordinate cash transactions throughout Europe
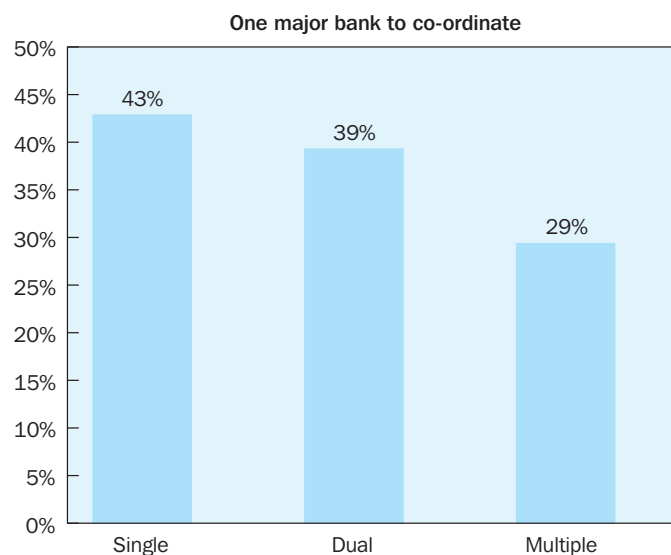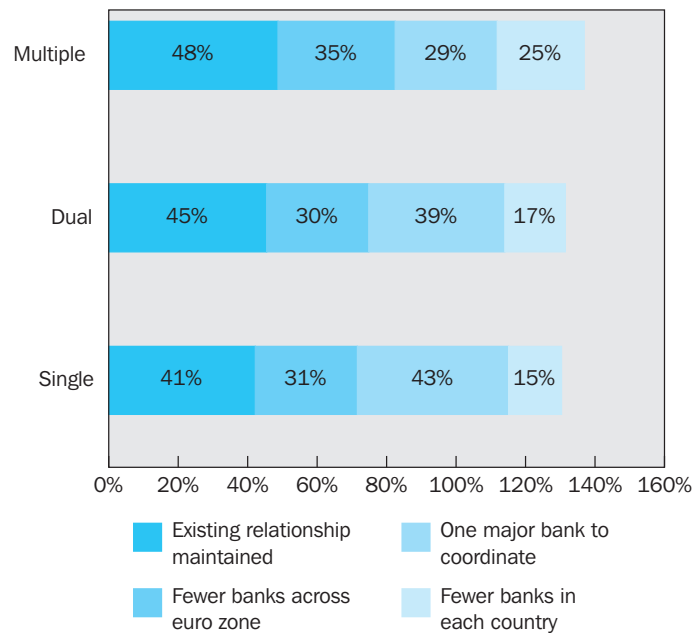
Figure 25.5
Stacked bar chart

data analyses by representing the nature and interconnection of ideas that have been uncovered (see Chapter 9). Another useful form of these charts is classification diagrams. Examples of classification charts for classifying secondary data were provided in Chapter 4 (Figures 4.1 to 4.4). An example of a flowchart for questionnaire design was given in Chapter 13 (Figure 13.3).[12]
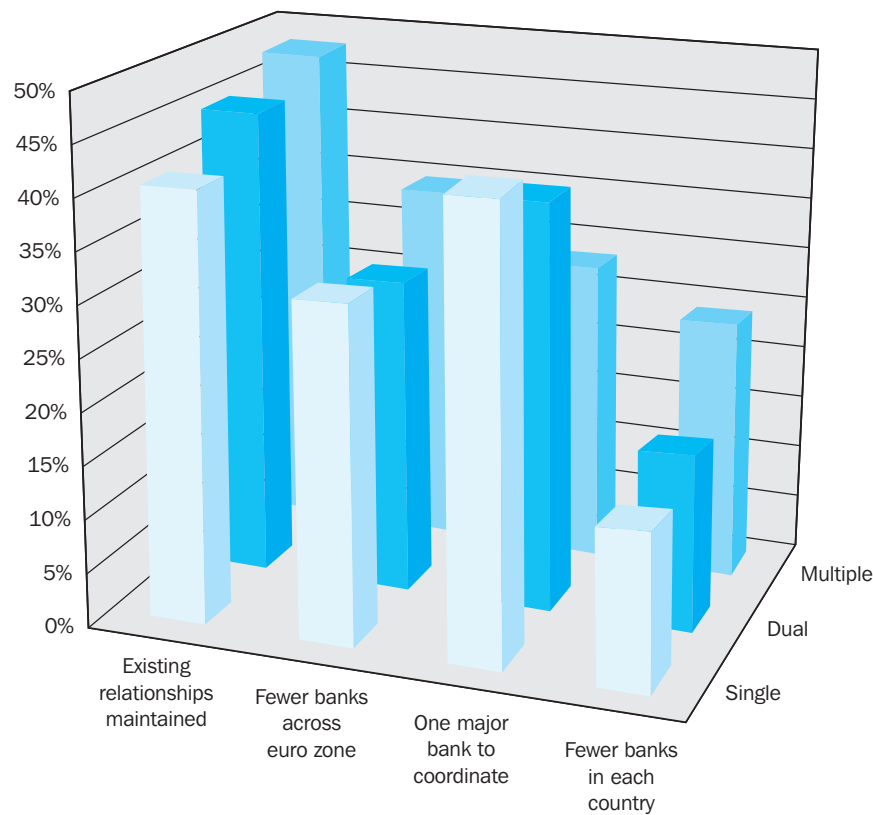
Figure 25.6
3D cluster bar

Oral presentation

The entire marketing research project should be presented to the management of the client firm. This presentation will help management understand and accept the written report. Any preliminary questions that the management may have can be addressed in the presentation. Because many executives form their first and lasting impressions about the project based on the oral presentation, its importance cannot be overemphasised.[13]

The key to an effective presentation is preparation. A written script or detailed outline should be prepared following the format of the written report. The presentation must be geared to the audience. For this purpose, the researcher should determine the backgrounds, interests and involvement of those in the project, as well as the extent to which they are likely to be affected by it. The presentation should be rehearsed several times before it is made to the management.

Visual aids such as tables and graphs should be displayed with a variety of media. Flip charts of large pads of blank paper mounted on an easel enable the researcher to manipulate numbers. They are particularly useful in communicating answers to technical questions. Visual aids can also be drawn on the pages in advance, and the speaker flips through the pages during the presentation. Although not as flexible, magnetic boards and felt boards allow for rapid presentation of previously prepared material. Overhead projectors can present simple charts as well as complex overlays produced by the successive additions of new images to the screen. The use of computer packages such as Microsoft's PowerPoint can also be of immense help. They can be used for making computer-controlled presentations or for presenting technical information such as analytical models. However, the presenter must not lose sight of the message, as illustrated in the following two examples.

example

Bridging the gap[14]

The following summarises the views of Cristina Stuart, managing director of SpeakEasy Training, and Khalid Aziz, Chairman of the Aziz Corporation. In essence they argue that the oral presentation is not about the slick use of technology – the presenter should add something to the visual presentation.

Actors understand that there is a gap between the speaker and the audience and that you have to do a certain amount of work to bridge that gap. Actors spend their lives at interviews and are constantly having to present themselves to new people. Like actors, some speakers use props to enhance their performance but these need to be handled with care. It is the person that must be persuasive, not the PC. People often hide behind their visuals but if that is all that you are presenting, you might as well have sent them in the post. Technology may be used where it is appropriate, but too often people overdo the visuals to the detriment of the message. Even at Microsoft, where with its PowerPoint software the medium is the message, the presenter presides over the technology, not the other way around. ■

example

Screen saviours[15]

Phillip Redding, Deputy Managing Director of The Presentation Company, emphasises that technology is only part of the presentation picture. For him, it is just as important to get the content and structure of the argument right, and to create designs that put them over in a powerful way. 'If the content and structure are not clear, the presentation is not going to stand a chance. People think that by adding 3D animation the audience will be impressed. They will, but they still won't think the argument is very good.'

Video recorders (VCRs) and large-screen projectors are particularly effective in presenting focus groups and other aspects of fieldwork that are dynamic in nature. It is important to maintain eye contact and to interact with the audience during the presentation. Sufficient

opportunity should be provided for questions, both during and after the presentation. The presentation should be made interesting and convincing with the use of appropriate stories, examples, experiences and quotations. Filler words like 'uh', 'y'know' and 'all right' should not be used.

Body language should be employed. Descriptive gestures are used to clarify or enhance verbal communication. Emphatic gestures are used to emphasise what is being said. Suggestive gestures are symbols of ideas and emotions. Prompting gestures are used to elicit a desired response from the audience. The speaker should vary the volume, pitch, voice quality, articulation and rate while speaking. The presentation should terminate with a strong closing. To stress its importance, the presentation should be sponsored by a top-level manager in the client's organisation. ■

After the presentation, key executives in the client firm should be given time to read the report in detail.

Research follow-up

The researcher's task does not end with the oral presentation. Two other tasks remain. The researcher should help the client understand and implement the findings and take follow-up action. Second, while it is still fresh in the researcher's mind, the entire marketing research project should be evaluated.

Assisting the client

After the client has read the report in detail, several questions may arise. Parts of the report, particularly those dealing with technical matters, may not be understood and the researcher should provide the help needed. Sometimes the researcher helps implement the findings. Often, the client retains the researcher to help with the selection of a new product or advertising agency, development of a pricing policy, market segmentation or other marketing actions. An important reason for client follow-up is to discuss further research projects. For example, the researcher and management may agree to repeat the study after two years. Finally, the researcher should help the client firm make the information generated in the marketing research project a part of the firm's marketing (management) information system (MIS) or decision support system (DSS), as discussed in Chapter 1. A key element in researchers being able to assist marketing decision-makers is the level of trust that exists between the two parties. The nature of personal interaction between managers and researchers is very important in creating trust in the researcher and consequently in the results of the research. The quality of personal interaction affects managers' perceptions of the overall quality of the report itself.[16] Trust between the decision-maker and the researcher has been found to influence the perceived quality of user–researcher interactions, the level of researcher involvement, the level of user commitment to the relationship and the level of market research utilisation.[17]

Evaluation of the research project

Although marketing research is scientific, which may seem to imply a rigid, systematic process, it clearly involves creativity, intuition and personal judgement. Hence, every marketing research project provides an opportunity for learning, and the researcher should critically evaluate the entire project to obtain new insights and knowledge. The key question to ask is 'Could this project have been conducted more effectively or efficiently?' This question, of course, raises several more specific questions. Could the problem have been defined differently so as to enhance the value of the project to the

client or reduce the costs? Could a different approach have yielded better results? Was the research design that was used the best? How about the method of data collection? Should street interviews been used instead of telephone interviews? Was the sampling plan employed the most appropriate? Were the sources of possible design error correctly anticipated and kept under control, at least in a qualitative sense? If not, what changes could have been made? How could the selection, training and supervision of fieldworkers be altered to improve data collection? Was the data analysis strategy effective in yielding information useful for decision-making? Were the conclusions and recommendations appropriate and useful to the client? Was the report adequately written and presented? Was the project completed within the time and budget allocated? If not, what went wrong? The insights gained from such an evaluation will benefit the researcher and the subsequent projects conducted.

International marketing research

The guidelines presented earlier in this chapter apply to international marketing research as well, although report preparation may be complicated by the need to prepare reports for management in different countries and in different languages. In such a case, the researcher should prepare different versions of the report, each geared to specific readers. The different reports should be comparable, although the formats may differ. The guidelines for oral presentation are also similar to those given earlier, with the added proviso that the presenter should be sensitive to cultural norms. For example, making jokes, which is frequently done in many countries, is not appropriate in all cultures (which may also include particular organisational cultures). Most marketing decisions are made from facts and figures arising out of marketing research. But these figures and how they have been arrived at have to be credible to decision-makers. The subjective experience and gut feeling of managers could vary widely across countries, necessitating that different recommendations be made for implementing the research findings in different countries. This is particularly important when making innovative or creative recommendations such as in advertising campaigns.

Ethics in marketing research

Many issues pertaining to research integrity arise during report preparation and presentation. A survey of 254 marketing researchers found that 33% believed that the most difficult ethical problems they face pertain to issues of research integrity. These issues included ignoring pertinent data, compromising the research design, deliberately misusing statistics, falsifying figures, altering research results and misinterpreting the results with the objective of supporting a personal or corporate point of view, and withholding information.[18] It is important that researchers deal with these issues in a satisfactory manner and prepare a report which accurately and fully discloses the details of all the procedures and findings.

Objectivity should be maintained throughout the research process. For example, when data are analysed and no meaningful results are found, researchers are tempted to see findings which are not supported by the analysis. One example is meaningfully interpreting a regression equation when all the independent variables turn out to be non-significant (Chapter 20). Ethical dilemmas can arise in these

instances. The researchers are being paid for their expert interpretation of data, and can nothing meaningful be said? 'To arrive at some rational, logical, and convincing conclusion is so much more satisfying intellectually than to admit that the findings are inconsistent and inconclusive. No wonder we find ourselves mentally selecting and shaping what might otherwise be shapeless into a coherent, well-defined story.'[19] Such temptations must be resisted to avoid unethical conduct.

Like researchers, clients also have the responsibility for full and accurate disclosure of the research findings and are obligated to employ these findings honourably. For example, the public can be negatively affected by a client who distorts the research findings to develop a more favourable television advertising campaign. Ethical issues also arise when client firms, such as tobacco companies, use marketing research findings to formulate questionable marketing programmes.

➤➤➤
See Professional Perspective 7.

Internet and computer applications

Marketing research reports are being published or posted directly on the Internet or on intranets. Go to the Companion Website and read Professional Perspective 7 by Arno Hummerston, 'Net reporting comes of age'. Arno describes how the Internet has allowed the readers of research reports to tailor findings to suit their own unique needs.

Reports on the Internet are not located in publicly accessible areas but in locations protected by passwords or corporate intranets. The various word-processing, spreatsheet and presentation packages have the capability to produce material in a format that can be posted directly to the Web, thus facilitating the process.

There are a number of advantages to publishing marketing research reports on the Web. These reports can incorporate all kinds of multimedia presentations, including graphs, pictures, animation, audio and full-motion video. The dissemination is immediate and the reports can be accessed by authorised persons online on a worldwide basis. These reports can be electronically searched to identify materials of specific interest. For example, a manager in Kuala Lumpur can electronically locate the portions of a report that pertain to Malaysia or Southeast Asia. Storage or future retrieval is efficient and effortless. It is easy to integrate these reports to become part of a decision support system. The main disadvantage is that the readers may not have permanent access to the reports, as Websites may change periodically.

Given the ability to electronically search and tailor reports on the Internet, information published on Websites is now becoming more 'pull' oriented, as opposed to the 'push' orientation of a printed report. The standard analysis that the researcher generates can continue to be written, but can now be accessed as a result of a search, instigation of a link, or even just rolling the mouse over an icon. In addition, using database-driven technology it is possible to have a completely interactive Website that allows data interrogation through the specification of questions, filters, cross-tabulations and even applied weighting.

The basic structure of a Website, and the ease with which it is possible to navigate around a large amount of information, ensures that managers can quickly find exactly what they want. It is also possible to have index areas constantly visible on the screen to ensure that areas of information contained within a single report, or indeed multiple reports, can be accessed quickly.

Examples of how marketing research reports on the Internet are making decision-makers' lives easier are as follows: [20]

- Reporting and interrogating real-time data (not just from Web interviews but from CATI and CAPI data)
- Linking different research projects' reports together to create a more detailed overview
- Building charts and tables by adding different elements (such as confidence limits and explanation of chart movements).
- Applying rules to the reporting to ensure the robustness of the presented results.
- Applying complex modelling calculations and processes to data as it is made available.

Viewing marketing research reports using the Internet is effectively the same as conducting a search for secondary data as detailed in Chapter 4. In order to get a feel for the different styles of report presentation, based upon the types of research technique used, the country in which it was conducted and the industry supported, visit the Websites of leading marketing research agencies. Click on **www.tnsofres.com**, the Web address of Taylor Nelson Sofres, and **www.gfk.de**, the Web address of the German marketing research company GfK. On both sites (which cover the many countries they operate in, with different languages) you will see case studies, descriptions of special studies related to specific industry sectors and examples of reports that they produce.

Summary

Report preparation and presentation is the final step in the marketing research project. This process begins with interpretation of data analysis results and leads to conclusions and recommendations. Next, the formal report is written and an oral presentation made. After management has read the report, the researcher should conduct a follow-up, assisting management and undertaking a thorough evaluation of the marketing research project.

In international marketing research, report preparation may be complicated by the need to prepare reports for management in different countries and in different languages. Several ethical issues are pertinent, particularly those related to the interpretation and reporting of the research process and findings to the client and the use of these results by the client.

The final example presents a metaphor of the use of the guitar in supporting presentations. It is a final reminder that the power of computing software can never replace the creative skills of conveying the story and impact of a piece of research upon a decision-making situation.

e x a m p l e

My paradigm is the guitar[21]

Developments in modern technology have had a profound impact on the art of business presentation, most notably through PC-driven presentations. There is no doubt that the standard of visualisation in presentations has improved immeasurably, but has the presentation itself? Technically good presentations are becoming commonplace, perhaps even predictable. Predictability precedes boredom.

Presenters spend too much time at the PC creating a slide show and not enough on their performance. Presenters have forgotten to plan their personal involvement and the involvement of their audience.

Some presenters forget to plan their personal involvement in a presentation.

My paradigm is the guitar. The guitar represents a tool that supports presentation, rather like a PC, but that can never do the performance for you. From my own experience as a guitarist, presentations and gigs have many parallels. You have to prepare diligently, and have a good plan for the progress of the performance. You should know your material. You must be able to excite the audience and get them involved. You must be able to improvise and respond to requests. You should have a good guitar, but the good guitar on its own won't carry the day. You will. ■

Questions

1 Describe the process of report preparation.

2 Why is the quality of report presentation vital to the success of a marketing research project?

3 Describe a commonly used format for writing marketing research reports.

4 Describe the following parts of a report: title page, table of contents, executive summary, problem definition, research design, data analysis, conclusions and recommendations.

5 Why is the 'limitations and caveats' section included in the report?

6 Discuss the importance of objectivity in writing a marketing research report.

7 Describe the guidelines for report writing.

8 How should the data items be arranged in a table?

9 What is a pie chart? For what type of information is it suitable? For what type of information is it not suitable?

10 Describe a line chart. What kind of information is commonly displayed using such charts?

11 What are the advantages and disadvantages of presenting data using 3D charts?

12 What is the purpose of an oral presentation? What guidelines should be followed in an oral presentation?

13 To what extent should marketing researchers interpret the information they present in a report?

14 Describe the evaluation of a marketing research project in retrospect.

15 Graphically represent the consumer decision-making process described in the following paragraph:

The consumer first becomes aware of the need. Then the consumer simultaneously searches for information from several sources: retailers, advertising, word of mouth, and independent publications. After that a criterion is developed for evaluating the available brands in the marketplace. Based on this evaluation, the most preferred brand is selected.

Notes

1 Acreman, S. and Pegram, B., 'Getting to know you', *Research* (November 1999), 36–41.

2 Tufte, E.R., *Visual Explanations: Images and Quantities, Evidence and Narrative* (Cheshire, CT: Graphic, 1997); Fink, A., *How to Report on Surveys* (Thousand Oaks, CA: Sage, 1995).

3 Birks, D.F., 'Market research', in Baker, M.J. (ed.), *The Marketing Book*, 3rd edn (Oxford: Butterworth Heinemann, 1994), 262.

4 Keys Jr, T., 'Report writing', *Internal Auditor* 53(4) (August 1996), 65–6.

5 Britt, S.H., 'The writing of readable research reports', *Journal of Marketing Research* (May 1971), 265. See also Mort, S., *Professional Report Writing* (Brookfield, IL: Ashgate, 1995); Shair, D.I., 'Report Writing', *HR Focus* 71(2) (February 1994), 20.

6 Boland, A., 'Got report-o-phobia?: Follow these simple steps to get those ideas onto paper', *Chemical Engineering* 103(3) (March 1996), 131–2.

7 Wilson, L.D., 'Are appraisal reports logical fallacies', *Appraisal Journal* 64(2) (April 1996), 129–33; Leach, J., 'Seven steps to better writing', *Planning* 59(6) (June 1993), 26–7; Ehrenberg, A.S.C., 'The problem of numeracy', *American Statistician* 35 (May 1981), 67–71.

8 Wallgren, A., Wallgren, B., Persson, R., Jorner, U. and Haaland, J.A., *Graphing Statistics and Data* (Thousand Oaks, CA: Sage, 1996); Tufte, E.R., *Visual Display of Quantitative Information* (Cheshire, CT: Graphic, 1992).

9 Kauder, N.B., 'Pictures worth a thousand words', *American Demographics* (Tools Supplement) (November/December 1996), 64–8.

10 Hinkin, S., 'Charting your course to effective information graphics', *Presentations* 9(11) (November 1995), 28–32.

11 Chen, M.T., 'An innovative project report', *Cost Engineering* 38(4) (April 1996), 41–5; Zelazny, G., *Say It with Charts*, 3rd edn (Homewood, IL: Business One Irwin, 1996).

12 Johnson, S. and Regan, M., 'A new use for an old tool', *Quality Progress* 29(11) (November 1996), 144; Parr, G.L., 'Pretty-darned-quick flowchart creation', *Quality* (August 1996), 62–3.

13 Verluyten, S.P., 'Business communication and intercultural communication in Europe: the state of the art', *Business Communication Quarterly* 60(2) (June 1997) 135–43; Hynes, G.E. and Bhatia, V., 'Graduate business students' preferences for the managerial communication course curriculum', *Business Communication Quarterly* 59(2) (June 1996), 45–55.

14 Miller, R., 'In the spotlight', *Marketing* (1997), 35.

15 Condon, R., 'Screen saviours', *Marketing* (8 January 1998), 24.

16 Deshpande, R. and Zaltman, G., 'Factors affecting the use of market research information: a path analysis', *Journal of Marketing Research* 19 (February 1982), 25.

17 Moorman, C., Deshpande, R. and Zaltman, G., 'Factors affecting trust in market research relationships', *Journal of Marketing* 57 (January 1993), 81–101.

18 Milton-Smith, J., 'Business ethics in Australia and New Zealand', *Journal of Business Ethics* 16(14) (October 1997), 1485–97; Chonko, L.B., *Ethical Decision Making in Marketing* (Thousand Oaks, CA: Sage, 1995).

19 Day, R.L., 'A comment on "Ethics in marketing research"', *Journal of Marketing Research* 11 (1974), 232–3.

20 Hummerston, A., 'Net reporting comes of age', *Research* (May 2000), 36.

21 Willetts N.J., 'Going Live', *Marketing Week* (13 November 1997), 47–8.

International marketing research

Stage 1
Problem definition

Stage 2
Research approach developed

Stage 3
Research design developed

Stage 4
Fieldwork or data collection

Stage 5
Data preparation and analysis

Stage 6
Report preparation and presentation

Objectives

After reading this chapter, you should be able to:

1 develop a framework for conducting international marketing research;

2 explain in detail the marketing, governmental, legal, economic, structural, informational and technological, and socio-cultural environmental factors and how they have an impact on international marketing research;

3 describe characteristics of the use of secondary data, qualitative techniques, telephone, personal and mail survey methods in different countries;

4 discuss how to establish the equivalence of scales and measures, including construct, operational, scalar and linguistic equivalence;

5 describe the processes of back translation and parallel translation in translating a questionnaire into a different language;

6 discuss the ethical considerations in international marketing research.

Marketing across national boundaries remains one of the most difficult activities for many manufacturers in many product and service areas. Likewise, the research problems involved in multi-country coordination can be formidable.[1]

Overview

This chapter starts by evaluating the need for international marketing research. It then discusses the environment in which international marketing research is conducted, focusing on the marketing, government, legal, economic, structural, informational and technological, and socio-cultural environment.[2] Although illustrations of how the six steps of the marketing research process should be implemented in an international setting have been presented in earlier chapters, here we present additional details on secondary data, qualitative techniques, survey methods, scaling techniques and questionnaire translation. Relevant ethical issues in international marketing research are identified. We begin with some examples illustrating the role of marketing research in international marketing. The first example illustrates the use of ethnographic techniques, as introduced in Chapter 6. It also illustrates the point that research techniques may have to be adapted and developed to suit the nature of consumers in different cultures. The second example illustrates the great potential in international markets and some of the technical challenges that must be overcome to conduct good quality marketing research.

e x a m p l e   Tasty wisdom[3]

'The most accurate research is done for pet food companies because it is not possible to ask pets what they think, only observe what they do.' Such an approach is just as feasible with humans! Economies of scale may be driving some research companies to harmonise their methodologies when conducting international marketing research, but Nestlé decided it needed more than just traditional focus groups to understand why its products were not thriving in Russia. It needed an ethnographic study of culinary habits to help develop a unified marketing approach. Incorporating video observation into a study of cooking habits by Russian housewives enabled Nestlé to sift out how people see themselves versus how they really are. Nestlé's study, using a range of different techniques, unmasked the myth of the creative cook interested in exotic recipes. The food giant reviewed its product portfolio accordingly. ■

e x a m p l e   Breaking down the great wall[4]

Although it has never been easy to do business in China, few multinationals have been able to resist the lure of potentially huge growth in the largest untapped market on earth. Research proved essential in tackling this unknown region, and international marketing research networks followed fast on the heels of their clients, while local agencies mushroomed. However, curbs on marketing research have made it even harder to conduct business.

Any survey conducted for or carried out by a company with foreign investment now has to be cleared twice by the authorities: once when the questionnaire is drafted and again before findings are handed over to the client. Marketing research sources believe China is more concerned with social than with consumer research. 'China is not ready for published opinion polling,' one Hong Kong based researcher said. Another, based in mainland China, said: 'after they check all the projects, maybe they will think "this is about washing powder, this is about TV sets. It has no relation to national security at all".'

In recent years, timely research has become even more important: the competition has been hotting up while the economy has been slowing down. The time added on to projects (around a week) is a real hindrance. Perhaps even more worrying is the handling of data. Despite State Statistical Bureau promises to safeguard the data, clients point out the office is undermanned. 'You really don't know how confidential it is. You do your research and submit your report to the bureau, there's a probability for a research study to leak out to a competitor.' ■

Another example comes from the many technical challenges in the GlobalCash study which was conducted in 20 countries across Europe. The following example illustrates the dilemma faced over just one question in the questionnaire that was developed.

Open or closed questions?[5]

In the development of codes for a questionnaire, it is a simple task to assign numbers to closed questions such as 'do you use netting systems?', assigning '1' to a *yes* response and '2' to a *no*. In the GlobalCash study a coding problem arose in a question that was open-ended in presentation but in reality was a closed list of banks and software companies. As an open-ended question, a single line was needed to establish a lead bank used in 21 European countries. If the question was presented in a closed format with boxes to tick, it would have had to cope with over 350 banks in Germany, 250 in Switzerland, 200 in France, 200 in Italy and 180 in Britain, to start with. As well as being too cumbersome for the respondent, there might have been some confusion as to the correct name or initials of a bank which may be used within a particular country. Banks may also have merged and respondents may use the pre-merger bank name. Each business school had to 'translate' the shorthand versions of bank names used by respondents in their country. Thus there was much work that had to be done to code the open-ended question, but the alternative of creating a closed question both would have been cumbersome to the respondent and would miss the cultural nuances of how they name the banks they work with. ■

What is international marketing research?

The term international marketing research can be used very broadly. It denotes research for true international products (international research), research carried out in a country other than the country of the research-commissioning organisation (foreign research), research conducted in all important countries where the company is represented (multinational research), and research conducted in and across different cultures (cross-cultural research). The latter category of cross-cultural research does not have to cross national boundaries. Many European countries have a vast array of ethnic groups, giving the marketing researcher many challenges in understanding consumers in these groups within their home country. The following example is a brief illustration of the array of problems faced by researchers working in a London local authority.

Qualitative research among ethnic minority communities in Britain[6]

Although the language barrier remains a problem among non-English speaking minorities in Britain, communicating or conducting research with them can still be done provided their traditional patterns of behaviour are observed. Two examples of the issues faced in studies of ethnic minorities in the London Borough of Newham include:

■ Is there a difference between the Punjabi spoken by Pakistanis and Indians in Britain? This will depend on the exact area from which the people come: the written forms are entirely different, and the more formal spoken forms may differ too, but there can be no hard and fast rules. Punjabi is the language spoken in the Punjab region of India and Pakistan. Indian Punjabis usually speak Hindi (the national language of India) as well, while Pakistani Punjabis usually speak Urdu (the national language of Pakistan) as well.
■ Should one use Bengali or Sylheti for written stimulus directed at the Bangladeshi population? Sylheti is not a written language, so if people from the Sylhet region are literate in their mother tongue, they will read Bengali. ■

The example above supports the contention that, methodologically, there is no difference between domestic and international marketing research. In other words, marketing researchers have to adapt their techniques to their target respondents, and the subtle cultural differences within a country offer the same challenges as the much more apparent differences over thousands of miles. The following example does not

present an actual case but the opinions of the respected international marketing researcher Mary Goodyear (MBL). Her views support the contention that there is no difference between domestic and international research but then show why international marketing research should be examined as a distinct subject area.

The world over, a group is a group is a group[7]

It has become the convention to talk about international research as if it were a discrete sector of the industry. Methodologically, in qualitative research at least, it is not really very different from what is done at home. A group is a group is a group. However, it is important to realise that international research is different: firstly, in terms of analysing the social and cultural dynamics of the society, how the foreign society is structured, how power is gained and expressed, the roles allotted to men and women, the interactions between different sectors of society and so on; and secondly, in the practical problems involved in organising the research. ■

The first point established in the above example is that many of the cultural and societal assumptions that we take for granted have to be examined. This we discussed when considering how researchers 'see' in Chapter 9, 'Qualitative research: data analysis'. Without such an examination we have the potential to be naive in the questions we pose, whom we pose those questions to, and the manner in which we interpret the answers we generate. The second point reminds us of the features of the infrastructure that supports the research process, features that we may get used to and take for granted. For example, without the benefits of accurate sampling frames, accurate and up-to-date secondary data and reliable and widespread use of telephones, our research plans may be weak. Our whole approach to organising and conducting research in international markets has to reflect these physical conditions.

Overseas expansion has become more important for European companies due to increasing economic integration, the lowering of trade barriers and in some instances intensive competition in domestic markets. International markets such as the vast developing economies of China, India and Russia offer huge prospective returns. As attractive as these markets may be, companies must realise that setting up operations in these countries does not guarantee success. Understanding the nature and scope of these markets is vital to realising the opportunities and presents a major challenge to marketing researchers. This challenge also lies in overcoming the 'red tape' in countries that may well have laws and policies designed to protect their own countries' businesses. Realising the opportunities and overcoming the challenges may mean, at one extreme, that marketing researchers work with decision-makers in their own domestic markets, trying to coordinate research in distant lands. It may mean that companies have operations in overseas markets which extend to indigenous marketing research operations; an example of this is the Indian subsidiary of Unilever.

Unilever's Indian subsidiary has a Marketing Research department that the global group considers to be a laboratory for research in developing markets. The marketing research team is manned by 65 people, including 25 highly respected senior managers, a fieldforce of over 1,600 and an annual budget of over €7 million. The team operates in a wide range of areas, from basic product and advertising testing, to trend forecasting and media research. The urban consumer panel it runs, of 50,000 households in 43 towns, must be one of the largest in the world. The following example illustrates how their rich local knowledge helped to create a successful research programme and financially successful marketing decisions.

Mud and lemons [8]

Hindustan Lever Limited (HLL), Unilever's Indian subsidiary, is big in India but has regularly faced competition from small, local entrepreneurs who have competed on price. This was the situation that HLL faced in 1997, when sales of the Vim soap bar for dishwashing had stagnated, being undercut by the Odopic Bar which was selling at half the price of Vim. HLL panicked, but its marketing research department began a general study of consumer behaviour in this category, which found that Vim bar's basic position seemed strong if one looked only at dishwashing bars, but compared with all other products, it wasn't.

Competition came from low-cost detergents, ash, lemons and even mud, which were essentially scourers. They got dishes fairly clean and they were cheap – free, in fact, in the case of ash taken from the residue of cooking. Taking these products into account, Vim's share of the market dropped from 52% to just 8%. Clearly the real issue for HLL was not Odopic, but how to grow the category by encouraging conversions from these dishwashing rivals. To determine consumer attitudes, HLL Marketing Research set up in-depth research projects, which concluded that no-one particularly liked the alternatives to Vim, but their low prices made them hard to resist.

HLL decided to avoid the price game, where the smaller players always had an advantage. Instead it would create a significantly better product for only a small increase in the price, one that would have to be better in quantifiable terms, like the ability to remove grease, and in terms of consumer perceptions. HLL Marketing Research used a sequential recycling technique where product development went hand-in-hand with continuous research with the same group of consumers. HLL and its consumer panel reached a product and price they were satisfied with. The end result was that Vim brand sales rose nearly threefold in the year 2000. ■

It could well be argued that the above example does not represent international marketing research, as it is domestic research conducted by an Indian company. This adds further support to the contention that, methodologically, there is no difference between domestic and international marketing research. For companies with such a global reach as Unilever, whose economies of scale allow them to operate in this manner, their international research does ultimately become a domestic issue. However, for companies without the financial muscle of Unilever that wish to realise the opportunities in international markets, decisions have to be made about how they plan and implement research from their domestic base. To see an example of a strong domestic company seeking opportunities in international markets, go to the Companion Website and read Professional Perspective 19 'The role of marketing research in the international strategies of German trade fair companies' by Bernd Aufderheide. Bernd describes why strategies to exploit international markets became vital to the German trade fair industry and how, in developing these strategies, marketing research plays a vital role.

A framework for international marketing research

Conducting international marketing research can be much more complex than conducting domestic marketing research. Although the basic six-step framework for domestic marketing research is applicable, the environment prevailing in the countries, cultural units or international markets that are being researched influences the way the six steps of the marketing research process should be performed. Figure 26.1 presents a framework for conducting international marketing research.

The differences in the environments of countries, cultural units or foreign markets should be considered when conducting international marketing research. These differences may arise in the marketing, government, legal, economic, structural, informational and technological, and socio-cultural environments, as shown in Figure 26.1.
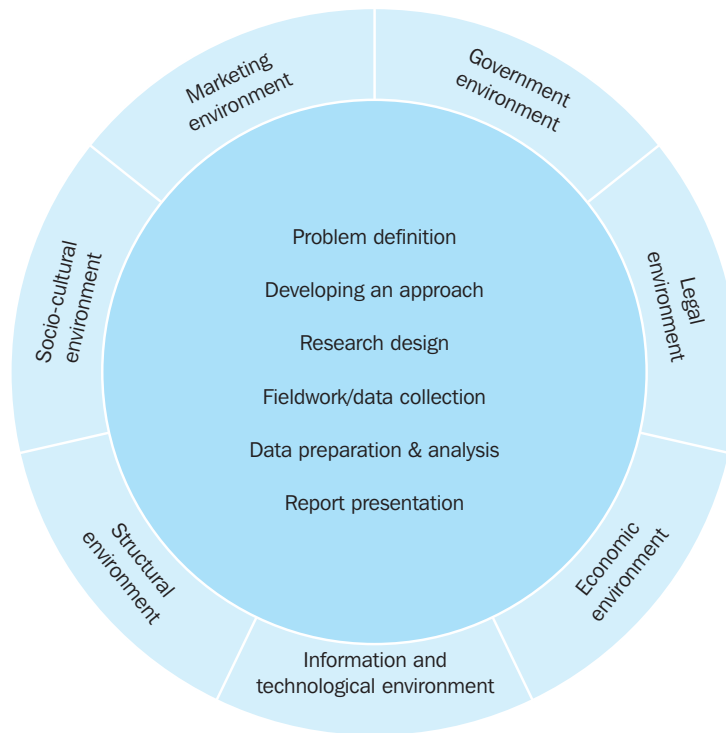
Figure 26.1
A framework for
international marketing
research

Marketing environment

The role of marketing in economic development varies in different countries. For example, many developing countries are frequently oriented towards production rather than marketing. Demand typically exceeds supply, and there is little concern about customer satisfaction, especially because the level of competition is low. In assessing the marketing environment, the researcher should consider the variety and assortment of products available, pricing policies, government control of media and the public's attitude towards advertising, the efficiency of the distribution system, the level of marketing effort undertaken, and the unsatisfied needs and behaviour of consumers. For example, surveys conducted in Europe usually involve questions on the variety and selection of merchandise. These questions would be inappropriate in many African countries, which are characterised by shortage economies. Likewise, questions about pricing may have to incorporate bargaining as an integral part of the exchange process. Questions about promotion should be modified as well. Television advertising, an extremely important promotion vehicle in Europe, is restricted or prohibited in many countries where TV stations are owned and operated by the government. Certain themes, words and illustrations used in Europe are taboo in some countries. This is illustrated in the following example which describes characteristics of research in the Middle East.

example

Warming responses from the Levant to the Gulf [9]

The Middle East is one of the most fascinating and challenging areas for research. Market research is no longer a novelty for people in the Middle East and is relatively well developed in most countries. Local Arab females' mobility is generally restricted and all interviews have to be conducted at the recruiter's or respondent's home. However, this means that the chances of finding women at home and obtaining an interview from them are much higher than in Europe. Using visuals in research can be problematic if they are not carefully selected before they are sent and presented to local custom. It is particularly essential to control the amount

of female skin visually exposed. One should never forget that sexual connotations, verbal as well as visual, are unacceptable in most Middle Eastern countries.

One of the main difficulties encountered in interviewing Arabs (and females in particular) is their natural tendency to want to please others, including researchers. Careful formulation of questions is essential to a fruitful interview or discussion, together with a non-complacent interviewer or moderator. Indeed, true feelings, opinions or preferences are often disguised, if felt to be socially unacceptable: ratings are generally higher, especially when related to product attributes or purchase propensity. ■

Government environment

An additional relevant factor is the government environment. The type of government has a bearing on the emphasis on public policy, regulatory agencies, government incentives and penalties, and investment in government enterprises. Some governments, particularly in developing countries, do not encourage foreign competition. High tariff barriers create disincentives to the efficient use of marketing research approaches. Also, the role of government in setting market controls, developing infrastructure and acting as an entrepreneur should be carefully assessed. The role of government is also crucial in many advanced countries, where government has traditionally worked with industry towards a common national industrial policy. At the tactical level, the government determines tax structures, tariffs and product safety rules and regulations, often imposing special rules and regulations on foreign multinationals and their marketing practices. In many countries, the government may be an important member of the distribution channel. The government purchases essential products on a large scale and then sells them to consumers, perhaps on a rationed basis.

Legal environment

The legal environment encompasses common law, foreign law, international law, transaction law, antitrust, bribery and taxes. From the standpoint of international marketing research, particularly salient are laws related to the elements of the marketing mix. Product laws include those dealing with product quality, packaging, warranty and after-sales service, patents, trademarks and copyright. Laws on pricing deal with price fixing, price discrimination, variable pricing, price controls and retail price maintenance. Distribution laws relate to exclusive territory arrangements, type of channels, and cancellation of distributor or wholesaler agreements. Likewise, laws govern the type of promotional methods that can be employed. Although all countries have laws regulating marketing activities, some countries have only a few laws that are loosely enforced and others have many complicated laws that are strictly enforced. In many countries the legal channels are clogged and the settlement of court cases is prolonged. In addition, home-country laws may also apply while conducting business or marketing research in foreign countries.

Economic environment

Economic environmental characteristics include economic size (gross domestic product, or GDP); level, source, and distribution of income; growth trends; and sectoral trends. A country's stage of economic development determines the size, the degree of modernisation, and the standardisation of its markets. Consumer, industrial and commercial markets become more standardised and consumers' work, leisure and lifestyles become more homogenised by economic development and advances in technology. The following example illustrates the problems inherent in understanding the characteristics of a country's stage of economic development and in forecasting the potential that exists there.

The East gets out of the red and into the black[10]

Dr Rudolf Bretschneider, Managing Director of Austrian-based FESSEL-GfK, is credited as the frontiersman in the GfK network (**www.gfk.de**) to make the first bold steps into Eastern Europe: GfK credits itself with being the first Western agency to spot market research potential in the region. Bretschneider cautions against relying too greatly on official economic indicators which, given the substantial size of the black and grey markets, can never give more than an illusory and fragmentary picture. GfK discovered that they had overestimated their knowledge of the region. Their experience highlights the pace of change and the meteoric swings that are commonplace, requiring extra vigilance from all observers. ■

Informational and technological environment

Elements of the informational and technological environment include information and communication systems, computerisation, use of electronic equipment, energy, production technology, science and invention. For example, in India, South Korea and many Latin American countries, advances in science and technology have not had a proportionate impact on the lifestyle of the majority of citizens. Computers and electronic information transfer have still to make an impact at grassroots level. Information handling and record keeping are performed in the traditional way. Again, this has an impact on the type of information that can be solicited from consumers, businesses and other enterprises.

Structural environment

Structural factors relate to transportation, communication, utilities and infrastructure. For example, telephone usage in the Far East, where many households do without telephones, is much lower than in Europe. Mail service is inefficient in many developing countries. Personal contact with respondents is difficult because city people work during the day and rural residents are inaccessible. Block statistics and maps are not available or can be obtained only with great difficulty. Many households and homes are unidentified.

Based on geography alone, national samples in China and Indonesia are almost unthinkable. China is predominantly rural and Indonesia consists of several thousand islands. In Thailand, only 21% of the population live in urban areas and two-thirds of the total urban population live in a single city, Bangkok. In these three countries, interviewing by telephone is, for the most part, not a consideration.[11]

Socio-cultural environment

Socio-cultural factors include values, literacy, language, religion, communication patterns, and family and social institutions. Relevant values and attitudes towards time, achievement, work, authority, wealth, scientific method, risk, innovation, change, and the Western world should be considered. The marketing research process should be modified so that it does not conflict with the cultural values. In many developing countries, 60% or more of the population is illiterate. In tradition-directed, less-developed societies, the ability of respondents to formulate opinions of their own seems to be all but absent; consequently, it is difficult to solicit information from these respondents. As a result, the sophisticated rating scales employed in Europe are not useful. There may also be several distinct spoken languages and dialects in a given nation or region.

In India, for example, there are 19 major languages and more than 200 dialects. India is divided into linguistic states. The country can be described as a mini-Europe, each state like a separate country within Europe, with its own language and cultural peculiarities. A survey which even approaches national representation in scope will generally be printed in at least 12 languages.[12]

A country with a homogeneous family structure is likely to be more culturally homogeneous than a country with multiple family structures. For example, Japan is culturally more homogeneous than many European or African countries, which have many different kinds of family structures.

Finally, what is meant by terms such as 'the youth market'? Many marketers believe that youth markets on a global basis are becoming more alike in their attitudes, lifestyles and aspirations. The following two examples urge caution in such assumptions, presenting small sketches of the forces shaping 'youth markets' in Asia and Europe.

'I want my MTV, but in Mandarin, please'[13]

Many magazines would have you believe that Asian teens, in appearance, attitude and posture, increasingly resemble American and European teens. However, this convergence view of global teen lifestyle overlooks the reality of Asian cultural diversity. The befuddling paradox is that, as Asian teens have become more globally aware, they have also grown increasingly confident in their local identity. This conflict between global and local is becoming more and more important, being reflected in food, music, language and brands. To appreciate this conflict, it is important to understand what 'Asianness' means to Asian teens. What makes Asianness a valid concept is its communities' resistance to deep cultural change, despite rapid technological and economic progress. Asia's open and purposeful resistance to change in general is rooted in fear of failure and its resultant loss of 'face'. In risk-averse Asia, the status quo is not such a bad place and any change is incremental. The salience of race, community and the family unit, downsizing notwithstanding, is unchallenged. The obvious signs of these include the enduring popularity of traditional costumes and aesthetics; the continued strength of extended families, family associations and networks; the continued power of religion; and a Confucian respect for elders. Even more interesting is how these values have over time been institutionalised into formal political or business practice. ■

Euro-youth: myth or reality?[14]

A major area of difference in European youth stems from the role of the family and attitudes to parenting. As expectations regarding living standards rise, and parental working hours extend to compensate, we are increasingly seeing a 'hands-off' style of parenting.

In Northern Europe we are increasingly seeing a 'hands-off' style of parenting.

Consequently, family related need-states become more difficult to satisfy, as parents are increasingly absent. This is especially true in Northern Europe, and nowhere more so than in Britain (which has the highest teen pregnancy and alcohol/drug abuse rates). Southern and certain eastern European Catholic countries tend to have more united family units, with extended support networks firmly in place. A distinct approach to personal space means that family members are likely to be resident close by and very much involved in young people's lives. Offspring remain living within the family home until much later (as late as 30 in Italy, for example) and parenting is much more 'hands-on'. This makes for a slower transition to emotional and financial independence. Tighter families provide an all-important sense of belonging to southern/eastern European teenagers. This is often denied to those from the north (with its more fragmented families), who are forced to make up for this deficit by turning to communal activities such as sports and clubbing to feel they are part of something. ■

Each country's environment is unique, so international marketing research must take into consideration the environmental characteristics of the countries or foreign markets involved. For companies that wish to expand into a large geographic region, this can cause problems. Should they invest in research into every country or even regions within countries, or will these be prohibitive in time and cost terms? If they cannot manage to research individual countries in a region, which countries should they focus upon? Questioning the resource issues of this dilemma, Martin Stollan, Associate Director of the research company IPSOS-Insight, offers the following argument.[15] 'Americans often tend to treat Europe as a single market, so that they feel that doing one country is enough, often choosing Britain because of the language, and that can be very misleading. An improvement is to choose one country in northern Europe and another in southern Europe. But ideally, you need to research each local market. Think of Britain and the differences between England, Wales and Scotland and you can understand the dangers of assuming that Portuguese consumers have the same attitudes as Spanish consumers (and by Spanish consumers do we mean Catalan and Basque as well as Castilian?).'

Researching from the security of a key market may miss subtleties that local researchers would be quick to pick up on. Another benefit of local researchers is that desk research and their knowledge can help to fill in many gaps. The ideal is to have a multi-country project, coordinated from one location with local suppliers carrying out individual country 'legs' within a common framework. The research approach should ensure quality control and consistency, and it should lead to actionable and global strategic recommendations, with all local markets represented. Such a maxim is fine, provided that international marketing researchers can demonstrate the potential returns from such an investment, which is inherently very difficult to achieve.

The extent to which companies may invest in research targeted at individual countries may be evaluated in terms of not only the potential returns and idiosyncrasies of that country, but also the complexities of implementing research techniques that will work well. To help understand the problems of implementing international marketing research, we provide additional details for implementing secondary data, qualitative techniques, survey methods, measurement and scaling techniques and questionnaire translation.[16]

Secondary data

Secondary data was covered in detail in Chapters 4 and 5. It is worth recalling where secondary data can support the research process, especially given some of the difficulties of conducting primary research in international markets. Secondary data can help to:

1 Diagnose the research problem
2 Develop an approach to the problem

3 Develop a sampling plan

4 Formulate an appropriate research design (for example, by identifying the key variables to measure or understand)

5 Answer certain research questions and test some hypotheses

6 Interpret primary data with more insight

7 Validate qualitative research findings.

Obtaining secondary data on international markets to gain support for the above is much simpler with the advent of the Internet. The Internet may allow access into international markets which in the past may have required travel to the country and a great amount of time-consuming searching. Judicious use of the Internet does not mean that the international researcher will automatically gain support in the above areas. There may still be little or no secondary data that relate to the issues we wish to research in an international market. For what data we can obtain, the principles of evaluating secondary data, as set out in Table 4.2, still apply. Conducting an evaluation of the nature, specifications, accuracy, currency and dependability may be far more difficult in international markets. The process and specifications of research conducted in international markets may not be explicit; one may need a deep understanding of conducting primary research in a country to understand why data have been collected in a particular manner. Language is also a major issue, not only in reading and interpreting data, but in the definitions used in measurements. Finally, secondary research from a target country may have been heavily influenced by political forces. The researcher may need to be aware of such influences in order to interpret and make use of the findings.

The following example illustrates the limitations of existing data in Russia, even though there are very competent researchers in the country. Overcoming this problem has helped to produce data that can be relied upon and is more relevant to a number of marketers.

example

Rise, rise, you Russian middle classes . . .[17]

All marketers need information to help understand their markets. One major problem is simply knowing who their consumers are. As one brand manager said: 'All my distribution is in Moscow and St Petersburg, but maybe only 30% of my final consumers are in these cities'. 1995 saw the introduction of the Target Group Index into Russia. Under licence from the British Market Research Bureau International, the Russian agency Comcon 2 has established the first single-source product and media survey covering all Russia. As in Britain, the objective is to gather information about the consumption and consumers of different goods and services, alongside their media preferences and attitudinal or lifestyle characteristics. Much previous research, while of high quality (the result of the high level of theoretical training of researchers within Russian universities), had been unsystematic and concentrated on the two major cities. ■

Another major development in the use of secondary data is in the development of geographic information systems. In the use of geodemographic classifications of consumers, systems already exist in Australia, Belgium, Britain, France, Germany, Hong Kong, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, South Africa, Spain, Sweden and the USA. The systems work well in defining consumers within each country and, as described in Chapter 5, can be used as a foundation to add transactional data generated within an organisation operating within each country. The international research problem emerges from the means of comparing classifications across the different countries. This problem emerges from the different data sources that are used to build the classifications in each country. There are different types of data that may be accessible, different standards and definitions used in data collection and different laws allowing the use, or not, of certain types of data.

Geodemographic databases can give an excellent introduction to a particular country and can form a foundation upon which other data sources can be built. At present, comparing consumer types across borders is problematic.

Qualitative techniques

Chapter 6 discussed the differences between European and US researchers in their approach to focus groups. In essence, a US approach is far more structured and tends to be a foundation for subsequent surveys, while the European approach tends to be more evolutionary and exploratory. The question of the perspective from which one plans and conducts qualitative research is not really important until other researchers or moderators from international markets become involved in the process. The approach adopted may not match the expectations of the decision-maker who will be supported by the research. These circumstances are outlined and illustrated in the following example.

example

When East meets West, quite what does 'qualitative' mean?[18]

International marketers have always been aware that qualitative research as it developed in the US and in Europe were quite different practices, stemming from different premises and yielding different results. American-style qualitative research started from the same evaluative premise as quantitative research but on a smaller scale. European-style qualitative research started from the opposite premise: it was developmental, exploratory and creative rather than evaluative.

The rest of the world tended to adopt one or the other style of qualitative research depending upon the history and trading partners of the country. For example, India adopted the European model, whereas Japan largely followed the American model. The difficulty for the Asia Pacific region is that the influences of both schools are found there concurrently. The founders of commercial research in the region came from the UK, Australia, New Zealand and America, bringing with them varying approaches. From a marketer's point of view, the problem is that both styles of qualitative research are currently on offer under the same description. Marketers may not be sure what they are getting; even worse, some research agencies seem unclear about what they are offering, serving up a confused mixture of both styles of qualitative research, falling short of achieving the objectives of either approach. ■

Decision-makers often require global answers to develop global marketing campaigns, yet data are often single-market related and cannot be easily interpreted across various countries. Researchers are wary of using varying methods across different countries and cultures and are especially wary of qualitative approaches. Using surveys, where far more control of the design and implementation can be afforded, many forms of bias can emerge when working across cultures. The problems of cross-cultural comparisons grow when qualitative techniques are used. Some researchers argue that international marketing researchers would be better off using direct observational approaches, rather than being misled by findings based upon the 'dubious' application of qualitative research techniques. [19]

Direct observational approaches are not feasible or even desirable for all international marketing researchers. Conducting qualitative research in a number of countries may mean approaching a local research company which is aware of local customs, language issues and administrative requirements. Nearly all international marketing researchers choose an in-country marketing research firm to complete the actual fieldwork, whether for focus group moderating or for interviews.[20] However, for some multi-country projects that are qualitative in nature, it may be impossible to find researchers with the cultural, linguistic and administrative experience that can

allow some degree of consistency and means to compare findings. The following example illustrates how bilingual researchers working in teams tackle this problem.

e x a m p l e

In-depth research? A bit controversial [21]

Consumer markets such as soft drinks or processed foods are important enough to be regularly researched across many countries. Wherever you go there are professional qualitative researchers who really understand them. The same cannot be said for most business-to-business markets. Just sending a topic guide and recruitment questionnaire to local agencies is not enough.

Some research agencies have responded by recruiting bilingual researchers, so that the same research team that does the 'home' interviews can, with as much ease, do interviews in target foreign countries. They feel that this is the only way to guarantee consistency in how the research is carried out. More fundamentally, it is the only way to keep the approach flexible. Why is flexibility so important? True qualitative research entails wording questions to suit respondents: if necessary, to keep rewording the question until the answer makes sense. It means the freedom to adapt the interview to suit the interviewer's developing understanding of the subject. ■

This example also illustrates the essence of qualitative research in terms of exploring through rewording, of adapting an interview until sense is made. In international marketing research this calls for linguistic skills matching the native speaker in any target country. This may not be a solution that is feasible for all international research projects. Depending upon the array of countries to be researched, recruiting researchers with the relevant linguistic skills may not be possible, but it is the best alternative should local research companies not exist or be considered to be of poor quality. Good quality local researchers should be sought out, because, even if researchers with the relevant linguistic abilities are available, there are other cultural nuances that can affect the trust, rapport and comfort of qualitative research respondents. Often participants are unfamiliar with the process or simply do not trust the researcher. In many Asian cultures, respondents are unwilling to offer an individual opinion that may differ from that of the group. For instance, the need for politeness can thwart the researcher's desire to get at respondents' real feelings in a typical Japanese focus group.

In general, patience and sensitivity are required to overcome cultural problems in conducting qualitative research in international markets.[22] An excellent working relationship with a local research company can be absolutely vital to success and in many cultures can only be achieved by developing a personal relationship. Listening to the advice of local research companies will also enhance sensitivity to respondents. Patience is expressed by a preparedness to lengthen the research schedule, extra care in translation, pre-tests of wording and tact, and respect for respondents and local researchers. Great care should be taken to be more selective in recruiting and grouping respondents for focus groups.

Whether shaped by European or American paradigms, good qualitative research should be open and a learning process. This openness and learning should be shared among decision-makers, the researchers coordinating research activities and local researchers working in individual countries. All parties should clearly understand the premises for exploring particular issues, probing certain individuals and the bases for interpreting data. In qualitative international marketing research, where there may be confusion about the premises for the whole approach, there is a much greater need for openness between all parties involved.

Survey methods

The following sections discuss the major interviewing methods in light of the challenges of conducting research in foreign countries, especially Europe and developing countries.[23]

Telephone interviewing and CATI

In the USA and Canada, the telephone has achieved almost total penetration of households. In North America, telephone interviewing is the dominant mode of questionnaire administration. Throughout Europe there are also many countries, regions and cities that have almost total telephone penetration. Over recent years, telephone interviewing techniques in Europe have grown enormously. However, even with high penetration levels, many Europeans are reluctant to divulge personal details over the telephone. This means that the technique does not have the dominance seen in North America.

The successful use of the telephone in international research depends upon three factors. The first is the level of telephone penetration (see Chapter 10 for a discussion of the issues related to the selection of probability samples in telephone interviewing). The second is the completeness and accuracy of telephone directories (the growth of the use of mobile phones and of 'ex-directory' numbers has compounded this factor). The third is the cultural acceptance of using the telephone to divulge personal details.

In Hong Kong, for example, 96% of households (other than those on outlying islands and on boats) can be contacted by telephone. With some persistence, evening telephone interviewing can successfully achieve interviews with 70–75% of selected respondents. Residents are uninhibited about using telephones and relaxed about telephone interviews. Even in these circumstances, this is not the most important mode of data collection.[24]

In most developing countries, telephone penetration is low. There may be relatively high concentrations of telephone usage in particular cities and with particular types of consumer (especially in professional classes). Telephone interviews are most useful with respondents in these countries who are accustomed to business transactions by phone.

With the decline of costs for international telephone calls, multi-country studies can be conducted from a single location.[25] This greatly reduces the time and costs associated with the organisation and control of the research project in each country. Furthermore, international calls can obtain a high response rate, and the results have been found to be stable (i.e. the same results are obtained from the first 100 interviews as from the next 200 or 500). It is necessary to find interviewers fluent in the relevant languages, but in most major European cities this is not a problem.

In-home and in-office personal interviews

In many European countries, home interviewing is the dominant means to conduct surveys. Given that in-home interviews require a large pool of qualified interviewers and are time-consuming and costly, this may seem odd. However, when one considers the quality of rapport that can be built up between the interviewer and respondent, the amount of probing and the quality of audiovisual stimuli that can be used, there are clear benefits that outweigh the costs. In many areas of industrial marketing research, the only means to contact certain managers may be through the personal interview, held in their office.

In international research this means understanding the cultural implications of entering someone's home or office. In the example 'Warming responses from the Levant to the Gulf' on page 667 it was noted that local Arab females' mobility is gen-

675

erally restricted and all interviews have to be conducted at the recruiter's or respondent's home. This meant that the chances of finding women at home and obtaining an interview from them were much higher than in Europe.

In the GlobalCash study, where in-office interviews were conducted throughout Europe, two interviewers were used. The first (who had a high level of technical knowledge of cash management) attended all the interviews to ensure that a consistent approach was used, especially when probing. The second was from a business school in the country being surveyed. The second interviewer guided the means to obtain the interview in the first instance, i.e. to gain access. They then ensured that the protocol was correct (given the type of respondent faced) and finally helped with any language difficulties. The following example further illustrates a cultural problem faced with in-office interviews.

e x a m p l e | **Russian roulette: exciting once you learn the new rules**[26]

Executives find it odd, and sometimes sinister, that people should just come to ask questions which often seem irrelevant. One respondent, a high official in a major city, kept interrupting the interview with offers to sell the product under discussion and asked our agency to supply some other products. He considered it fatuous that he should face a stream of questions without seeing any concrete result at the end. It is usually assumed that the researcher belongs to the company commissioning the research, even when the relationship is made clear. Guarantees that the respondents' identity will remain confidential can be treated with disbelief. ■

Street interviews

In many countries, with low telephone penetration and comparatively low literacy rates, the street interview offers an excellent means to survey. This is illustrated in the following example, showing Chinese respondents enjoying the marketing research task.

e x a m p l e | **Foreign policy**[27]

Interviewing is not usually an obstacle in mainland China, says Taylor Nelson Audits of Great Britain Group Director for Greater China, Richard Necchi. The problem is ensuring that you have a representative sample to work from. Census data are simply not up to Western standards, he says, and researchers try to rely on government statistics as little as possible. The huge migrant population of about 100 million also means that the actual population of certain cities is much greater than official figures suggest, and this problem is likely to grow. 'In China we have to do face-to-face every time, as there are so few telephones. But people love answering questions.' Taylor Nelson compiles the TV ratings in China and has recently enlarged its sample to 12,200 households in 62 cities. ■

Provided that the environment of the street allows an interview to be conducted, the street interview is an excellent means to identify and interview individuals where there are poor or non-existent sampling frames. Given the congestion in many major international cities, the locations where interviews will be conducted have to be carefully selected. However, this does not differ from planning and conducting street interviews in 'home' markets. As with in-home and in-office interviews, the differences lie in the culture of approaching someone in the street and their willingness to divulge information to a stranger.

Mail interviews

Because of their low cost, mail interviews continue to be used in most developed countries where literacy is high and the postal system is well developed. Mail surveys are, typically, more effective in industrial international marketing research, although

it is difficult to identify the appropriate respondent within each firm and to personalise the address. This point was illustrated in the GlobalCash study where 6,000 telephone calls were made throughout Europe to establish the identity of the target respondent and to check the postal details.

The growth of Internet usage has also meant a growth in the use of electronic mail surveys. Given the global nature of the Internet, for certain topics (such as home shopping using the Internet) this may be the cheapest and quickest means to conduct an international survey. However, given the sampling problems associated with Internet usage, great care must be taken in validating samples from such surveys. As the penetration of Internet usage increases globally, the technique has great potential for both survey work and qualitative techniques; it is ideally suited to reaching out across national boundaries.

The criteria for the selection of survey methods were discussed in Chapter 10. As discussed and illustrated in that chapter, an important consideration in selecting the methods of administering questionnaires is to ensure equivalence and comparability across countries. This is illustrated in the following example, where comparability between countries is of paramount importance. The example leads on to issues of equivalence in measurement and scaling.

e x a m p l e

Does luxury have a home country?[28]

In the late 1980s, RISC, an international consulting company, decided to launch a study on luxury goods. Consequently, they developed a number of questions about luxury brands which were added to the questionnaire administered to national representative samples of major European countries. In 1992 they decided to explore, in France, Germany and Italy, the particular issue of country images in relation to luxury. As a result, 2,500 respondents in each country answered the question '*In your opinion, which country best understands the idea of luxury and reflects it through its products and brands*?'

Respondents could choose from a list of seven countries: Britain, France, Germany, Italy, Japan, Spain and the USA. They were then asked to select from the same list the country they spontaneously associated with 10 general product-related dimensions.

The fact that the same survey was conducted at the same time with exactly the same research design in three countries was important because it allowed the magnitude of response biases due to '*patriotism*', a factor often presented as an explanatory variable for country-of-origin effects. ■

Measurement and scaling

In international marketing research, it is critical to establish the equivalence of scales and measures used to obtain data from different countries.[29] As illustrated in Figure 26.2, this requires an examination of construct equivalence, operational equivalence, scalar equivalence and linguistic equivalence.[30]

Construct equivalence deals with the question of whether the marketing constructs (for example opinion leadership, variety seeking and brand loyalty) have the same meaning and significance in different countries. In many countries, the number of brands available in a given product category is limited. In some countries, the dominant brands have become generic labels symbolising the entire product category. Consequently, a different perspective on brand loyalty may have to be adopted in these countries.

Construct equivalence comprises conceptual equivalence, functional equivalence and category equivalence. **Conceptual equivalence** deals with the interpretation of brands, products, consumer behaviour and marketing effort. For example, sales pro-

Construct equivalence
A type of equivalence that deals with the question of whether the marketing constructs have the same meaning and significance in different countries.

Conceptual equivalence
A construct equivalence issue that deals with whether the interpretation of brands, products, consumer behaviour and the marketing effort are the same in different countries.
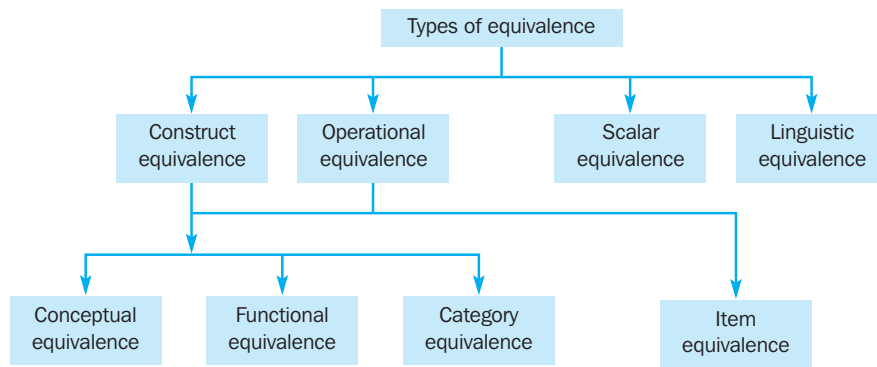
Figure 26.2
Scaling and measurement equivalence in international research

Functional equivalence
A construct equivalence issue that deals specifically with whether a given concept or behaviour serves the same role or function in different countries.

Category equivalence
A construct equivalence issue that deals specifically with whether the categories in which brands, products and behaviour are grouped are the same in different countries.

Operational equivalence
A type of equivalence that measures how theoretical constructs are operationalised in different countries to measure marketing variables.

Item equivalence
Use of the same instrument in different countries.

Scalar equivalence
The demonstration that two individuals from different countries with the same value on some variable will score at the same level on the same test. Also called metric equivalence.

Linguistic equivalence
The equivalence of both spoken and written language forms used in scales and questionnaires.

motion techniques are an integral component of marketing effort throughout Europe. On the other hand, in countries with shortage economies, where the market is dominated by the sellers, consumers view sales with suspicion because they believe that the product being promoted is of poor quality. **Functional equivalence** examines whether a given concept or behaviour serves the same role or function in different countries. For example, in many developing countries, bicycles are predominantly a means of transportation rather than of recreation. Marketing research related to the use of bicycles in these countries must examine different motives, attitudes, behaviours, and even different competing products from such research would in Europe. **Category equivalence** refers to the category in which stimuli such as products, brands and behaviours are grouped. In Europe, the category of the principal shopper may be defined as either the male or female head of household. This category may be inappropriate in countries where routine daily shopping is done by a domestic servant. Furthermore, the category 'household' itself varies across countries.

Operational equivalence concerns how theoretical constructs are operationalised to make measurements. In Europe, leisure may be operationalised as playing golf, tennis or other sports, watching television, or basking in the sun for example. This operationalisation may not be relevant in countries where people do not play these sports or do not have round-the-clock TV transmission. Lying in the sun is generally not normal behaviour in countries with hot climates. **Item equivalence**, which is closely connected to operational equivalence, presupposes both construct and operational equivalence. To establish item equivalence, the construct should be measured by the same instrument in different countries.

Scalar equivalence, also called metric equivalence, is established if the other types of equivalence have been attained. This involves demonstrating that two individuals from different countries with the same value on some variable, such as brand loyalty, will score at the same level on the same test. Scalar equivalence has two aspects. The specific scale or scoring procedure used to establish the measure should be equivalent. The equivalence of response to a given measure in different countries should be considered. For example, do scores from the top box or from the top two boxes on a purchase-intent scale reflect similar likelihood of purchase in different countries? Finally, **linguistic equivalence** refers to both the spoken and the written language forms used in scales, questionnaires and interviewing. The scales and other verbal stimuli should be translated so that they are readily understood by respondents in different countries and have equivalent meaning.[31]

Questionnaire translation

Translation into 10 languages[32]

Robert Kirby, who was the Telephone Research Director at Research Services Ltd and has many years' experience of pan-European telephone surveys, offers the following advice in coping with translation:

> *Translations require the most rigorous checking. In my experience it is not advisable to use someone from a home country who can speak the language fluently, because frequently a current idiomatic knowledge is required. A translation agency with no expertise could have the idiom but not the interviewing skills.*

His advice was valuable in the GlobalCash Project. The GlobalCash questionnaire was translated from English into Czech, Finnish, French, German, Greek, Hungarian, Italian, Polish, Portuguese and Spanish. In the project there was the added dimension of confusing technical terms with terms such as 'netting', 'pooling' and 'concentration', which have a day-to-day usage that is entirely different from the financial usage. The translation for the GlobalCash Project was undertaken by the business schools in the target countries that understood the idiom, the elicitation task required of finance directors in their country and the technical terms used. ■

e x a m p l e

Taking the people's temperature – right across Europe[33]

The continuous tracking survey of EU citizens' opinion of European issues uses 19 variants of their questionnaire, to include people living in countries where more than one national language is common. All the questionnaires have to go through extensive checking, translation, back translation, harmonisation and piloting in the short time between finally agreeing the questionnaire and starting work. ■

As in the above examples, questions may have to be translated for administration in different cultures. Direct translation, in which a bilingual translator translates the questionnaire directly from a base language to the respondent's language, is frequently used. If the translator is not fluent in both languages and is not familiar with both cultures, however, direct translation of certain words and phrases may be erroneous. Procedures like back translation and parallel translation have been suggested to avoid these errors. In **back translation**, the questionnaire is translated from the base language by a bilingual speaker whose native language is the language into which the questionnaire is being translated. This version is then retranslated back into the original language by a bilingual whose native language is the initial or base language. Translation errors can then be identified. Several repeat translations and back translations may be necessary to develop equivalent questionnaires, and this process can be cumbersome and time-consuming.[34]

An alternative procedure is **parallel translation**. A committee of translators, each of whom is fluent in at least two of the languages in which the questionnaire will be administered, discusses alternative versions of the questionnaire and makes modifications until consensus is reached. In countries where several languages are spoken, the questionnaire should be translated into the language of each respondent subgroup. It is important that any non-verbal stimuli (pictures and advertisements) also be translated using similar procedures. The following example underscores the importance of translation that does not impose the structure of a 'home country' from which the research is commissioned, in this case Britain.

Back translation
A translation technique that translates a questionnaire from the base language by a translator whose native language is the one into which the questionnaire is being translated. This version is then retranslated back into the original language by someone whose native language is the base language. Translation errors can then be identified.

Parallel translation
A translation method in which a committee of translators, each of whom is fluent in at least two languages, discusses alternative versions of a questionnaire and makes modifications until consensus is reached.

The search for focus – brand values across Europe[35]

Land Rover wished to understand the values associated with their brand in Belgium, Britain, France, Germany, Italy, the Netherlands, Portugal and Spain. The project involved the use of CAPI and required the development of a questionnaire that could allow comparisons between the countries. The issue of questionnaire translation was a sensitive and pertinent one. What was crucial was that the translation was meaningful in all languages rather than forcing the English language requirements into the local language. The translation was a difficult process and required an understanding of both the research process and the cultural context of the individual markets. It was the case that certain concepts that are expressive in English could not easily be translated into other languages. For instance, the word 'aspirational' does not easily translate into French, Spanish or Italian with the same meaning as in English. Consequently, in order to ensure comparability the translation process involved a series of iterative steps changing the master questionnaire to fit what was achievable overall in each market. ■

Ethics in marketing research

Ethical responsibilities for marketing research conducted abroad are very similar to those for research conducted domestically. In conducting marketing research across Europe (and indeed globally), ESOMAR produce a code of conduct that guides professional practice that protects the interests of all research stakeholders. In general, for each of the six stages of the marketing research design process, the same four stakeholders (client, researcher, respondent and public) should act honourably and respect their responsibilities to one another. For individual countries, a key development in honourable practice lies in the development of professional associations. These can guide researchers through the development of specific codes of practice that may reflect their culture and industrial heritage. For example, in the Czech Republic, Hungary, Poland and Russia, major market research suppliers are forming associations and introducing or discussing quality standard systems for fieldwork. Research was not unknown in the region prior to the collapse of communism and most countries have had a tradition of government-sponsored research spanning 40–50 years. Most major suppliers are investing in their future by training their staff in Western Europe and North America.[36]

For all the similarities that may exist in conducting research in international markets, some ethical issues become more difficult to solve. Consider the dilemma presented by David Mendoza, Director of Gobi International, an Anglo-Russian industrial research company: 'Some people will, Soviet style, still refuse to answer any questions of any kind. Others, often the same type of former Soviet apparatchik, but totally disillusioned, will promise to tell you whatever you need in return for dollars. Frankly, this is highly unsatisfactory and something we try to avoid, although often there is no choice'.[37]

Internet and computer applications

The Internet and computers can be extensively used in all phases of the international marketing research process. These uses parallel those discussed in earlier chapters and hence will not be repeated here. The fact that the Internet can be used to communicate with respondents and marketing decision-makers anywhere in the world has given a new dimension to international marketing research. For example, the online survey overcomes geographic boundaries and differences in postal systems to solicit responses from around the world. The online survey also takes advantage of one interviewer (the computer) that can present the same survey in several different translations.

The environmental characteristics of international markets detailed in this chapter present a formidable research task, especially when first learning about a country. The Internet is the ideal means to quickly access material that helps to shape an understanding of the environmental context of a particular country. This is particularly relevant in tracking down secondary data sources that may not be available in the home country of a researcher.

The Internet has also been extremely beneficial in conducting qualitative interviews. One of the most difficult administrative tasks in conducting focus groups is getting participants together at an agreed location. Because of the travel problems for certain respondents, many focus groups may have to be conducted to cover a wide geographic area. The Internet can help to overcome these problems, first of all by the convenience it offers the participants – not having to leave their homes – and second by the breaking down of geographic boundaries.

Summary

With the globalisation of markets, international marketing research is burgeoning rapidly. As well as the technical requirements of conducting successful marketing research as outlined in this text, the researcher has to cope with new cultures and languages in targeted international markets. Given the array of ethnic groups within most European countries, the challenges of understanding new cultures and languages can exist within a 'home' country.

The environment prevailing in the international markets being researched influences all six steps of the marketing research process. Important aspects of this environment include the marketing, government, legal, economic, structural, informational and technological, and socio-cultural environment.

In collecting data from different countries, it is desirable to use techniques with equivalent levels of reliability rather than the same method. Repeating identical techniques across borders may result in subtle cultural and linguistic differences being ignored, which may have a great effect upon the nature and quality of data that are generated. It is critical to establish the equivalence of scales and measures in terms of construct equivalence, operational equivalence, scalar equivalence, and linguistic equivalence. Questionnaires used should be adapted to the specific cultural environment and should not be biased in favour of any one culture or language. Back translation and parallel translation are helpful in detecting translation errors.

The ethical concerns facing international marketing researchers are similar in many ways to the issues confronting domestic researchers. Working in countries where professional marketing research associations exist should allow access to codes of conduct that reflect the culture and industrial heritage of that country. However, the international researcher being exposed to an array of cultures and economic scenarios should always expect to find new situations and ethical dilemmas.

Questions

1 Evaluate the meaning of 'international' from the perspective of the marketing researcher.

2 What characteristics distinguish international marketing research from domestic marketing research?

3 Describe the aspects of the environment of each country that should be taken into account in international marketing research.

4 Describe the importance of considering the marketing environment in conducting international marketing research.

5 What is meant by the structural environment? How do the variables comprising the structural environment influence international marketing research?

6 What is meant by the informational and technological environment? How do the variables comprising the informational and technological environment influence international marketing research?

7 What is meant by the socio-cultural environment? How do the variables comprising the socio-cultural environment influence international marketing research?

8 How should the researcher evaluate secondary data obtained from foreign countries?

9 Describe the factors that may influence the approach to qualitative techniques in different countries.

10 Select a country, using environmental characteristics to illustrate why CATI works particularly well as a survey technique.

11 Select a country, using environmental characteristics to illustrate why in-home interviewing does not work particularly well as a survey technique.

12 Select a country, using environmental characteristics to illustrate why Internet surveys work particularly well as a survey technique.

13 How should the equivalence of scales and measures be established when the data are to be obtained from different countries or cultural units?

14 What problems are involved in the direct translation of a questionnaire into another language?

15 Briefly describe the procedures that may be adopted to ensure translation is correctly conducted in the development of a questionnaire.

Notes

1 Goodyear, M., Guest editorial – 'International Research and Marketing', *Journal of the Market Research Society* 38(1) (January 1996), 1.

2 See Malhotra, N.K., 'Administration of questionnaires for collecting quantitative data in international marketing research', *Journal of Global Marketing* 4(2) (1991), 63–92; and Malhotra, N.K., 'Designing an international marketing research course: framework and content', *Journal of Teaching in International Business* 3 (1992), 1–27.

3 Petrin, K., 'Tasty wisdom for ESOMAR delegates', *Research* (1 December 2001).

4 Savage, M., 'Breaking down the great wall', *Research* (October 1999), 14.

5 Birks, D.F., 'Researching cash management practices', in Birks, D.F. (ed.), *Global Cash Management in Europe* (Basingstoke: Macmillan 1998), 39.

6 Sills, A. and Desai, P., 'Qualitative research amongst ethnic minorities in Britain', *Journal of the Market Research Society* 38(3) (July 1996), 247.

7 Goodyear, M., 'The world over, a group is a group is a group', *ResearchPlus* (November 1992), 5.

8 Doctor, V., 'Levering the market', *Research* (1 August 2001).

9 Dunlop, C., 'Warming responses from the Levant to the Gulf', *Research* (November 1995), 24.

10 Savage, M., 'The East gets out of the red and into the black', *Research* (October 1997), 22.

11 Hutton, G., 'If you board the Asian 'bus, better mind your language', *ResearchPlus* (February 1996), 9.

12 *Ibid.*

13 Lau, S., 'I want my MTV, but in Mandarin, please', *Admap* (March 2001), 34–6.

14 Pasco, M., 'Euro-youth: myth or reality?', *Admap* (June 2001) 14–15.

15 Miles, L., 'Finding a balance in global research', *Marketing* (29 November 2001), 33.

16 See Dawar, N. and Parker, P., 'Marketing universals: consumers' use of brand name, price, physical appearance, and retailer reputation as signals of product quality', *Journal of Marketing* 58 (April 1994), 81–95.

17 Wicken, G. and Koneva, E., 'Rise, rise, you Russian middle-classes', *ResearchPlus* (January 1996), 5.

18 Broadbent, K., 'When East meets West, quite what does "qualitative" mean?', *ResearchPlus* (March 1997), 15.

19 Choudry, Y.A., 'Pitfalls in international marketing research: are you speaking French like a Spanish cow?', *Akron Business and Economic Review* 17(4) (Winter 1986), 18–28.

20 Zimmerman, A.S. and Szenberg, M., 'International primary market research techniques: availability, effectiveness and relative costs', *Proceedings of the International Trade and Finance Association*, Vol. III (San Jose, Costa Rica, May 1995), 773–806.

21 Bloom, N., 'In-depth research? A bit controversial!', *ResearchPlus* (June 1993), 13.

22 Zimmerman, A.S. and Szenberg, M., 'Implementing international qualitative research: techniques and obstacles', *Qualitative Market Research: An International Journal* 3(3) (2000), 158–64.

23 The work on survey methods is drawn from Malhotra, N.K., 'Administration of questionnaires for collecting quantitative data in international marketing research', *Journal of Global Marketing* 4(2) (1991), 63–92.

24 Davies, R.W.B., Minter, C.J.W., Moll, M. and Bottomley, D.T., 'Marketing research in Hong Kong', *European Research* (May 1987), 114–20.

25 De Houd, M., 'Internationalized computerized telephone research: is it fiction?', *Marketing Research Society Newsletter* 190 (January 1982), 14–15.

26 Mendoza, D., 'Russian roulette: exciting once you learn new rules', *Research* (January 1995), 20.

27 Gander, P., 'Foreign policy', *Marketing Week* (30 April 1998), 46.

28 Dubois, B. and Paternault, C., 'Does luxury have a home country? An investigation of country images in Europe', *Marketing and Research Today* (May 1997), 80.

29 See also Min-Han, C., Lee, B.W. and Ro, K.K., 'The choice of a survey mode in country image studies', *Journal of Business Research* 29 (February 1994), 151–62.

30 See Bhalla, G. and Lin, L., 'Cross-cultural marketing research: a discussion of equivalence issues and measurement strategies', *Psychology and Marketing* 4(4) (1987), 275–85. A similar discussion is also found in Douglas, S.P. and Craig, C.S., *International Marketing Research* (Englewood Cliffs, NJ: Prentice Hall, 1983).

31 Andrews, J.C., Durvasula, S. and Netemeyer, R.G., 'Testing the cross-national applicability of U.S. and Russian advertising belief and attitude measures', *Journal of Advertising* 23 (March 1994), 17–26.

32 Kirby, R., 'A pan-European view of the executive at lunch', *ResearchPlus* (October 1992), 7.

33 Phillips, A., 'Taking the people's temperature – right across Europe', *ResearchPlus* (November 1996), 7.

34 For an application of back translation, see Wharton, R., Baird, I.S. and Lyles, M.A., 'Conceptual frameworks among Chinese managers: joint venture management and philosophy', *Journal of Global Marketing* 5(1–2) (1991), 163–81.

35 Bull, N. and Oxley, M., 'The search for focus – brand values across Europe', *Marketing and Research Today* (November 1996), 243.

36 Bartonova, M., 'The markets are emerging – and research is hard on their heels', *ResearchPlus* (January 1996), 4.

37 Mendoza, D., 'Russian roulette: exciting once you learn the new rules', *Research* (January 1995), 21.

Business-to-business (b2b) marketing research

Objectives

After reading this chapter, you should be able to:

1 explain what b2b marketing is;

2 understand the differences between b2b and consumer marketing;

3 explain the marketing research challenges that emerge from the differences between b2b and consumer marketing decision making;

4 understand how the concepts of networks and relationships underpin the distinctive challenges faced by b2b decision makers;

5 understand what competitor intelligence is and how it may supplement or even take the place of marketing research for some businesses;

6 discuss the cultural challenges faced in conducting b2b marketing research in an international setting;

7 discuss the ethical considerations in b2b marketing research by comparing the challenges set in competitor intelligence with conventional marketing research.

The management of customer relationships is the core task of the b2b marketer. The marketing researcher must help to unlock the nature and dynamism of these relationships.

Overview

This chapter starts by defining the nature of b2b marketing, presenting a definition that introduces the concept of 'relationships'. This is followed by a short introduction to the differences between b2b and consumer marketing. The evaluation of differences is crucial if one is to consider whether these have any impact upon the practice of b2b marketing research. The chapter then sets out the bases in which to evaluate the key challenges faced by b2b marketing decision-makers. These bases describe the importance of nurturing and developing relationships and the use of networks. From these bases, the key differences between b2b and consumer marketing are explained in more detail. There are five key areas of difference that are explored, followed in each case by the marketing research challenges these differences present. The 'threat' to conventional marketing research is then explored with a review of the nature and use of competitor intelligence. The challenges faced by b2b marketing researchers in an international setting are then evaluated, discussing how language and cultural differences impact upon gaining access to businesses. This is followed by a discussion of the ethical issues raised in implementing competitive intelligence compared with conventional marketing research. The chapter finally evaluates the implications of using the Internet in b2b research, especially the benefits of briefing and supporting fieldworkers when they have been given the task to interview respondents about technical questions.

We begin with two examples illustrating how marketing research may be used in b2b marketing. The first example demonstrates the use of quantitative and qualitative techniques in an international setting. The case could well refer to any consumer marketing organisation, investing in the use of exhibitions and wishing to evaluate the worth of that investment. It basically illustrates that all the techniques that have been presented and discussed in this text can be used in a b2b context. It could also be used to support the argument that there is no difference between marketing research in b2b contexts compared with consumer marketing, as the case clearly shows that all techniques can be utilised. What it does not reveal are the many challenges facing b2b marketing researchers when generating information in which b2b decision-makers will have confidence.

Fair's fair for Ericsson[1]

As the world's largest supplier of mobile systems and a leading player in fixed line networks, Ericsson operates in more than 140 countries. It provides total solutions from systems and applications to services and core technology for mobile handsets. Its customers are network operators from around the world; indeed, the top ten largest mobile operators are all customers of Ericsson. Telecoms companies like Ericsson recognise that trade shows play a crucial role in building the brand as well as providing showcases to demonstrate new products and technologies. They also provide the means to nurture and develop relationships by meeting and entertaining customers and other parties.

Ericsson has a clearly defined strategy for participation at trade shows and typically takes part in the 3GSM World Congress (France), CTIA (USA), CeBIT (Germany), CommunicAsia (Singapore) and ITU Telecom Asia (Hong Kong). Given the high levels of expenditure to participate and the importance of this communication channel, Ericsson determined that they needed to ensure that:

- the return on investment in a show could be measured;
- best practice guidelines were developed;
- the Ericsson brand was being communicated consistently.

In order to support the decisions involved in these areas, Ericsson used marketing research to provide feedback in five key areas:

■ the 'quality' of visitors;
■ their performance at an individual show;
■ the impact on customer/employee relationships;
■ the impact on the brand;
■ the value of participation relative to other communication tools.

Working with the research company BPRI, an array of research techniques were used. Depending upon the nature of their target audience, CAPI, ethnography, telephone, Web and self-completion surveys were used. The CAPI approach was used daily and enabled them to capture findings from the previous day and to make changes to the visitor experience. For example, at the beginning of the 3GSM World Congress, some concern was expressed about the processes used for directing visitors to key Ericsson staff, as well as about the need for a more tailored approach. The daily briefings by the BPRI team highlighted these areas of concern and alternative solutions to these issues were implemented on the same day. Marked improvement was subsequently reflected in much higher scores for these aspects of performance. ■

The second example comes from the many technical challenges in the GlobalCash study. The following serves as a reminder of the GlobalCash 'journey' throughout this text. It may come as a surprise to recall that all of the GlobalCash examples used in the previous 26 chapters to illustrate different techniques and challenges faced by marketing researchers have been set in a b2b context. Again, the whole array of research techniques have been used, which may add further weight to the belief held by many that there are no differences in marketing research for b2b or consumer marketing.

example

GlobalCash Project

Supporting b2b decision making

The GlobalCash study was designed to support b2b decision-makers. The array of techniques used in the study included:

■ a traditional postal survey to the largest companies in Europe;
■ a telephone survey to these companies prior to the launch of the euro;
■ an Internet survey regarding company responses to the introduction of the euro;
■ qualitative in-depth interviews with key decision-makers;
■ a feedback workshop in Brussels where survey findings were presented to respondents, followed by qualitative workshops based on group discussions.

These techniques were targeted at Chief Executives, Finance Directors and Cash Managers. The decision-makers in major banks from Britain, France, Germany, Ireland, Spain and the USA that required decision support included: Customer Relationship Managers, Systems Designers, Brand Managers, Marketing Managers and Strategists.

Selections of the areas of decision support, taken from the initial proposal, include:

■ **Understand your market and your consumers** 'Obtain an increased understanding of what the most demanding customers are looking for in terms of cash management products, services, systems and quality standards across Europe.'
■ **Benchmark yourself** 'Learn what major companies in each country and across Europe think about their banks and related service providers and the transactional banking services they use.'
■ **Identify market opportunities** 'Understand the strengths and weaknesses of products and services as perceived by the leading companies in Europe.'
■ **Product development opportunities** 'Introduce new cash management strategies based on an in-depth understanding of present and future market needs and competitor activity, strengths and weaknesses.' ■

The main challenge in the GlobalCash study lay in gaining access to such prominent and busy executives. Over the 10 years that the study unfolded, different research techniques were tried out and developed to draw out the most meaningful, valid and reliable information from the target executives. The GlobalCash researchers were fortunate to be able to repeat the study every two years and to learn very valuable lessons in implementing the next study. The majority of b2b marketing researchers do not have this luxury, having to do their best in a one-off situation, learning from their experiences and the case material of b2b research practices in many specific contexts.

There are many marketing researchers who believe that b2b marketing research is quite distinctive from consumer marketing research. There is a professional group, the Business Intelligence Group (BIG) **www.b2bresearch.org**, that was lauched in 2001 to support the interests of individuals involved in b2b marketing research. We will progress by evaluating whether there are distinctive characteristics of b2b marketing decision making that may require a different approach to marketing research. To begin such an evaluation we continue by evaluating the nature of b2b marketing and what are contended to be the similarities and distinctions when compared with consumer marketing.

What is b2b marketing?

b2b marketing is a relatively new term used to describe business marketing, industrial marketing or even organisational marketing. It is a simple acronym for the term business-to-business marketing that was initially used to distinguish online business transactions, but is now used as the generic term for business marketing. In very simple terms b2b means businesses that sell products or services to other businesses, such as Intel selling processors to Dell or other computer manufacturers.

Such a simple term masks the challenges faced by b2b marketers. In the following definition, Dibb *et al.*[2] highlight the activities involved in the process of exchange between two parties. This definition presents little distinction between b2b marketing and the generic process of marketing, as discussed on pages 4 and 5 in Chapter 1.

> *Industrial [b2b] marketing concerns activities directed towards facilitating and expediting exchanges between industrial markets and industrial producers.*

In the second definition, Ford *et al.*[3] present a view of b2b marketing that goes beyond the more traditional marketing notions of exchange as presented above.

> *Business [b2b] marketing is the task of selecting, developing and managing customer relationships for the advantage of both customer and supplier, with regard to their respective skills, resources, technologies, strategies and objectives.*

In this definition the key element to emphasise is the task of managing customer relationships. This can be seen as a process that goes beyond the facilitating and expediting exchanges of goods, although it could be argued that the facilitation of exchanges could include the management of relationships. Nonetheless, Ford *et al.*'s emphasis upon the management of relationships forming the essence of b2b marketing could lead to the case that there are more differences than similarities between b2b marketing and consumer marketing.

Introducing the distinction between b2b and consumer marketing

There are many products and services which may be clearly identified as b2b transactions, such as electrical components used in the manufacture of earthmoving equipment. Other products and services, which at first glance may appear to be b2b transactions, may also be sold to consumer markets. For example, electrical components used in car manufacturing may be sold direct to consumers who wish to service, repair or even transform their own cars. Other more obvious examples of products and services which may be bought by both businesses and consumers include: cars, private health insurance, meals in restaurants, laptop computers and stationery equipment – the list could go on! Littler[4] contends that the key factor that distinguishes between b2b and consumer transactions is *the purpose for which they are sold*. In b2b transactions, he argues that products are bought for resale or for the production of other goods, and services are bought to facilitate the resale and production of goods. In consumer transactions, products and services are generally bought for the self, the household or as gifts. This distinction clearly does not rule out consumers reselling products, as any glance in the classified columns of newspapers for 'unwanted Christmas presents' or a search on eBay (**www.ebay.com**) would quickly reveal!

If there are different purposes for which goods and services are sold in the two sectors, the approaches to marketing them could differ quite significantly, the nature of target markets may be fundamentally different, and the marketing techniques that communicate and generate sales could differ. However, the underlying reason for such a difference may be simply that business consumers tend to be seen as rational buyers, whilst consumers tend to be seen as more intuitive and emotive. In examining the differences between business and consumer purchases, Wilson[5] argues that it is difficult to dispel the persistent images of professional committees deciding proactively on organisational purchases, and wilful or manipulated individual consumers engaging in semi-rational or impulsive buying behaviour. Such images can be defended if one is circumspect in choosing examples to illustrate the argument, from either perspective. For example, if one were to contrast the buying decision for a train manufacturer selecting the supplier of braking mechanisms compared with a consumer buying chocolates, there would very different time frames and stages to the decision-making process. If one were to compare the differences between the two sectors in the decision to buy a Christmas hamper as a gift, the time frame and stages to the decision-making process may be very similar. The purpose behind the decision to buy a Christmas hamper may also be very similar. There are also many examples of consumer purchasing that may be very 'professional' in terms of knowledge of prices, quality, source reliability, product availability and product applications which is not only highly professional but can make the difference between competitive survival or failure – in the sense of sustaining accustomed standards of organisational (i.e. family) performance[6]. There are clearly many grey areas between a notion of business purchasing = professional and consumer purchasing = emotive and impulsive. The task of setting out clear differences is difficult given the breadth of products and services that are purchased by businesses and consumers and the great overlap in the nature and intention of the purchase. One of the foremost industrial marketing writers, Frederick Webster,[7] recognised that many of the differences may be only by degree, but that there are substantial differences:

> *If we are to understand and intelligently attack industrial marketing problems, a number of substantial differences between industrial and consumer marketing must be recognized. While it may be true that these are often differences of degree rather than kind, the degrees of difference are substantial.*

In examining the similarities and differences between b2b and consumer marketing, the strength of the argument lies in the examples chosen to make the case for the difference to be 'substantial'. These differences may not be absolutes; there may always be specialised b2b and consumer markets that can be used to counter-argue a point. The emphasis may be upon the propensity or tendency for b2b and consumer marketing problems to be different. The 'tendencies' for b2b purchases developing distinctive characteristics do have implications for the b2b marketing researcher. Before examining these implications, two factors at the heart of the defence for the distinction of b2b marketing need addressing, that is **networks** and **relationships**. These factors are founded in Ford *et al.*'s definition of business marketing and are the basis for the contention that there are fundamental differences between b2b and consumer marketing, and thus differences in the approaches and challenges faced by b2b and consumer marketing researchers.

Bases for the distinction of b2b marketing research

The network

Ford *et al.*[8] argue that the transactions between a business supplier and its customers are not isolated events that are unrelated to each other. Instead, each of the transactions takes place in a relationship and each is related to the previous experience of the companies and what they plan for the future. Business marketers have relationships with their suppliers and these relationships are vital if the marketer aims to build strong relationships with their customers. Each of their suppliers will have relationships with other suppliers and with their customers. All of these relationships are affected by the other ones and are intertwined in a network, and it is in this network that the b2b market operates. The distinctive element of seeing b2b marketing as a series of relationships in a network is that, if the marketing researcher wishes to understand the behaviour of an individual company, then they have to understand its relationships with other companies. Businesses are not seen as free agents, able to develop and implement their strategy in isolation; each is dependent upon the other in order to act and react to the strategies of others. In other words, the basic assumption of network thinking is that 'no business is an island'.[9]

Figure 27.1 illustrates a distribution network from the perspective of a large supplying company: IBM. The figure demonstrates the different ways that computers can reach customers in Italy.

In order to understand a network, a focal relationship is usually chosen. By choosing a particular relationship the network can be examined from that perspective. As soon as the focal relationship changes, a new set of relationships emerges. By examining a network carefully a b2b marketer can achieve a realistic and balanced understanding of their own position in the minds of customers and suppliers. It is very important to evaluate the network from different focal points as drawing the network from the perspective of just one business can lead it to believe that it is more important to others than it actually is. The existence of a network should not cloud the fact that the key task of b2b marketing is to manage each single relationship. It is through its relationships that a business learns and adapts to its surrounding network; that a business exploits and develops its own abilities and gains access to others, and it is also through relationships that a business can influence different businesses elsewhere in the network.

Relationships

Ford *et al.* use the term 'relationship' to describe 'the pattern of interactions and mutual conditioning of behaviours over time, between a company and a customer, a supplier or other organisation'.[10] The time dimension of a relationship requires decision-makers to shift their emphasis away from each single discrete transaction, towards tracking how things unfold in the relationship over time and changing them
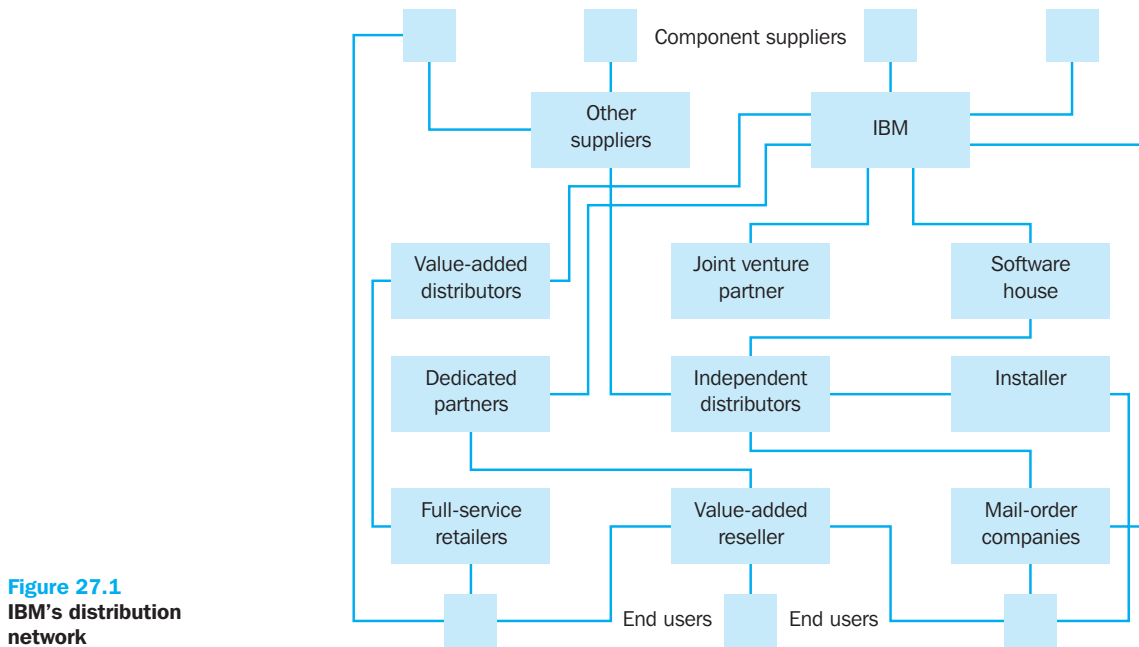
Figure 27.1
IBM's distribution
network

when appropriate. Relationships with business customers are likely to be complex. A business's relationships with its suppliers, development partners and financial institutions that constitute its network all depend upon its relationships with its customers and on solving their problems in those relationships. Ford *et al.* make the clear distinction between the growth of Customer Relationship Management (CRM) systems and the stream of books and articles dealing with 'relationship marketing'. They contend that the idea that underpins most relationship marketing literature is that relationships are the creation of the marketing company. Such views do not help b2b marketers who have to cope with:

■ The history of their relationship and the impact of those events upon how different individuals in separate businesses react to each other.
■ Relationships that are not necessarily positive through coping with differences in aims and understanding that may result in conflict as well as cooperation.
■ Both businesses interacting with each other, both attempting to manage the relationship in a way that they think appropriate.

The b2b relationship is not a relationship that is managed in a unilateral manner, i.e. a relationship that the marketer can 'drive'. CRM systems that hold the details relating to every customer, that are typically used to ensure all interactions with the customer are consistent and knowledge driven, are typically 'unilateral'. Though CRM systems may measure and model characteristics of transactions that develop over time, they are electronic observations of measurable characteristics of those transactions. CRM systems are not a means to acquire a realistic understanding of the rich and qualitative complexities of continually developing b2b relationships.

In the context of understanding networks and relationships, the prime objective of marketing researchers in business markets therefore is to establish and support an understanding of how the network operates and how successful relationships evolve. There is a need to understand how networks develop and how businesses and individuals in it relate to each other. By being able to recognise patterns of behaviour in the network, businesses should gain some guidance on how they should act and react. To develop such an understanding of patterns requires more than analysis in the traditional sense of gathering and interpreting the available data on patterns of sales and market shares. It means

capitalising on the tacit knowledge inside the business and in other networked businesses. Tacit knowledge in this sense could mean, for example, understanding the psychological barriers that may exist in managers when presented to new offerings that evidently offer better monetary value than they get from their present suppliers. They may have worked with their present suppliers for many years and developed economically and socially together; what price would they pay to sever that relationship?

Faced with such a challenge in understanding networks and relationships, the b2b marketing researcher has all the techniques as detailed and illustrated in this text at their disposal. Having the techniques available is one thing; implementing them is another. The challenge for the b2b marketing researcher is to creatively adapt and develop the techniques to overcome the obstacles in the b2b environment. In order to understand the challenge set for the b2b marketing researcher we now examine the implications they face of the differences between business and consumer purchases.

Implications of the differences between business and consumer purchases for the marketing researcher

The differences between b2b and consumer purchases have implications for the approach, research design and individual data gathering and analysis techniques that may be employed by marketing researchers. We now examine the nature of the differences between business and consumer marketing and the challenges faced in applying marketing research techniques using the framework of arguments as presented by Ford et al.[11] As we work through the individual challenges you will also see how these challenges become interrelated. The five key differences are summarised in Figure 27.2 and then individually evaluated.

The number of people who may be involved in a b2b purchasing decision

The concept of the 'buying centre' is well established in business purchasing. Collective decisions are more typical of complex, expensive or controversial purchasing decisions in larger businesses, whereas routine decisions and/or decisions made by smaller organisations are typically made by individuals rather than by groups.[12] However, even in cases of decisions apparently made by an individual, there could still be many other roles of influence involved in the phases of the purchasing process before and after the actual decision itself.

Webster and Wind[13] famously identified five buying 'roles' within the context of 'buying centres'. We can illustrate these roles by using an example of a b2b scenario involving an FMCG business buying Marketing Research services!

Figure 27.2
Summary of differences between b2b and consumer marketing

Key differences between b2b and consumer marketing	b2b marketing tends towards	Consumer marketing tends towards
1 Number of people involved in purchasing decision	High	Low
2 Professionalism of those involved in the purchasing decision	More rational	More emotive
3 Time taken to negotiate from decision to buy to actual purchase	Lengthy	Short
4 Importance of individual consumer to supplier	High	Low
5 Relationship development	Joint negotiation	Unilateral

1   *Users*. This may be a Brand Manager seeking marketing research support for the decisions faced in the launch of a new product.
2   *Influencers*. This may be a Human Resources Director who has used a range of research organisations for staff surveys over the past two years.
3   *Buyers*. This may be a Purchaser who is responsible for drawing up a research brief, contacting research organisations, gathering research proposals and managing the negotiation process.
4   *Deciders*. This may be the Marketing Director if the budget for project exceeds a set amount.
5   *Gatekeeper*. This may be another Brand Manager who may be supportive or down-right secretive and may help or hinder in feeding information and the interpretation of that information.

Further categories have been suggested such as those of:

6   *Initiator*.[14] This may be a more Senior Manager to the Brand Manager that recognises the decision problems faced by the Brand Manager/user and encourages them to purchase marketing research support or even champions their case for marketing research support.
7   *Analyst*.[15] This may be an individual who has the technical ability to critically evaluate the offerings and associated costs as presented in the proposals from competing research agencies.
8   *Spectator*.[16] This may be the Brand Manager of a competing brand who may react in the way the Brand Manager/User desires – 'sheer panic' at the thought that the competition is employing a major ethnographic study to support the development of a new integrated communications strategy.

This example shows eight people involved in the decision. In some organisations, these roles may be performed by one person, in others there may be numerous influencers and gatekeepers, sometimes working in different managerial functions, hierarchical levels and geographical locations. The people in the 'buying centre' will have different needs because of their different responsibilities, so a buyer is likely to be more interested in price and contractual details than a user is. Behind these roles there are individuals who will have their own perceptions, expectations and objectives based on their personalities and backgrounds.[17] The marketing researcher needs to help decision-makers measure and understand these individuals and their inter-relationships.

Marketing research challenges. The main challenge of the number of people involved in a business purchase is one of sampling. In a team such as that illustrated above, who would be the target **element** when putting together a sampling plan? Even if there were sampling frames that listed the managerial roles of individuals in an organisation, it is highly improbable that they will be listed by their roles in a 'buying centre'. The views of the marketing researcher Ruth McNeil illustrate this point.

example

Open for business?[18]

Sampling is becoming ever more demanding. Clients are requiring specialist samples, targeting particular people with particular job titles or skill sets. Such detailed specifications require the fieldwork company to know both where to find and how to access good lists. The absence of good, globally relevant sampling frames is still an issue. Given that SIC codes are rapidly becoming out of date, the increasingly specific demands for respondent type mean that a b2b researcher's skill now lies as much in finding the right respondent as in conducting and analysing a good interview. ■

The above example illustrates something of the 'detective' skills in tracking down the right respondent for a survey or interview. Even when we have tracked down the right individual, the challenge of getting them to respond in the way we would like is a great challenge. This challenge is developed in the next difference between b2b and consumer purchases discussed below.

The people involved in a b2b purchase are professional

The concept of the professional buyer is well established in business purchasing. Businesses use professional buyers to make economically sound purchases that support their strategic and operational decisions. Such professionals establish auditable systems of purchasing to cope with relatively cheap and routine goods through to expensive, complex and risky one-off products and services such as the purchase of a new accounting information system in a multinational company. Purchasing professionals in essence aim to manage b2b transactions in a rational and economic manner. The following example briefly describes the nature and mission of the professional body that represents purchasing professionals.

e x a m p l e

The professional buyer – a Chartered Purchaser

The Chartered Institute of Purchasing and Supply (CIPS) is an international education and qualification body representing purchasing and supply chain professionals. CIPS exists to promote and develop high standards of professional skill, ability and integrity among all those engaged in purchasing and supply chain management. It is the largest organisation of its kind in Europe and a central reference point worldwide on matters relating to purchasing and supply chain management. Its Professional Code of Ethics is the model for the international code and the domestic codes of many countries. The Institute acts as a centre of excellence for the whole profession of purchasing and supply chain management. ■

A body of professionals with rational, economic and ethical practices and that are auditable in a business should make the research challenge of measuring their practices relatively easy when compared with emotive and fickle consumers. However, even assuming that it is the professional buyer that makes the purchasing decision, one has to consider why they should divulge or share their practices with marketing researchers.

Marketing research challenges. The first major challenge of the people involved in a business purchase being professional is one of **access**. Access can be described as working at two levels.

1 *Initial contact:* getting to meet a target respondent in the first instance so that they understand the nature, purpose and benefits of taking part in an interview or completing a questionnaire.
2 *True feelings:* assuming that a researcher has managed to persuade a target respondent to participate, are they able to access what they really feel or think about the research subject? How guarded or sensitive may the respondent be in responding to the set questions and probes? Will they be responsive to research techniques that seek to uncover issues that are difficult to conceive and/or express?

Even when one has gained access to a target respondent, the challenge still remains to gain access to what they really feel or think about a subject.

The example on page 32 illustrated how increasingly difficult it is to gain access to Chief Executive officers, especially in a b2b context. It illustrates the despair faced by marketing researchers when they are briefed to interview managers who the researchers know have been targeted in many other

studies, and they realise the difficulty or impossibility of gaining access to them, at both levels of access described above. The views of the marketing researcher Richard Field further illustrate this point.

Is nothing straightforward in b2b research?[19]

Organisational changes are making it increasingly difficult to complete business research. It is, for example, harder to contact business executives. Routes to them are less likely to be through a secretary or even a switchboard. Phone numbers are more likely to be direct, unlisted or routed through a central department or an automatic exchange that is immune to the persuasive interviewer. On a wider level, researchers have also had to contend with both the seemingly inexorable conglomeration of the business world and policy changes on compliance. Increasingly, corporate policy forbids interviews – even in corporations that need, and commission, research themselves. In areas where there is a restricted pool of potential respondents, one cannot afford low strike rates. It doesn't just take longer to complete the quotas; one just runs out of contacts. ■

The second level of access ('true feelings') also presents major challenges for the b2b marketing researcher. The issue raised by professional purchasers at this level is whether they actually behave in a rational and economic manner and whether they understand or are willing to discuss any emotive or subjective reasons for their purchasing behaviour. The following quote illustrates that, although business purchasers may be professional and strive to be rational in economic terms, the reality is somewhat different.

> *No organisation has the resources (in terms of time, energy, money and brain capacity) to look at all the possibly relevant information about all the competitive alternatives available.*[20]

At some point in the evaluation of potential alternatives, professional purchasers will determine that they have considered enough information to be able to make a decision. Professional purchasers may strive for and project an image of rationality. The marketing researcher Neil McPhee[21] does not see it this way:

> *There is a strand of business teaching that has put forward a model that decision-making in corporates is logical and unemotional. This has never been true and is not true now. Research respondents remain human, even when they are at work, and human emotions have a habit of clouding rationalism.*

The following example illustrates the techniques used by Insight Research Group to clarify the issues that may cloud rationalism in b2b purchases.

A spoonful of research helps the medicine go down[22]

The only difference between doctors and other consumers is that this is a classic b2b market, in which customers are not in themselves the end consumers. However, they still have a set of beliefs about the value of a brand for themselves and their patients. It is imperative therefore that researchers get beyond doctors' rational and logical outer shell. They need to find a robust way to underpin functional data with the kind of emotionally driven information that could be used to feed into differentiated pharmaceutical products. Having long known that doctors were adept at assuming and maintaining a professional, logical 'distance' in research, the Insight Research Group began looking at how this behaviour might affect the depth of the overall findings. The often thoughtful, cogent, technical and articulate responses from doctors were in themselves something of a barrier to getting at the more fundamental drivers and triggers for prescribing which were required to make really compelling campaigns. This led to the development of a research approach which placed a greater emphasis on non-direct ques-

tioning, increased observation and interpretation of materials, generated from purpose designed exercises. In small workshops, doctors participated in group and individual work that helped to dismantle the 'doctor' behaviour and facilitate access to their deeper 'sensing' levels. At the analysis stage, the specific sorts of communication, imagery, language and even tone that would help the brand to trigger response at both rational and emotional levels became the focus. ■

It could be quite easily argued that such research approaches are commonplace in consumer marketing research; there is nothing distinctive in such an approach for b2b marketing researchers. What makes it distinctive is the juxtaposition of the access issues discussed above, i.e. getting to interview professional purchasers in the first instance and, if access is gained, getting them to cooperate – with the implementation of qualitative techniques. The following example from the GlobalCash project illustrates the difficulties and expense involved in making such approaches work.

<table>
<tr><td>

e x a m p l e

GlobalCash Project

</td><td>

Engineering the group discussion

Trying to bring together finance directors and cash managers from the largest companies in Europe to participate in focus groups could be viewed as an impossible task. The logistics of bringing together such senior managers, the rewards they may demand for attending such an event and the question of how much they would reveal to other managers and researchers in such a forum make the whole technique questionable. The GlobalCash solution demanded many months of planning, great expense to create an event that would draw managers together and the careful setting of a context that made managers relaxed enough to be open with their responses.

Part of the incentive to complete the GlobalCash questionnaire was an invitation to take part in a 'closed forum for respondents'. This meant that a date was set to present findings from the survey exclusively to respondents. The date and meeting place of The Management Centre in Brussels (a forum for 'serious' discussion of high quality) were established well in advance to allow managers the chance to put the date in their diaries. The day started with an initial session of presenting statistical findings to the whole group of managers who had responded to the questionnaire. As questions were fielded in this forum, particular topics of interest were identified and focus groups (never named as focus groups but as 'workshops') were built around these topics. By mid-morning, groups of around 10–12 managers were together tackling a theme of importance to them. With a loose set of topics to develop into questions and probes, the format for a focus group was achieved. ■

</td></tr>
</table>

The second major challenge of the people involved in a business purchase being professional is one of **interviewer credibility**. The subject of an interview may be technical in nature given the complexity of the product or service. This was discussed in Chapter 8 in the context of qualitative depth interviewing. The interviewer with credibility in the eyes of the respondent would appreciate the significance of particular revelations, understand the language (even technical language and jargon in certain areas) and ultimately get the respondent to open up, i.e. they will gain a greater level of access.

The nature of networks or relationships within an industry may have a level of complexity that could prove to be difficult for an interviewer to comprehend. These complexities may not present problems when a structured interview is being conducted. However, if there are open-ended questions or areas where probing is required, as would be necessary in any qualitative interview, the 'technical competence' of the interviewer will be revealed. If the interviewer has technical knowledge, probing can be much more meaningful to the respondent, the process of building a rapport can be much stronger and a richer and more revealing response may be obtained. In the GlobalCash study, the very precise technical nature of the subject and

the seniority of the respondents meant that interviewer credibility had to be tackled when the series of in-depth interviews was planned. The supporting researchers from business schools in individual European countries helped in gaining access to key respondents, and a Professor would be at hand to help with any language and cultural issues. The nature of the questions and probes were handled by a consultant with an international reputation in the field of cash management, that the respondents respected, wanted to talk to and could relate to at their technical level. The consultant had knowledge of the banking techniques, but, just as important, the nature of networks and relationships that operated in a highly competitive industry.

A b2b purchase may take a long time from the moment when the issue is first raised until final delivery

Given the complexity of a business customer's requirements, the importance or value of their purchase, the number of individuals involved in the purchase and the level of their knowledge and the help and advice that negotiating suppliers can give, a business purchase can be a very lengthy process. The following GlobalCash example illustrates the time taken for what may seem to be a simple decision, i.e. to choose to work with a new bank.

<table>
<tr><td>e x a m p l e
GlobalCash Project</td><td></td></tr>
</table>

Interfering parents

Hitachi Europe operate from offices in UK and Germany. They manage transactions with wholesalers and retailers throughout Europe involving the euro and other currencies, e.g. Swiss francs and UK pounds. They manage these transactions and the relationships with the parent company in Japan through a number of domestic and international banks.

One of the key challenges that Hitachi Europe faced was trying to simplify the process of managing transactions and cutting down on the expense incurred in their thousands of transactions. The first route that they took was to purchase and implement a new accounting system, which in itself took three years from the moment of recognising the nature of their problem through to evaluating possible solutions, negotiations with alternative suppliers and on to buying and implementing the system through to changes in practices and training their staff. The second route was to concentrate their banking business by cutting down the number of banks they had accounts with and to develop a new business relationship with a bank that had the international expertise, compatible information systems and service support of the quality level they sought.

For an organisation like Hitachi, choosing a new bank and setting up a new account is not like a consumer going into a bank branch or to a Website and selecting a product from a range of standard products. The main obstacle to change faced by Hitachi was the parent company's insistence that the European offices worked with certain Japanese banks. It is left to your imagination to appreciate the nature and intricacy of the networks and relationships between individuals working in Hitachi headquarters, regional offices, country offices, wholesalers, retailers and the banks that support their operations. Developing a new relationship between Hitachi Europe and a new bank, which would involve the demise of many establishing relationships, would involve much negotiation within Hitatchi and between existing banks and the new bank. For Hitatchi Europe the process started at the time of choosing a new accounting system. Working with the new bank took another two years to complete, five years in all. ■

In this example, the compounding factors of the time taken from the moment when an issue is first raised to final delivery emerge from the numbers of individuals who may be involved in the negotiations and the fact that they are professionals working in an environment of clear cost/benefit analyses. Assuming then that marketing researchers can identify who to talk to and gain access to them, what new challenges emerge given lengthy decision-making processes?

Marketing research challenges. The main challenge of the time taken to make a business purchase is one of **respondent error**.

Telescoping
A psychological phenomenon that takes place when an individual telescopes or compresses time by remembering an event as occurring more recently than it actually occurred.

In Chapter 13 we discussed the errors caused when administering questionnaires through the inability of respondents to remember events. **Telescoping** takes place when an individual telescopes or compresses time by remembering an event as occurring more recently than it actually occurred.[23] The ability to remember an event is influenced by (1) the event itself, (2) the time elapsed since the event, and (3) the presence or absence of things that would aid memory. We tend to remember events that are important or unusual or that occur frequently. Decision-makers that are part of a 'buying centre' are more likely to remember the negotiations, events and individuals associated with their most recent decisions. In the GlobalCash example the time taken to conduct negotiations developed over five years. During that time, negotiations may unfold in different geographical locations, involving different business functions and types of decision-maker. The personnel involved may even change as decision-makers move on to other projects or other organisations. For the marketing researcher this presents challenges in terms of accessing potential respondents who were present through the process of negotiations. If such decision-makers can be accessed, the next challenge is to get them to remember key events and individuals. Questions that do not provide the respondent with cues to the event, and that rely on unaided recall, can underestimate the actual occurrence and nature of an event.

In Chapter 3 (pages 75 to 77) we discussed the non-sampling errors termed as response errors. It was noted that a non-sampling error is likely to be more problematic than a sampling error. Its problematic nature emerges from the difficulties involved in identifying the error, measuring it and making changes in estimates. Response errors can be broken down into 'respondent inability' and 'respondent unwillingness'.

■ *Inability error* results from the respondent's inability to provide accurate answers. Respondents may provide inaccurate answers because language and logic used in a survey or questionnaire bear no relevance to their experiences, or the question format and content may be taking them into issues that are meaningless in their work context. Such an experience may result in respondents suffering from fatigue, boredom and a general lack of engagement with the process of questioning. For example, in the GlobalCash example, imagine a respondent being asked to respond to a series of Likert scale items about the service quality of online technical support. Though the organisation may use such support, a professional buyer responding to the scales may see the overall issue of online technical support as being of very low priority in the context of all the issues involved in negotiating with a new bank. Could the marketing researcher trust that the respondent has thought through the measured issues carefully and is being honest?

■ *Unwillingness error* arises from the respondent's unwillingness to provide accurate information. Respondents may intentionally misreport their answers because of a desire to provide socially acceptable answers, to avoid embarrassment or to please the interviewer. More especially in the b2b scenario, their answer has great commercial sensitivity and responding to it truthfully would reveal their commercial intellectual property. For example, to impress the interviewer a respondent may intentionally say that their business works with *The Cooperative Bank* in order to portray themselves as having relationships with organisations that profess a socially responsible agenda. They may not have any business with this bank, but would the interviewer be able to validate this response? As another example, the interviewer may pose questions about weaknesses in the services offered by a par-

ticular bank. The respondent may consider this sensitive as the weaknesses may reveal flaws and, by implication, the individuals involved in these flaws in their own organisation. The respondent may also question what may happen to their responses. If given directly to other banks in an unaggregated form, it would present all the elements required for another bank to make a sales pitch that may upset an established relationship.

Overcoming telescoping and respondent errors raises fundamental questions about the techniques used to elicit information and the benefits and rewards given to respondents. These types of error cannot be simply managed by larger samples as we will explore in the next section.

In many cases each business customer is individually important to a supplier and responsible for a significant proportion of its total sales

An example to illustrate this difference between b2b and consumer marketing is the extreme situation of dependence which can occur in commercial aviation, where there are only two main suppliers of airliners and only a few dozen major customers. Of course, not all b2b marketers have so few target customers. A manufacturer selling photocopier paper could have potentially huge numbers of businesses to sell their wares to, from major multinationals through to home businesses. In consumer marketing, a manufacturer of bespoke tailor-made shoes, selling at €2,000 a pair, may have relatively small numbers of target consumers that are responsible for a significant proportion of their total sales.

In b2b and consumer marketing, examples can be found of businesses that have a few high-value customers and those that have huge numbers of customers with low-value transactions. In general, however, b2b marketers have a higher propensity to deal with fewer customers of high value when compared with consumer marketers. The comparison between the two is illustrated in Figure 27.3.

To further illustrate the comparison between b2b and consumer marketing, an example from the GlobalCash study is presented below.
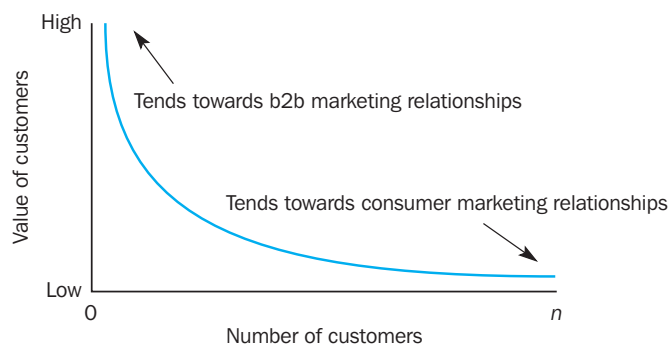
Figure 27.3
Customer number/value trade-off

Size matters

One would imagine that in commercial banking the number of target businesses is huge: every business requires banking services of some sort. 'Some sort' of services covers a vast array of services that banks can deliver. When these services are individually analysed, the individual number of consumers is not so big. We will examine such an analysis in the case of cash management banking.

Imagine yourself going on holiday and wanting to exchange your local currency for Australian dollars. All banks can help you make currency exchanges. The transaction is paid for through an exchange rate that is constantly changing and a commission fee. If you were a small business conducting transactions that involved a few currency exchanges per year, then you may have to pay the 'going rates' at your local bank or shop around for the best rates. However, for a company like Peugeot, imagine the variation in their profits depending upon the relative value of the euro to the US dollar and how much it costs for each currency transfer. Think of the currencies that Peugeot may have to deal in and the number of transactions and flows of funds. To Peugeot, the fluctuations in currency values and timing of cash transfers are crucial factors that can make the difference between sound profits and a loss. Given the volume of their business and the size of their transactions, Peugeot would not pay the 'going rate' that yourself or a small business would. They would expect a cash management banking service tailored specifically to their operations that may be radically different from the operations of BMW.

The target markets in cash management banking sought by major banks are those companies that operate internationally or even globally, dealing in huge sums and a variety of currencies. There may be millions of individual transactions to manage, huge networks of transactions across many countries or very high values to the transactions. When examining the nature of businesses that fulfill these requirements, there are relatively few compared with the total number of businesses, which is why the GlobalCash study, conducted on behalf of 15 of the largest pan-European banks, targeted just the largest 5,000 companies in Europe, a very small fraction of the total number of European businesses. ∎

The GlobalCash example illustrates how the essence of relatively few consumers may be of high value – through careful target marketing. Some businesses may have a mass marketing strategy that aims to target a great breadth of business types; others may have to specifically tailor their offerings to suit the specific needs of quite distinctive businesses. From the following example you are encouraged to track down the Kelly's directory (see the Web link references below) for a UK and global illustration of how businesses may be classified and the extent to which b2b marketers may finely tune the nature of their offerings to generate business with a few high-value businesses.

Kelly's – a 205-year-old success story[24]

Kelly's, a subsidiary of the publishers Reed Elselvier, is the UK's premier brand for industrial purchasing and supplies. Its print directory, popularly known as 'the Big Red Book' for industrial product and part manufacturers, comes out yearly, and has done since 1799. Its current directory aggregates more than 150,000 companies in 110,000 categories. Kelly's added a CD-ROM version of the directory in 1989 so that users could more easily search the listings. In 2001 a dynamic search-based site was established enabling customers to look for anything from abrasives to water pumps. Since then, the number of searches on **www.kellysearch.com** has risen from 10,000 to 12.5 million a month. The number of companies it represents has also ballooned from 150,000 to 1.3 million and each business receives quarterly figures tracking the responses the site has generated. ∎

Marketing research challenges. The main challenge of each business customer being individually important to a supplier and responsible for a significant proportion of its total sales is one of coping with relatively few target consumers. As most quantitative marketing research techniques tend to assume large numbers of potential target consumers and survey respondents, the challenge faced by b2b marketing researchers is one of **sampling methods** and the subsequent **analyses that may be performed given sample selection**.

We have already presented the challenges of 'element' definition in a sampling plan, gaining access to respondents (and the effect upon response rates) and response error. If we add on the challenge of conducting surveys when there are relatively few target respondents, it is clear to see that the whole sampling process becomes very challenging.

In choosing a sampling method for a small target population, probability methods such as random and systematic sampling can be utilised. With simple random sampling it is often difficult to construct a sampling frame that will permit a simple random sample to be drawn. Simple random sampling may or may not result in a representative sample; this is the paradox faced when drawing a random sample – one does not know if it will be representative of a target population, only the probability that it will be within margins of error. Although samples drawn will represent the population well on average, a given simple random sample may grossly misrepresent the target population. This is more likely if the size of the sample is small, as faced in many b2b surveys.

With systematic sampling, when the ordering of the sample elements is related to the characteristic of interest the technique increases the representativeness of the sample. If businesses within an industry are arranged in increasing order of annual sales, a systematic sample will include some small and some large firms. A simple random sample may be unrepresentative because it may contain, for example, only small firms or a disproportionate number of small firms. Choosing the right characteristics by which to order sample elements is crucial. With the wrong criteria, systematic sampling may decrease the representativeness of the sample. Thus, the power of systematic sampling lies in the accuracy of the sampling frame and the criteria used to order or arrange the sample elements.

Non-probability sampling
Sampling techniques that do not use chance selection procedures but rather rely on the personal judgement of the researcher.

With the time and cost factors of sourcing and building up-to-date and accurate sampling frames, with details of sample elements, many b2b marketing researchers will favour administrative, i.e. timing and cost, criteria in preference to statistical criteria, and ultimately be forced to use non-probability sampling techniques. **Non-probability sampling** relies on the personal judgement of the researcher rather than on chance to select sample elements. The b2b marketing researcher would aim to build the elements to include in a sample, rather than starting from a point of having an up-to-date and accurate sampling frame that resembles their target population. Building up the sample in this manner leads the b2b researcher to commonly use judgemental sampling and snowball sampling techniques (as discussed in Chapter 14, pages 364 to 367). Samples created using these techniques may yield good estimates of the population characteristics, but they do not allow for objective evaluation of the precision of the sample results. Because there is no way of determining the probability of selecting any particular element for inclusion in the sample, the estimates obtained are not statistically projectable to the population.

The following example illustrates the challenges faced in the IT industry in sourcing accurate sampling frames and how researchers spend time building their own.

Traditional market share research is dead in the water[25]

Data concerning IT market share has been one of the most visible aspects of b2b research, with IT trade magazines headlining leading IT companies' latest research findings. Traditionally, these data have been collected from large panels of retailers and resellers and often audited against shipment data provided by many of the manufacturers. Major changes in the IT market now mean that such panels provide only a partial view. Many key IT vendors are working to reposition themselves with their customers, targeting after-sales revenue which may spread over many years. For example, IBM is evolving from a manufacturer of hardware and associated software to a provider of IT services. As a result, for many key IT products, market size bears only an indirect relationship to the most recent purchase data. Measuring it therefore requires complex calculations based on combining the installed base of the product with the after-sales revenue flow it generates. In most cases, the raw data can only be collected by researching customers directly.

Contacting customers directly means that researchers have to collect data about a product which could be purchased by multiple decision-makers at dozens of different locations – and at site, department, or an individual level. These factors present quite formidable sampling challenges. Incompatibility between business sampling frames and official statistics regarding the number of businesses, where the latter exist at all, is common. Many research agencies are forced to build their own sampling frames, often from telephone directories. The result can be sampling frames that miss whole sections of the business population, such as sub-branches or government offices and perhaps the most 'invisible' group, the home-office sector. ■

Summary of sampling challenges. Whilst the theory of sampling is the same for all types of markets whatever their nature, there are key differences between sampling in b2b and consumer markets. These may be summarised as:[26]

1 *Unit definition*. Is it the site (establishment), organisation, legal entity or some other business unit? Each definition offers different benefits and drawbacks, depending upon the research objectives. In consumer research only the individual or household has to be considered.

There is much more variance between a small business and a large organisation than between any two consumers.

2 *Element definition.* In b2b research, an individual is interviewed as a representative of the business. We have to be sure to interview the right individual in each business, and be consistent across businesses. This can be complex as business titles and responsibilities vary widely across businesses of different sizes and sectors.

3 *Heterogeneity.* There is a lack of homogeneity in the business universe. There is much more variance between a small business and a large organisation than between any two consumers. There is also great variation between different types of small business.

4 *Sampling frames.* The available sampling frames of businesses and the nature of their detail that could allow stratification vary considerably in accuracy and coverage.

5 *Public sector.* The public sector generally counts as business, though the usage and buying patterns are frequently very different from the private sector.

A b2b purchase or sale is not an isolated event

Salespersons, purchasing professionals and indeed anyone making decisions in a 'buying centre' do not just meet, negotiate a deal and move on to other things. Imagine a cement manufacturer wishing to buy a furnace that is expected to be part of their cement production process for over 20 years. There would be many meetings with staff from the cement company, the various competing furnace manufacturers and many potential suppliers of key components for the furnace. Even when the decision to buy from one particular supplier has been made, the interactions between technical, financial, production and legal staff could take place over many months or even years. The purchase could be the forerunner to servicing and repair contracts and the installation of similar equipment in other cement-making plants.

In the above example of buying a new furnace, there could be a time when modifications need to be made. There are locations where local authorities allow cement manufacturers to burn old car tyres in their furnaces. Adding old tyres to a gas-fired furnace may require modifications to allow the tyres to be fed in, changes to the chimneys to get rid of different fumes and the lining of the furnace. All of these developments may again unfold over time involving purchasing, technicians, production, logistics, marketing and legal personnel. There may be components that need to be replaced frequently given the immense temperatures in the furnace. The delivery of these components may involve a repeat order which may need a simple telephone call or even be generated by computer.

In all of these cases, the customer and supplier interact on the basis of their experiences of previous purchases between them, and given the nature of their expectations for future purchases. The interaction between a customer and supplier is a single episode in the relationship between them. Each episode is affected by the relationship of which it forms part and each interaction in turn affects the relationship itself.[27] Each interaction and each purchase or sale can only be understood in the context of that relationship, and ultimately within the network or businesses that enable the relationships to work.

Marketing research challenges. The main challenge of a b2b purchase or sale not being an isolated event is finding a research design that is appropriate to understanding complex networks and relationships. The cross-sectional study is the most frequently used descriptive design in marketing research. **Cross-sectional designs** involve the collection of information from any given sample of population elements only once. They may be either single cross-sectional or multiple cross-sectional designs. In **single cross-sectional designs**, only one sample of respondents is drawn from the target population, and information is obtained from this sample only once.

Cross-sectional design
A type of research design involving the collection of information from any given sample of population elements only once.

Single cross-sectional design
A cross-sectional design in which one sample of respondents is drawn from the target population and information is obtained from this sample once.

These designs are also called sample survey research designs. In **multiple cross-sectional designs**, there are two or more samples of respondents, and information from each sample is obtained only once. In either variation of the design, much of the richness of the context of individual relationships and the network of relationships is lost. A more appropriate though more expensive and resource-intensive is the longitudinal design. In **longitudinal designs**, a fixed sample (or samples) of population elements is measured repeatedly. A longitudinal design differs from a cross-sectional design in that the sample or samples remains the same over time. In other words, the same people are studied over time. In contrast to the typical cross-sectional design, which gives a snapshot of the variables of interest at a single point in time, a longitudinal study provides a series of 'pictures'. These 'pictures' give an in-depth view of the situation and the changes that take place over time.

A key feature of creating such 'pictures' of relationships and networks is the building of contextual material from sources such as sales data, customer data, secondary data statistics, intelligence through press and trade commentary that focuses upon particular markets or industries and a vast resource of intelligence held within the employees of a business that may be seeking research support. In the following example, the SevenHR company was built around secondary data and intelligence sources, plus in-depth interviews within SevenHR.

e x a m p l e

A new approach to FMCG and retail recruitment – SevenHR[28]

Frustrated by current practices in the recruitment industry, a group of FMCG and retail professionals decided to form a new recruitment consultancy. SevenHR considered that a researched understanding of its markets and target customers should be at the core of its business. They commissioned the research agency Propaganda to develop their launch strategy. Propaganda was briefed to evaluate the perceptions of key industry decision-makers and to explore the opportunities of developing a differentiated recruitment brand. Propaganda used a two-stage strategy. First they conducted desk research using secondary data and intelligence to audit the competitors of SevenHR, with focus on how they portrayed themselves and what reputation they had earned among industry influencers. Second, internal research was conducted which included in-depth interviews with each of the founding members of the company. The findings of the research project enabled Propaganda to establish a cultural mission statement for SevenHR, and formed the basis of their communications tactics. Having launched in October 2002, within six months the business was already 75% ahead of forecast. Significantly, almost every client had placed repeat business, some returning up to five times. ■

There may be a lack of resources devoted to allowing b2b marketing researchers to use longitudinal studies and uncover the wider contextual issues that would allow detailed knowledge of relationships and networks. There could be a limitation in the way that b2b researchers view the value of contextual material. Bairfelt and Wilkinson[29] argue that many b2b marketing researchers demonstrate a 'tunnel vision', which fails to recognise the wider business context of business findings. Many researchers work comfortably within the standard research process of: a brief, proposal, research set-up, fieldwork, presentation and report. They continue to argue that 'desk research is a dying art and that the multitude of excellent Internet resources are sadly neglected'; many b2b researchers are consistently failing to put their research into a wider market or business context. Such a failing can affect the briefing/proposal stage as well as the analysis/reporting stage. This means that research findings are often set in a 'vacuum' without reference to the challenges faced by b2b marketers. The danger of this is that b2b marketing research is seen by decision-makers as a 'nice-to-know' or luxury item rather than a key part of the planning and decision-making process. Adding value by drawing on the various sources of

market intelligence available reduces the risk of marketing research appearing irrelevant to decision-makers. It can also improve the research process so that decision makers value the information provided.

The weaknesses of cross-sectional research designs, the resources required to perform longitudinal designs and the limitations of some b2b researchers to appreciate the wider context of their investigations may result in b2b decision-makers looking to other sources for decision support. Over recent years there has been a growth in the use and practice of competitor intelligence. Many organisations have developed competitor intelligence systems that allow them to have a relatively low-cost 'picture' of their environment, competitors and sometimes even their customers. Competitive intelligence uses many marketing research techniques and by using multiple sources aims to set a broad context to allow the b2b decision-maker to understand the networks in which they operate. The next section briefly explores the nature of competitor intelligence and its 'threat' to the nature and standards of conventional marketing research.

The growth of competitive intelligence

The research designs and techniques employed in competitive intelligence may in some respects differ from the structured methods of marketing research; in other respects there are many similarities. Over the last 15 years or so, those who professionally gather competitive intelligence have continued to refine research techniques and to distinguish their approach from industrial espionage. In general, the research design employed for competitive intelligence is exploratory, using primarily secondary data, intelligence sources and qualitative primary data.[30] Competitive intelligence has been defined as a systematic process by which a firm gathers, analyses and disseminates information about its rivals[31] or as a process of researching a competitor's organisation, products, prices, financial performance, technology and strategy.[32]

From these definitions we can see in 'systematic' and 'process' the continuous development of an understanding or knowledge of competitors. In the first definition we see the elements of managing data that replicate the marketing research process in terms of gathering, analysing and disseminating data. In the second definition we see the focus of data gathering in terms of products, prices, financial performance, technology and strategy. These are areas of measurement and understanding that go beyond the remit of conventional marketing researchers.

While the nature of competitive intelligence has evolved there has been much confusion over what competitive intelligence is and what it is not. Competitive intelligence is not corporate espionage, stealing competitive documents or accessing competitors' computer files. Rather, the discipline is a structured approach to a different type of research which has evolved as an outgrowth of strategic planning and market research. The Society of Competitive Professionals (SCIP) was formed in 1986 to develop behavioural standards and a code of ethics.[33]

The most commonly used techniques in competitive intelligence are:

1 *Review of public records:* monitoring secondary sources of information such as government records, commercial databases, media reports and news clippings, and company-produced literature, for potential threats and opportunities.
2 *Observational techniques:* directly observing competitors' activities or facilities through overt surveillance is another way to acquire competitive data, though such practices may be viewed with suspicion. Likewise, learning about a competitor's capabilities, strengths and weaknesses through physical inspection of its products

or actual use of its services is seen to be within the bounds of acceptable practice. Observers contribute to the gathering of 'informal' information concerning the market and business at different levels. Observers are not formal researchers but members of business departments, marketing specialists, technicians in engineering departments or plant-building departments. Employees in every department, including sales, purchasing and research and development, hear industry news well before it appears in a trade journal or in databases.[34] These observers and their contacts form key components of b2b relationships set in the context of networks.

3 *Personal interviews:* conversations with informed sources are beneficial when hard-to-find information about a competitor's strategies and intentions are used to fill critical gaps in secondary data and provide a more complete understanding of the competition.[35]

In the gathering and interpretation of 'public records', conducting personal interviews and overt observations, competitive intelligence targets individuals from the following types of organisation:[36]

- competitor companies;
- customers;
- suppliers;
- distributors;
- government organisations;
- trade associations;
- financial institutions.

Given the array of data sources, data-gathering techniques, institutions and individuals that may be targeted (including consumers in some instances), and the array of employees that may be used to 'observe', competitor intelligence is clearly a 'broad brush' approach. Competitor intelligence can offer the broad context and in many instances detail specific areas to b2b decision-makers that may be missing in conventional marketing research. Marketing research techniques may be more error free, less subjective and carried out with much stronger and more well-established codes of conduct when compared with competitor intelligence. However, for the b2b decision-maker needing to understand the long-term development of relationships and the dynamics of the network in which those relationships work, conventional views of marketing research may not offer the support they seek. They may forego an amount of error, interpret subjective and biased views and even push the boundaries of ethical practice to gain support that is more attuned to the decisions they face. In the following example we see how Texas Instruments, as long ago as 1995, moved this way by utilising their marketing research practices and findings as a foundation for their competitor intelligence system.

example

The evolution of marketing research archives to a comprehensive competitive intelligence system[37]

When Alicia Helton was Director of Corporate Market Research at Texas Instruments, she championed the placing of semiconductor marketing research archives on the corporate intranet. They began with current research reports and gradually added two years of earlier reports. At the time they used Microsoft PowerPoint for all research presentations, so it was easy to convert to HTML. They required all research suppliers to follow the same standards with a policy that the entire research presentation was available on the intranet within one week of the original presentation. The whole site used a portion of the semiconductor marketing site and, as this was already established, it proved to be the ideal foundation for their competitive intelligence system. From this point, Texas Instruments began to improve its acquisition, analysis and dissemination of competitive intelligence. The strategic leadership team worked with

consultants The Futures Group to fully integrate competitive intelligence into strategic decision making. Marketing research was seen as but one component of competitive intelligence. However, because the competitor intelligence system was built around marketing research archives, they found that the use and influence of marketing research grew. ■

This example illustrates that competitor intelligence need not replace marketing research. In this case, the well thought out use of competitor intelligence increased the use and influence of marketing research. This relationship happened because competitor intelligence helped to provide the broader context needed to get the most out of interpreting marketing research findings. There still remain ethical issues to consider with competitor intelligence, which will be tackled later in the chapter.

International marketing research

It is clear from the many examples throughout the text that b2b marketing research is practised in international settings. The challenges involved in conducting international research detailed in Chapter 26, i.e. adapting research techniques based upon different cultures and languages, simply add to the challenges of b2b marketing research detailed in this chapter. The b2b marketing researcher Stephen Connell of the Research Agency MORPACE International argues that in his experience international b2b research is getting easier for a number of reasons.[38] He contends that globalisation in b2b as well as consumer industries is making marketing research interviews more familiar and acceptable. The Web allows survey and focus-group respondents to be accessed in almost any country of the world. Technology allows for more centralisation of data collection and field operations alongside the more traditional options of using local sub-contractors or partners.

The following two examples summarise recent projects completed by Connell.[39] In both examples he highlights how small the potential sample may be and how that can help in managing the project from a central location.

example | Customer satisfaction survey in global telecoms

A survey was undertaken of 95 telephone companies from around the world, from Argentina to Inner Mongolia, in which up to six face-to-face interviews were carried out per company. The interviews, a total of 420 individuals in 200 organisations, combined scoring and rating questions with a considerable amount of open-ended discussion of business and technical issues. Appointments were set up by telephone and email, and interviewers travelled from the UK or, in one or two cases, from the USA. A total of nine languages, including Mandarin, Russian and Polish, were used to make appointments and to conduct the interviews.

The process worked well with response rates of over 85%; the no-show rate when the interviewer arrived was below 5%. The small team of interviewers achieved a very high standard of probing and issues analysis. The travel budget was only 15% of the total cost. ■

example | Revenue loss in telecoms

A survey was carried out of how telecom operators saw the problem of 'revenue loss' caused by billing problems, resulting in a failure to charge for calls, or by deliberate fraud. This was of major concern in the industry and the survey was designed to assess: the extent to which the problem was recognised; quality perceptions of the amount of revenue lost in this way, and approaches to detecting and minimising loss. A total of 50 telephone in-depth interviews were carried out, in more than 30 countries, using a list of fixed and mobile operators assembled by a leading consultancy. ■

In this last example, note that a specialist consultancy was used to develop a sampling frame. A further difficulty faced even when the name and telephone number were listed in the sampling frame was identifying the sample elements. Identifying the relevant decision-maker in each business was difficult as the persons responsible for limiting revenue loss had a wide variety of titles, e.g. Manager Network Operation, Vice President – Finance, Controller, Risk Management Manager, etc. Once these were tracked down, the next challenge was persuading them that the survey was bona-fide and worth responding to. In the end the survey was a success given that a wide range of companies participated despite the clearly sensitive nature of the subject matter.

Connell also presents some generalised observations of cultural differences in applying b2b marketing research which are worth debating:

■ Telephone and Web interviews with b2b respondents can be very difficult to achieve in China. Managers are often unwilling to discuss important issues without the reassurance of face-to-face contact. On the other hand, personal interviews, once started, can yield as much information and ideas as anywhere in the world, though it can be difficult to tie down respondents to precise appointments in advance.

■ b2b respondents in other countries, including some Southern European countries where marketing research is less familiar, require more reassurance about confidentiality and security. Faxes explaining the research agency, the sponsoring client and the project credentials are more likely to be required after initial contact by telephone (for a telephone or face-to-face survey). Once the interview is granted, excellent information can be provided.

■ Telephone and self-completion questionnaires sometimes elicit less detailed and informative responses from Nordic (Denmark, Finland, Iceland, Norway, Sweden) respondents than elsewhere. In this region b2b respondents were more cautious, using formal ways to respond to unstructured questions, although this behaviour can be modified in the context of a face-to-face interview.

This latter observation was certainly the same experience as in the GlobalCash studies. The Nordic/Scandinavian respondents were most cautious in responding to the more sensitive questions, with the Danes having the highest non-response levels to questions that referred to their experiences of specific banks.

With the spread of globalisation and increase of international trade, international marketing research has similarly grown. The marketing research industry has had to respond to the challenges set, with major agencies developing account or relationship teams that serve the worldwide operations of their customers.

Ethics in marketing research

One of the major differences in competitive intelligence and traditional market research is the knowledge of the purpose of the study. Ethical marketing research should reveal the purpose of the study to the respondent, and, sometimes, how the information will be used, especially when explaining why particularly sensitive questions are being asked. Competitive intelligence studies in contrast generally do not reveal the purpose of the study and how the information will be used. Essentially, the competitive intelligence interviewer must develop a 'story' as to why

they are calling the respondent, e.g. they are interested in their products, they are conducting a general industry study, they are conducting a study as part of their degree or they are looking at specific issues within the industry (e.g. legislation or product safety). This may generate a response and some respondents may be happy to cooperate. However, this would be clearly considered as unethical by traditional marketing research organisations.[40] Being duplicitous in gaining access to respondents could potentially damage the reputation of legitimate marketing researchers and ultimately damage response rates, making access continually more difficult.

For the bona-fide b2b marketing researcher, it is usual to disclose the identity of their clients, i.e. who commissioned the research, either up-front or in the course of an interview. This means that marketing research has to play an ambassadorial role. Marketing researchers are not speaking for their clients, but are representing them in asking respondents for their views. As a result the behaviour of researchers must be highly professional and care must be taken of client and respondent reputation. The size of respondent populations is often so small that treating respondents with care is becoming ever more important and sometimes b2b marketing researchers need to dissuade their clients from trying to cram more into ever-longer interviews. Although it is tempting to ask as much as possible in the one 'hit', this rebounds on researchers and their clients in the long run.[41] Looking back to Chapter 15, pages 393 and 394, where non-response issues were discussed, re-examine the proposed means to improve response rates. Now consider how feasible it is to increase response rates with a small pool of respondents who may be very irritated by the number of requests for interviews and the length of those interviews.

Internet and computer applications

For business respondents, the Web is a useful additional access and information exchange tool beyond the more traditional methods of gathering data. As a means to gain access to b2b respondents, it does not always offer the potential that many commentators believed it would. b2b respondents are busy and, although they may be delighted to talk if asked questions, they often do not have the time to complete Web-based questionnaires, however good their intentions. In the GlobalCash surveys, the Web-based survey, whilst much shorter than the postal survey, did not achieve anything like the same success in terms of response rate. This is only one example and does not dismiss the hugely important role of the Web in b2b interviewing. It can be used to send respondents stimulus or advertising material to look at, to confirm appointments, thank the respondent, and deliver incentives or summary reports; and to publish research results so that respondents, equipped with access passwords, can look at such data and even manipulate data as is relevant to them. A whole spectrum of management-tiered tailored results and interactivity has opened up through the Web.

The Web can also be used for interviewer briefing. One chemical company for instance commonly used to bring sheets of polythene and other materials used for insulation to briefing sessions. Now, however, it can point the interviewers to the Websites, which is a great advantage to all those managing the field force and the interviewers working in the field. Interviewers can develop a clear understanding of

esoteric topics such as insulation, aviation fuel, cabling or the solvents industry. The Web, together with the collection of accurate email addresses, is becoming increasingly important, but in limited, facilitating areas rather than as the overall vehicle for data collection.[42]

Summary

This chapter was founded upon the key question of whether there are substantial differences between b2b and consumer marketing and thus differences in the approaches and challenges faced by b2b and consumer marketing researchers. The two main bases of b2b marketing are the concepts of networks and relationships. In the context of understanding networks and relationships, the prime objective of marketing researchers in business markets therefore is to support an understanding of how the network operates and how successful relationships evolve. The nature of networks and relationships underpins five key differences between b2b marketing and consumer marketing. These can be summarised as differences in the: number of people who may be involved in a b2b purchasing decision; people involved in a b2b purchase being professionals; nature of b2b purchases taking a long time from the moment when an issue is first raised until final delivery; many cases where each business customer is individually important to a supplier and responsible for a significant proportion of its total sales; and b2b purchases or sales not being isolated events. These differences are the reason that there are distinctive differences in the approaches adopted and challenges faced by b2b and consumer marketing researchers. The key challenges they face are in:

- Administering effective sampling plans with the associated problems of finding appropriate sampling techniques and the limitations of inference from the statistics that they may generate.
- Gaining access to key respondents, in terms of getting to interview respondents in the first instance and getting them to be open and honest. A key element of this access lies in how credible interviewers are perceived by business respondents.
- Business respondents actually remembering what has happened and why they or their organisation behaved as they did. The events may have taken a long time to unfold, involved numerous individuals and taken a great deal of time.
- Creating research designs that can measure or help to create an understanding of the dynamics of networks and relationships.

b2b marketing researchers face further challenges from practitioners in competitor intelligence who have to work with decision-makers who may forego an amount of error, interpret subjective and biased views and even push the boundaries of ethical practice to gain support that is more attuned to the decisions they face. Competitor intelligence need not be a direct replacement for conventional marketing research; it can be a strong supplement that allows research findings to be understood in a more dynamic and broader context. Even though many of the techniques adopted in competitor research match those in marketing research and may be quite removed from notions of industrial espionage, there still remain clear ethical issues about such approaches.

b2b marketing researchers tackle the challenges summarised above, mostly in very creative manners, as illustrated in the examples throughout this chapter and indeed throughout the text. So well in fact that b2b researchers sometimes see themselves as

superior to consumer researchers![43] With the spread of globalisation and increase of international trade, international b2b marketing research has similarly grown. The marketing research industry has had to respond to the challenges set, with major agencies developing account or relationship teams that serve the worldwide operations of their customers. The use of the Web has enabled the management and control of international b2b marketing research to become far more cost effective. The Web has also enabled the briefing and training of field interviewers to become much more effective in terms of enabling them to understand the technical issues of the questions they are posing. This in turn can increase their credibility with business respondents and improve the quality of their response.

Questions

1 Why is it important to ask the question 'is b2b marketing research significantly different from consumer marketing research'?

2 What characteristics distinguish b2b marketing from consumer marketing?

3 From what focal point should a business network be analysed?

4 What challenges do b2b marketers face in managing relationships?

5 How do the concepts of business networks and relationships relate to each other?

6 What are the five major differences between b2b and consumer marketing?

7 Describe the sampling challenges that the b2b marketing researcher faces.

8 What does 'access' mean in the context of b2b marketing research?

9 Evaluate the reasons why gaining access to key respondents may be so difficult. Why may gaining access grow more difficult over time?

10 What is meant by 'interviewer credibility'? What may interviewers do to be seen as more credible in the eyes of target respondents?

11 Describe the potential respondent errors that may occur in b2b marketing research. What may be done to reduce these potential errors?

12 What is competitor intelligence? How does this definition differ from notions of 'conventional marketing research'?

13 What sampling frame challenges exist in international b2b marketing research?

14 To what extent do you think that competitor analysts have values to offer to potential respondents in any interviews they conduct?

15 Describe the means by which the Web may improve the process of conducting b2b marketing research.

Notes

1 Chilvers, V., 'Fair's fair at trade shows', *Research* (November 2003), 63.

2 Dibb, S., Simkin, L., Pride, W.M. and Ferrell, O.C., *Marketing: Concepts and Strategies*, 4th edn (Boston: Houghton-Mifflin, 2000), 158.

3 David Ford conducts his work with a team of researchers in the IMP (International Marketing and Purchasing) Group. This group was founded in 1976 by researchers from five European countries. The group's 'interaction approach' is based on the importance for both researchers and managers of understanding the *interaction* between *active* buyers and sellers in continuing business *relationships*. Ford, D., *The Business Marketing Course: Managing in Complex Networks* (Chichester: Wiley, 2002), 6. See www.impgroup.org

4 Littler, D., 'Organizational marketing' in Baker, M.J. (ed.), *The Marketing Book*, 3rd edn (Oxford: Butterworth-Heinneman, 1994), 610.

5 Wilson, D.F., 'Why divide consumer and organizational buyer behaviour?', *European Journal of Marketing*, 34(7) (2000), 780–96.

6 Wilson, D.F., 'Why divide consumer and organizational buyer behaviour?', *European Journal of Marketing*, 34(7) (2000), 789.

7 Webster, F.E., Jr, 'Management science in industrial marketing', *Journal of Marketing*, 42 (January 1978), 22.

8 Ford, D., *The Business Marketing Course: Managing in Complex Networks* (Chichester: Wiley, 2002), 29.

9 Hakansson, H. and Snehota, I., 'No business is an island: the network concept of business strategy', *Scandinavian Journal of Management*, 14(3) (1990), 177–200.

10 Ford, D., Gadde, L.E., Hakansson, H. and Snehota, I., *Managing Business Relationships*, 2nd edn (Chichester: Wiley, 2003), 38.

11 Ford, D., *The Business Marketing Course: Managing in Complex Networks* (Chichester: Wiley, 2002), 3–6.

12 Spekman, R.E. and Stern, L.W., 'Environmental uncertainty and buying group structure: an empirical investigation', *Journal of Marketing*, 43 (1979), 54–64.

13 Webster, F.E. Jr and Wind, Y., 'A general model of organizational buying behaviour', *Journal of Marketing*, 36(2) (1972), 12–19.

14 Bonoma, T.V., 'Major sales: who really does the buying?', *Harvard Business Review*, 60(3) (1982), 111–19.

15 Wilson, D.F., *Organizational Marketing* (London: Routledge, 1998).

16 *Ibid.*

17 Ford, D., *The Business Marketing Course: Managing in Complex Networks* (Chichester: Wiley, 2002), 79.

18 McNeil, R., 'Open for business', *Research* (May 2003), 63.

19 Field, R., 'Open for business', *Research* (May 2003), 63.

20 Minett, S., B2B *Marketing* (Harlow: Financial Times Prentice Hall, 2002), 63.

21 McPhee, N., 'Gaining insight on business and organisational behaviour: the Qualitative Dimension', *International Journal of Marketing Research*, 44(1) (January 2002).

22 Hamilton-Stent, S., 'A spoonful of research helps the medicine go down', *Research* (November 2003), 63.

23 Menon, G., Raghubir, P and Schwarz, N., 'Behavioural frequency judgments: an accessibility–diagnosticity framework', *Journal of Consumer Research*, 22(2) (September 1995), 212–28; Cook, W.A., 'Telescoping and memory's other tricks', *Journal of Advertising Research* (February–March 1987), 5–8; and Sudman, S., Finn, A. and Lannom, L., 'The use of bounded recall procedures in single interviews', *Public Opinion Quarterly* (Summer 1984), 520–24.

24 Hemsley, S., 'Book, online and tinker', *Marketing Week* (17 June), 39-41.

25 Stubington, P., 'Traditional market share research is dead in the water', *Research* (November 2004), 63.

26 Macfarlane, P., 'Structuring and measuring the size of business markets', *International Journal of Market Research*, 44(1) (2002).

27 Ford, D., Gadde, L.E., Hakansson, H. and Snehota, I., *Managing Business Relationships*, 2nd edn (Chichester: Wiley, 2003), 4.

28 Anon, Special supplement on Marketing Research Awards – Business and Professional, *Marketing* (7 July 2004).

29 Bairfelt, S. and Wilkinson, T., 'Future proof', *Research* (May 2003), 63.

30 Stanat, R., 'The relationship between market research and competitive intelligence', *ESOMAR Journal* (1998).

31 Kahaner, L., *Competitive Intelligence: From Black Ops to Boardrooms. How Businesses Gather, Analyze, and Use Information to Succeed in a Global Marketplace* (New York, NY: Simon and Schuster, Inc., 1996).

32 Stanat, R., 'The relationship between market research and competitive intelligence', *Proceedings of the ESOMAR Congress Berlin* (1998).

33 van Hamersveld, M., 'What are we talking about?', *ESOMAR Journal* (1999), www.warc.com/fulltext/ESOMAR/13800.htm.

34 Jakobiak, F., *ESOMAR Journal* (1999), www.warc.com/fulltext/ESOMAR/13801.htm.

35 Kassler, H.S., 'Mining the Internet for competitive intelligence', *Online* (September/October 1997), 34–45, www.onlineinc.com/onlinemag.

36 Stanat, R., 'The relationship between market research and competitive intelligence', *ESOMAR Journal* (1998).

37 Helton, A.S., 'Data, data everywhere, but no competitive intelligence in sight', *ESOMAR Journal* (1999) www.warc.com/print/13808p.asp.

38 Connell, S., 'Travel broadens the mind – the case for international research', *International Journal of Market Research*, 44(1) (2002), 7.

39 Connell, S., 'Travel broadens the mind – the case for international research', *International Journal of Market Research*, 44(1) (2002), 3–4.

40 Stanat, R., 'The relationship between market research and competitive intelligence', *ESOMAR Journal* (1998).

41 McNeil, R., 'Open for business', *Research* (May 2003), 63.

42 *Ibid.*

43 Britton, T. 'Superiority "devalues" b2b research', *Research* (21 May 2003), www.research-live.com/index.aspx?pageid=30&newsid=259.

Table 1 Simple random numbers

Line/col.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1	10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291	90700
2	22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615	99505
3	24130	48390	22527	97265	76393	64809	15179	24830	49340	32081	30680	19655	63348	58629
4	42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758	16379
5	37570	39975	81837	16656	06121	91782	60468	81305	49684	60072	14110	06927	01263	54613
6	77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394	42880
7	99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
8	96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
9	89579	14342	63661	10281	17453	18103	57740	84378	25331	12568	58678	44947	05585	56941
10	85475	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593	64952
11	28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610	78188
12	63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703	90322
13	09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613	74952
14	10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104	20285	29975	89868
15	07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372	74461	28551	90707
16	51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
17	02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835	44919	05944	55157
18	01011	54092	33362	94904	31273	04146	18594	29852	71685	85030	51132	01915	92747	64951
19	52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738	17752	35156	35749
20	07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916	19509	25625	58104
21	48663	91245	85828	14346	09172	30163	90229	04734	59193	22178	30421	61666	99904	32812
22	54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524	15227	96909	44592
23	32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012	64161	18296	22851
24	29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10307	07684	36188	18510
25	02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953
26	81525	72295	04839	96423	24878	82651	66566	14778	76797	14780	13300	87074	79666	95725
27	29676	20591	68086	26432	46901	20849	89768	81536	86645	12659	92259	57102	80428	25280
28	00742	57392	39064	66432	84673	40027	32832	61362	98947	96067	64760	64584	96096	98253
29	05366	04213	25669	26422	44407	44048	37937	63904	45766	66134	75470	66520	34693	90449
30	91921	26418	64117	94305	26766	25940	39972	22209	71500	64568	91402	42416	07844	69618
31	00582	04711	87917	77341	42206	35126	74087	99547	81817	42607	43808	76655	62028	76630
32	00725	69884	62797	56170	86324	88072	76222	36086	84637	93161	76038	65855	77919	88006
33	69011	65795	95876	55293	18988	27354	26575	08625	40801	59920	29841	80150	12777	48501
34	25976	57948	29888	88604	67917	48708	18912	82271	65424	69774	33611	54262	85963	03547
35	09763	83473	73577	12908	30883	18317	28290	35797	05998	41688	34952	37888	38917	88050
36	91567	42595	27959	30134	04024	86385	29880	99730	55536	84855	29088	09250	79656	73211
37	17955	56349	90999	49127	20044	59931	06115	20542	18059	02008	73708	83517	36103	42791
38	46503	18584	18845	49618	02304	51038	20655	58727	28168	15475	56942	53389	20562	87338
39	92157	89634	94824	78171	84610	82834	09922	25417	44137	48413	25555	21246	35509	20468
40	14577	62765	35605	81263	39667	47358	56873	56307	61607	49518	89656	20103	77490	18062
41	98427	07523	33362	64270	01638	92477	66969	98420	04880	45585	46565	04102	46880	45709
42	34914	63976	88720	82765	34476	17032	87589	40836	32427	70002	70663	88863	77775	69348
43	70060	28277	39475	46473	23219	53416	94970	25832	69975	94884	19661	72828	00102	66794
44	53976	54914	06990	67245	68350	82948	11398	42878	80287	88267	47363	46634	06541	97809
45	76072	29515	40980	07391	58745	25774	22987	80059	39911	96189	41151	14222	60697	59583

(*continued*)

Table 1 (*continued*)

Line/col.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
46	90725	52210	83974	29992	65831	38857	50490	83765	55657	14361	31720	57375	56228	41546
47	64364	67412	33339	31926	14883	24413	59744	92351	97473	89286	35931	04110	23726	51900
48	08962	00358	31662	25388	61642	34072	81249	35648	56891	69352	48373	45578	78547	81788
49	95012	68379	93526	70765	10592	04542	76463	54328	02349	17247	28865	14777	62730	92277
50	15664	10493	20492	38301	91132	21999	59516	81652	27195	48223	46751	22923	32261	85653
51	16408	81899	04153	53381	79401	21438	83035	92350	36693	31238	59649	91754	72772	02338
52	18629	81953	05520	91962	04739	13092	97662	24822	94730	06496	35090	04822	86774	98289
53	73115	35101	47498	87637	99016	71060	88824	71013	18735	20286	23153	72924	35165	43040
54	57491	16703	23167	49323	45021	33132	12544	41035	80780	45393	44812	12515	98931	91202
55	30405	83946	23792	14422	15059	45799	22716	19792	09983	74353	68668	30429	70735	25499
56	16631	35006	85900	98275	32388	52390	16815	69293	82732	38480	73817	32523	41961	44437
57	96773	20206	42559	78985	05300	22164	24369	54224	35083	19687	11052	91491	60383	19746
58	38935	64202	14349	82674	66523	44133	00697	35552	35970	19124	63318	29686	03387	59846
59	31624	76384	17403	53363	44167	64486	64758	75366	76554	31601	12614	33072	60332	92325
60	78919	19474	23632	27889	47914	02584	37680	20801	72152	39339	34806	08930	85001	87820
61	03931	33309	57047	74211	63445	17361	62825	39908	05607	91284	68833	25570	38818	46920
62	74426	33278	43972	10119	89917	15665	52872	73823	73144	88662	88970	74492	51805	99378
63	09066	00903	20795	95452	92648	45454	69552	88815	16553	51125	79375	97596	16296	66092
64	42238	12426	87025	14267	20979	04508	64535	31355	86064	29472	47689	05974	52468	16834
65	16153	08002	26504	41744	81959	65642	74240	56302	00033	67107	77510	70625	28725	34191
66	21457	40742	29820	96783	29400	21840	15035	34537	33310	06116	95240	15957	16572	06004
67	21581	57802	02050	89728	17937	37621	47075	42080	97403	48626	68995	43805	33386	21597
68	55612	78095	83197	33732	05810	24813	86902	60397	16489	03264	88525	42786	05269	92532
69	44657	66999	99324	51281	84463	60563	79312	93454	68876	25471	93911	25650	12682	73572
70	91340	84979	46949	81973	37949	61023	43997	15263	80644	43942	89203	71795	99533	50501
71	91227	21199	31935	27022	84067	05462	35216	14486	29891	68607	41867	14951	91696	85065
72	50001	38140	66321	19924	72163	09538	12151	06878	91903	18749	34405	56087	82790	70925
73	65390	05224	72958	28609	81406	39147	25549	48542	42627	45233	57202	94617	23772	07896
74	27504	96131	83944	41575	10573	03619	64482	73923	36152	05184	94142	25299	94387	34925
75	37169	94851	39117	89632	00959	16487	65536	49071	39782	17095	02330	74301	00275	48280
76	11508	70225	51111	38351	19444	66499	71945	05422	13442	78675	84031	66938	93654	59894
77	37449	30362	06694	54690	04052	53115	62757	95348	78662	11163	81651	50245	34971	52974
78	46515	70331	85922	38329	57015	15765	97161	17869	45349	61796	66345	81073	49106	79860
79	30986	81223	42416	58353	21532	30502	32305	86482	05174	07901	54339	58861	74818	46942
80	63798	64995	46583	09785	44160	78128	83991	42865	92520	83531	80377	35909	81250	54238
81	82486	84846	99254	67632	43218	50076	21361	64816	51202	88124	41870	52689	51275	83556
82	21885	32906	92431	09060	64297	51674	64126	62570	26123	05155	59194	52799	28225	85762
83	60336	98782	07408	53458	13564	59089	26445	29789	85205	41001	12535	12133	14645	23541
84	43937	46891	24010	25560	86355	33941	25786	54990	71899	15475	95434	98227	21824	19535
85	97656	63175	89303	16275	07100	92063	21942	18611	47348	20203	18534	03862	78095	50136
86	03299	01221	05418	38982	55758	92237	26759	86367	21216	98442	08303	56613	91511	75928
87	79626	06486	03574	17668	07785	76020	79924	25651	83325	88428	85076	72811	22717	50585
88	85636	68335	47539	03129	65651	11977	02510	26113	99447	68645	34327	15152	55230	93448
89	18039	14367	61337	06177	12143	46609	32989	74014	64708	00533	35398	58408	13261	47908
90	08362	15656	60627	36478	65648	16764	53412	09013	07832	41574	17639	82163	60859	75567
91	79556	29068	04142	16268	15387	12856	66227	38358	22478	73373	88732	09443	82558	05250
92	92608	82674	27072	32534	17075	27698	98204	63863	11951	34648	88022	56148	34925	57031
93	23982	25835	40055	67006	12293	02753	14827	23235	35071	99704	37543	11601	35503	85171
94	09915	96306	05908	97901	28395	14186	00821	80703	70426	75647	76310	88717	37890	40129
95	59037	33300	26695	62247	69927	76123	50842	43834	86654	70959	79725	93872	28117	19233
96	42488	78077	69882	61657	34136	79180	97526	43092	04098	73571	80799	76536	71255	64239
97	46764	86273	63003	93017	31204	36692	40202	35275	57306	55543	53203	18098	47625	88684
98	03237	45430	55417	63282	90816	17349	88298	90183	36600	78406	06216	95787	42579	90730
99	86591	81482	52667	61582	14972	90053	89534	76036	49199	43716	97548	04379	46370	28672
100	38534	01715	94964	87288	65680	43772	39560	12918	80537	62738	19636	51132	25739	56947
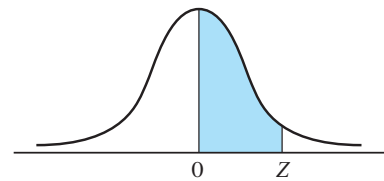
Table 2 Area under the normal curve

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.49865	.49869	.49874	.49878	.49882	.49886	.49889	.49893	.49897	.49900
3.1	.49903	.49906	.49910	.49913	.49916	.49918	.49921	.49924	.49926	.49929
3.2	.49931	.49934	.49936	.49938	.49940	.49942	.49944	.49946	.49948	.49950
3.3	.49952	.49953	.49955	.49957	.49958	.49960	.49961	.49962	.49964	.49965
3.4	.49966	.49968	.49969	.49970	.49971	.49972	.49973	.49974	.49975	.49976
3.5	.49977	.49978	.49978	.49979	.49980	.49981	.49981	.49982	.49983	.49983
3.6	.49984	.49985	.49985	.49986	.49986	.49987	.49987	.49988	.49988	.49989
3.7	.49989	.49990	.49990	.49990	.49991	.49991	.49992	.49992	.49992	.49992
3.8	.49993	.49993	.49993	.49994	.49994	.49994	.49994	.49995	.49995	.49995
3.9	.49995	.49995	.49996	.49996	.49996	.49996	.49996	.49996	.49997	.49997

Each entry represents the area under the standard normal distribution from the mean to Z.

$1-\alpha$

α

0 $\qquad\qquad$ $x^2_{(\alpha,df)}$ $\qquad$ x^2
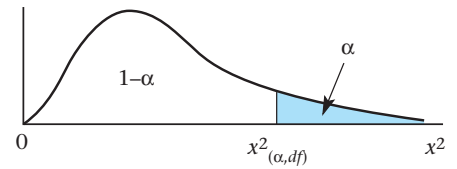
Table 3 Chi-square distribution

Degrees of freedom	Upper tail area (α)											
	.995	.99	.975	.95	.90	.75	.25	.10	.05	.025	.01	.005
1			0.001	0.004	0.016	0.102	1.323	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	0.575	2.773	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	1.213	4.108	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	1.923	5.385	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	2.675	6.626	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	3.455	7.841	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	4.255	9.037	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	5.071	10.219	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	5.899	11.389	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	6.737	12.549	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	7.584	13.701	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	8.438	14.845	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	9.299	15.984	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	10.165	17.117	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	11.037	18.245	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	11.912	19.369	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	12.792	20.489	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	13.675	21.605	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	14.562	22.718	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	15.452	23.828	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	16.344	24.935	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.042	17.240	26.039	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	18.137	27.141	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	19.037	28.241	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	19.939	29.339	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	20.843	30.435	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	21.749	31.528	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	22.657	32.620	37.916	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	23.567	33.711	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	24.478	34.800	40.256	43.773	46.979	50.892	53.672
31	14.458	15.655	17.539	19.281	21.434	25.390	35.887	41.422	44.985	48.232	52.191	55.003
32	15.134	16.362	18.291	20.072	22.271	26.304	36.973	42.585	46.194	49.480	53.486	56.328
33	15.815	17.074	19.047	20.867	23.110	27.219	38.058	43.745	47.400	50.725	54.776	57.648
34	16.501	17.789	19.806	21.664	23.952	28.136	39.141	44.903	48.602	51.966	56.061	58.964
35	17.192	18.509	20.569	22.465	24.797	29.054	40.223	46.059	49.802	53.203	57.342	60.275
36	17.887	19.233	21.336	23.269	25.643	29.973	41.304	47.212	50.998	54.437	58.619	61.581
37	18.586	19.960	22.106	24.075	26.492	30.893	42.383	48.363	52.192	55.668	59.892	62.883
38	19.289	20.691	22.878	24.884	27.343	31.815	43.462	49.513	53.384	56.896	61.162	64.181
39	19.996	21.426	23.654	25.695	28.196	32.737	44.539	50.660	54.572	58.120	62.428	65.476
40	20.707	22.164	24.433	26.509	29.051	33.660	45.616	51.805	55.758	59.342	63.691	66.766
41	21.421	22.906	25.215	27.326	29.907	34.585	46.692	52.949	56.942	60.561	64.950	68.053
42	22.138	23.650	25.999	28.144	30.765	35.510	47.766	54.090	58.124	61.777	66.206	69.336

(*continued*)

Table 3 (*continued*)

Degrees of freedom	Upper tail areas (α)											
	.995	.99	.975	.95	.90	.75	.25	.10	.05	.025	.01	.005
43	22.859	24.398	26.785	28.965	31.625	36.436	48.840	55.230	59.304	62.990	67.459	70.616
44	23.584	25.148	27.575	29.787	32.487	37.363	49.913	56.369	60.481	64.201	68.710	71.893
45	24.311	25.901	28.366	30.612	33.350	38.291	50.985	57.505	61.656	65.410	69.957	73.166
46	25.041	26.657	29.160	31.439	34.215	39.220	52.056	58.641	62.830	66.617	71.201	74.437
47	25.775	27.416	29.956	32.268	35.081	40.149	53.127	59.774	64.001	67.821	72.443	75.704
48	26.511	28.177	30.755	33.098	35.949	41.079	54.196	60.907	65.171	69.023	73.683	76.969
49	27.249	28.941	31.555	33.930	36.818	42.010	55.265	62.038	66.339	70.222	74.919	78.231
50	27.991	29.707	32.357	34.764	37.689	42.942	56.334	63.167	67.505	71.420	76.154	79.490
51	28.735	30.475	33.162	35.600	38.560	43.874	57.401	64.295	68.669	72.616	77.386	80.747
52	29.481	31.246	33.968	36.437	39.433	44.808	58.468	65.422	69.832	73.810	78.616	82.001
53	30.230	32.018	34.776	37.276	40.308	45.741	59.534	66.548	70.993	75.002	79.843	83.253
54	30.981	32.793	35.586	38.116	41.183	46.676	60.600	67.673	72.153	76.192	81.069	84.502
55	31.735	33.570	36.398	38.958	42.060	47.610	61.665	68.796	73.311	77.380	82.292	85.749
56	32.490	34.350	37.212	39.801	42.937	48.546	62.729	69.919	74.468	78.567	83.513	86.994
57	33.248	35.131	38.027	40.646	43.816	49.482	63.793	71.040	75.624	79.752	84.733	88.236
58	34.008	35.913	38.844	41.492	44.696	50.419	64.857	72.160	76.778	80.936	85.950	89.477
59	34.770	36.698	39.662	42.339	45.577	51.356	65.919	73.279	77.931	82.117	87.166	90.715
60	35.534	37.485	40.482	43.188	46.459	52.294	66.981	74.397	79.082	83.298	88.379	91.952

For a particular number of degrees of freedom, each entry represents the critical value of χ^2 corresponding to a specified upper tail area, α.

For larger values of degrees of freedom (df), the expression $z = \sqrt{2\chi^2} - \sqrt{2(df) - 1}$ may be used and the resulting upper tail area can be obtained from the table of the standardised normal distribution.
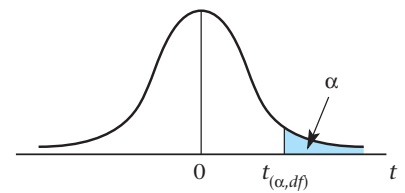
Table 4 *t* distribution

Degrees of freedom	Upper tail area					
	.25	.10	.05	.025	.01	.005
1	1.0000	3.0777	6.3138	12.7062	31.8207	63.6574
2	0.8165	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.7407	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0322
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2,1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7207	2.0796	2.5177	2.8314
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7969
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.6825	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.6822	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.6820	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.6818	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.6816	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.6814	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.6812	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.6810	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.6808	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.6807	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.6805	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.6804	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.6802	1.3016	1.681 1	2.0167	2.4163	2.6951

(*continued*)

717

Table 4 (*continued*)

Degrees of freedom	Upper tail area					
	.25	.10	.05	.025	.01	.005
44	0.6801	1.3011	1.6802	2.0154	2.4141	2.6923
45	0.6800	1.3006	1.6794	2.0141	2.4121	2.6896
46	0.6799	1.3002	1.6787	2.0129	2.4102	2.6870
47	0.6797	1.2998	1.6779	2.0117	2.4083	2.6846
48	0.6796	1.2994	1.6772	2.0106	2.4066	2.6822
49	0.6795	1.2991	1.6766	2.0096	2.4049	2.6800
50	0.6794	1.2987	1.6759	2.0086	2.4033	2.6778
51	0.6793	1.2984	1.6753	2.0076	2.4017	2.6757
52	0.6792	1.2980	1.6747	2.0066	2.4002	2.6737
53	0.6791	1.2977	1.6741	2.0057	2.3988	2.6718
54	0.6791	1.2974	1.6736	2.0049	2.3974	2.6700
55	0.6790	1.2971	1.6730	2.0040	2.3961	2.6682
56	0.6789	1.2969	1.6725	2.0032	2.3948	2.6665
57	0.6788	1.2966	1.6720	2.0025	2.3936	2.6649
58	0.6787	1.2963	1.6716	2.0017	2.3924	2.6633
59	0.6787	1.2961	1.6711	2.0010	2.3912	2.6618
60	0.6786	1.2958	1.6706	2.0003	2.3901	2.6603
61	0.6785	1.2956	1.6702	1.9996	2.3890	2.6589
62	0.6785	1.2954	1.6698	1.9990	2.3880	2.6575
63	0.6784	1.2951	1.6694	1.9983	2.3870	2.6561
64	0.6783	1.2949	1.6690	1.9977	2.3860	2.6549
65	0.6783	1.2947	1.6686	1.9971	2.3851	2.6536
66	0.6782	1.2945	1.6683	1.9966	2.3842	2.6524
67	0.6782	1.2943	1.6679	1.9960	2.3833	2.6512
68	0.6781	1.2941	1.6676	1.9955	2.3824	2.6501
69	0.6781	1.2939	1.6672	1.9949	2.3816	2.6490
70	0.6780	1.2938	1.6669	1.9944	2.3808	2.6479
71	0.6780	1.2936	1.6666	1.9939	2.3800	2.6469
72	0.6779	1.2934	1.6663	1.9935	2.3793	2.6459
73	0.6779	1.2933	1.6660	1.9930	2.3785	2.6449
74	0.6778	1.2931	1.6657	1.9925	2.3778	2.6439
75	0.6778	1.2929	1.6654	1.9921	2.3771	2.6430
76	0.6777	1.2928	1.6652	1.9917	2.3764	2.6421
77	0.6777	1.2926	1.6649	1.9913	2.3758	2.6412
78	0.6776	1.2925	1.6646	1.9908	2.3751	2.6403
79	0.6776	1.2924	1.6644	1.9905	2.3745	2.6395
80	0.6776	1.2922	1.6641	1.9901	2.3739	2.6387
81	0.6775	1.2921	1.6639	1.9897	2.3733	2.6379
82	0.6775	1.2920	1.6636	1.9893	2.3727	2.6371
83	0.6775	1.2918	1.6634	1.9890	2.3721	2.6364
84	0.6774	1.2917	1.6632	1.9886	2.3716	2.6356
85	0.6774	1.2916	1.6630	1.9883	2.3710	2.6349
86	0.6774	1.2915	1.6628	1.9879	2.3705	2.6342
87	0.6773	1.2914	1.6626	1.9876	2.3700	2.6335
88	0.6773	1.2912	1.6624	1.9873	2.3695	2.6329
89	0.6773	1.2911	1.6622	1.9870	2.3690	2.6322
90	0.6772	1.2910	1.6620	1.9867	2.3685	2.6316
91	0.6772	1.2909	1.6618	1.9864	2.3680	2.6309
92	0.6772	1.2908	1.6616	1.9861	2.3676	2.6303
93	0.6771	1.2907	1.6614	1.9858	2.3671	2.6297
94	0.6771	1.2906	1.6612	1.9855	2.3667	2.6291
95	0.6771	1.2905	1.6611	1.9853	2.3662	2.6286

(*continued*)

Table 4 (*continued*)

Degrees of freedom	Upper tail area					
	.25	.10	.05	.025	.01	.005
96	0.6771	1.2904	1.6609	1.9850	2.3658	2.6280
97	0.6770	1.2903	1.6607	1.9847	2.3654	2.6275
98	0.6770	1.2902	1.6606	1.9845	2.3650	2.6269
99	0.6770	1.2902	1.6604	1.9842	2.3646	2.6264
100	0.6770	1.2901	1.6602	1.9840	2.3642	2.6259
110	0.6767	1.2893	1.6588	1.9818	2.3607	2.6213
120	0.6765	1.2886	1.6577	1.9799	2.3578	2.6174
130	0.6764	1.2881	1.6567	1.9784	2.3554	2.6142
140	0.6762	1.2876	1.6558	1.9771	2.3533	2.6114
150	0.6761	1.2872	1.6551	1.9759	2.3515	2.6090
∞	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758

For a particular number of degrees of freedom, each entry represents the critical value of t corresponding to a specified upper tail area α.

$\alpha = .05$

$F_{(\alpha, df_1, df_2)}$
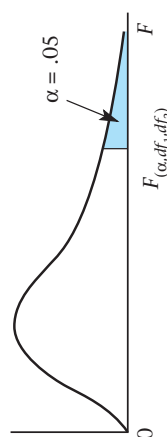
Table 5 F distribution

Numerator df_1

Denominator df_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

(*continued*)

$\alpha = .025$

$F_{(\alpha, df_1, df_2)}$

Table 5 *(continued)*

Numerator df_1

Denominator df_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

(continued)

$\alpha = .01$

$F_{(\alpha, df_1, df_2)}$

Table 5 *(continued)*
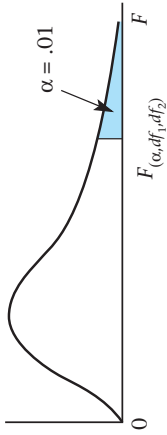
Numerator df_1

Denominator df_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	485	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

For a particular combination of numerator and denominator degrees of freedom, each entry represents the critical value of F corresponding to a specified upper tail area α.

A *posteriori* contrasts Contrasts made after conducting the analysis. These are generally multiple comparison tests.

A *priori* contrasts Contrasts determined before conducting the analysis, based on the researcher's theoretical framework.

Action research A team research process, facilitated by a professional researcher(s), linking with decision-makers and other stakeholders who together wish to improve particular situations.

Agglomerative clustering A hierarchical clustering procedure where each object starts out in a separate cluster. Clusters are formed by grouping objects into bigger and bigger clusters.

Alpha error (α) See Type I error.

Alternative hypothesis A statement that some difference or effect is expected. Accepting the alternative hypothesis will lead to changes in opinions or actions.

Alternative-forms reliability An approach for assessing reliability that requires two equivalent forms of the scale to be constructed and then the same respondents to be measured at two different times.

Analysis of covariance (ANCOVA) An advanced analysis of variance procedure in which the effects of one or more metric-scaled extraneous variables are removed from the dependent variable before conducting the ANOVA.

Analysis of variance (ANOVA) A statistical technique for examining the differences among means for two or more populations.

Analysis sample Part of the total sample used to check the results of the discriminant function.

Analytical model An explicit specification of a set of variables and their interrelationships designed to represent some real system or process in whole or in part.

Analytical services Companies that provide guidance in the development of research design.

Area sampling A common form of cluster sampling in which the clusters consist of geographic areas such as counties, housing tracts, blocks or other area descriptions.

Association techniques A type of projective technique in which respondents are presented with a stimulus and are asked to respond with the first thing that comes to mind.

Asymmetric lambda A measure of the percentage improvement in predicting the value of the dependent variable given the value of the independent variable in contingency table analysis. Lambda also varies between 0 and 1.

Audit A data collection process derived from physical records or performing inventory analysis. Data are collected personally by the researcher, or by representatives of the researcher, and are based on counts usually of physical objects rather than people.

Back translation A translation technique that translates a questionnaire from the base language by a translator whose native language is the one into which the questionnaire is being translated. This version is then retranslated back into the original language by someone whose native language is the base language. Translation errors can then be identified.

Balanced scale A scale with an equal number of favourable and unfavourable categories.

Bar chart A chart that displays data in bars positioned horizontally or vertically.

Bayesian approach A selection method where the elements are selected sequentially. The Bayesian approach explicitly incorporates prior information about population parameters as well as the costs and probabilities associated with making wrong decisions.

Beta error (β) See Type II error.

Bibliographic databases Databases composed of citations to articles in journals, magazines, newspapers, marketing research studies, technical reports, government documents, and the like. They often provide summaries or abstracts of the material cited.

Binominal test A goodness-of-fit statistical test for dichotomous variables. It tests the goodness of fit of the observed number of observations in each category to the number expected under a specified binominal distribution.

Bivariate regression A procedure for deriving a mathematical relationship, in the form of an equation, between a single metric-dependent variable and a single metric-independent variable.

Branching question A question used to guide an interviewer (or respondent) through a survey by directing the interviewer (or respondent) to different spots on the questionnaire depending on the answers given.

Branded market research products Specialised data collection and analysis procedures developed to address specific types of marketing research problems.

Broad statement of the problem The initial statement of the marketing research problem that provides an appropriate perspective on the problem.

Call disposition Call disposition records the outcome of an interview call.

Carryover effects Where the evaluation of a particular scaled item significantly affects the respondent's judgement of subsequent scaled items.

Cartoon tests Cartoon characters are shown in a specific situation related to the problem. Respondents are asked to indicate the dialogue that one cartoon character might make in response to the comment(s) of another character.

Case study A detailed study based upon the observation of the intrinsic details of individuals, groups of individuals and organisations.

Casewise deletion A method for handling missing responses in which cases or respondents with any missing responses are discarded from the analysis.

Category equivalence A construct equivalence issue that deals specifically with whether the categories in which brands, products and behaviour are grouped are the same in different countries.

Causal research A type of conclusive research where the major objective is to obtain evidence regarding cause-and-effect (causal) relationships.

Causality Causality applies when the occurrence of X increases the probability of the occurrence of Y.

Census A complete enumeration of the elements of a population or study objects.

Centroid method A variance method of hierarchical clustering in which the distance between two clusters is the distance between their centroids (means for all the variables).

Characteristic profile An aid to interpreting discriminant analysis results by describing each group in terms of the group means for the predictor variables.

Chi-square distribution A skewed distribution whose shape depends solely on the number of degrees of freedom. As the number of degrees of freedom increases, the chi-square distribution becomes more symmetrical.

Chi-square statistic The statistic used to test the statistical significance of the observed association in a cross-tabulation. It assists us in determining whether a systematic association exists between the two variables.

Classification information Socio-economic and demographic characteristics used to classify respondents.

Cluster sampling A two-step probability sampling technique where the target population is first divided into mutually exclusive and collectively exhaustive sub-populations called clusters, and then a random sample of clusters is selected based on a probability sampling technique such as simple random sampling. For each selected cluster, either all the elements are included in the sample, or a sample of elements is drawn probabilistically.

Codebook A book containing coding instructions and the necessary information about the questions and potential answers in a survey.

Coding Assigning a code to represent a specific response to a specific question along with the data record and column position that the code will occupy.

Coding and data entry services Companies whose primary service offering is their expertise in converting completed surveys or interviews into a usable database for conducting statistical analysis.

Coding data Breaking down qualitative data into discrete chunks and attaching a reference to those chunks of data.

Coefficient alpha A measure of internal consistency reliability that is the average of all possible split-half coefficients resulting from different splittings of the scale items.

Coefficient of variation A useful expression in sampling theory for the standard deviation as a percentage of the mean.

Cohort analysis A multiple cross-sectional design consisting of surveys conducted at appropriate time intervals. The cohort refers to the group of respondents who experience the same event within the same interval.

Common factor analysis An approach to factor analysis that estimates the factors based only on the common variance. Also called principal axis factoring.

Comparative scales One of two types of scaling technique in which there is direct comparison of stimulus objects with one another.

Complete linkage A linkage method that is based on maximum distance of the farthest neighbour approach.

Completion rate The percentage of qualified respondents who complete the interview. It enables researchers to take into account anticipated refusals by people who qualify.

Completion technique A projective technique that requires respondents to complete an incomplete stimulus situation.

Conceptual equivalence A construct equivalence issue that deals with whether the interpretation of brands, products, consumer behaviour and the marketing effort are the same in different countries.

Conclusive research A research design characterised by the measurement of clearly defined marketing phenomena.

Concomitant variation A condition for inferring causality that requires that the extent to which a cause, X, and an effect, Y, occur together or vary together is predicted by the hypothesis under consideration.

Concurrent validity A type of validity that is assessed when the data on the scale being evaluated and on the criterion variables are collected at the same time.

Confidence intervals The range into which the true population parameter will fall, assuming a given level of confidence.

Confounding variables Variables used to illustrate that extraneous variables can confound the results by influencing the dependent variable; synonymous with extraneous variables.

Conjoint analysis A technique that attempts to determine the relative importance consumers attach to salient attributes and the utilities they attach to the levels of attributes.

Conjoint analysis model The mathematical model expressing the fundamental relationship between attributes and utility in conjoint analysis.

Consistency checks A part of the data cleaning process that identifies data that are out of range, logically inconsistent, or have extreme values. Data with values not defined by the coding scheme are inadmissible.

Constant sum scaling A comparative scaling technique in which respondents are required to allocate a constant sum of units such as points, euros, chits, stickers or chips among a set of stimulus objects with respect to some criterion.

Construct equivalence A type of equivalence that deals with the question of whether the marketing constructs have the same meaning and significance in different countries.

Construct validity A type of validity that addresses the question of what construct or characteristic the scale is measuring. An attempt is made to answer theoretical questions of why a scale works and what deductions can be made concerning the theory underlying the scale.

Construction technique A projective technique in which respondents are required to construct a response in the form of a story, dialogue or description.

Content analysis The objective, systematic and quantitative description of the manifest content of a communication.

Content validity A type of validity, sometimes called face validity, that consists of a subjective but systematic evaluation of the representativeness of the content of a scale for the measuring task at hand.

Contingency coefficient A measure of the strength of association in a table of any size.

Contingency table A cross-tabulation table. It contains a cell for every combination of categories of the two variables.

Continuous rating scale A measurement scale that has respondents rate the objects by placing a mark at the appropriate position on a line that runs from one extreme of the criterion variable to the other. The form may vary considerably. Also called graphic rating scale.

Contrasts In ANOVA, a method of examining differences among two or more means of the treatment groups.

Contrived observation Observing behaviour in an artificial environment.

Controlled test market A test-marketing programme conducted by an outside research company in field experimentation. The research company guarantees distribution of the product in retail outlets that represent a predetermined percentage of the market.

Convenience sampling A non-probability sampling technique that attempts to obtain a sample of convenient elements. The selection of sampling units is left primarily to the interviewer.

Convergent validity A measure of construct validity that measures the extent to which the scale correlates positively with other measures of the same construct.

Cookie technology A group of letters and numbers stored in a Web surfer's browser that identify their computer.

Correspondence analysis An MDS technique for scaling qualitative data that scales the rows and columns of the input contingency table in corresponding units so that each can be displayed in the same low-dimensional space.

Covariance A systematic relationship between two variables in which a change in one implies a corresponding change in the other (COV_{xy}).

Covariate A metric-independent variable used in ANCOVA.

Cramer's V A measure of the strength of association used in tables larger than 2×2.

Criterion validity A type of validity that examines whether the measurement scale performs as expected in relation to other selected variables as meaningful criteria.

Critical request The target behaviour being researched.

Cross-cultural analysis A type of across-countries analysis in which the data could be aggregated for each country and these aggregate statistics analysed.

Cross-sectional design A type of research design involving the collection of information from any given sample of population elements only once.

Cross-tabulation A statistical technique that describes two or more variables simultaneously and results in tables that reflect the joint distribution of two or more variables that have a limited number of categories or distinct values.

Cross-validation A test of validity that examines whether a model holds on comparable data not used in the original estimation.

Customer database A database that details characteristics of customers and prospects that can include names and addresses, geographic, demographic and buying behaviour data.

Customised services Companies that tailor research procedures to best meet the needs of each client.

Data analysis services Firms whose primary service is to conduct statistical analysis of quantative data.

Data assembly The gathering of data from a variety of disparate sources.

Data cleaning Thorough and extensive checks for consistency and treatment of missing responses.

Data display Involves summarising and presenting the structure that is seen in collected qualitative data.

Data mining The process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition as well as statistical and mathematical techniques.

Data reduction The organising and structuring of qualitative data.

Data verification Involves seeking alternative explanations of the interpretations of qualitative data, through other data sources.

Datawarehouses This may be seen as a 'super database', but more specifically it may be defined as a process of gathering disparate data from database and survey sources, and converting it into a consistent format that can aid business decision-making.

Debriefing After a disguised experiment, informing test subjects what the experiment was about and how the experimental manipulations were performed.

Decision support system (DSS) An information system that enables decision-makers to interact directly with both databases and analysis models. The important components of a DSS include hardware and a communication network, database, model base, software base and the DSS user (decision-maker).

Decomposition of the total variation In one-way ANOVA, separation of the variation observed in the dependent variable into the variation due to the independent variables plus the variation due to error.

Deduction A form of reasoning in which a conclusion is validly inferred from some premises, and must be true if those premises are true.

Demand artefacts Responses given because the respondents attempt to guess the purpose of the experiment and respond accordingly.

Dependence techniques Multivariate techniques appropriate when one or more of the variables can be identified as dependent variables and the remaining ones as independent variables.

Dependent variables Variables that measure the effect of the independent variables on the test units.

Depth interview An unstructured, direct, personal interview in which a single respondent is probed by an experienced interviewer to uncover underlying motivations, beliefs, attitudes and feelings on a topic.

Derived approaches In MDS, attribute-based approaches to collecting perception data requiring respondents to rate the stimuli on the identified attributes using semantic differential or Likert scales.

Descriptive research A type of conclusive research that has as its major objective the description of something, usually market characteristics or functions.

Design control A method of controlling extraneous variables that involves using specific experimental designs.

Determinism A doctrine espousing that everything that happens is determined by a necessary chain of causation.

Diary media panels A data gathering technique composed of samples of respondents whose television viewing behaviour is automatically recorded by electronic devices, supplementing the purchase information recorded in a diary.

Diary purchase panels A data gathering technique in which respondents record their purchases in a diary.

Dichotomous question A structured question with only two response alternatives, such as yes and no.

Direct approach A type of qualitative research in which the purposes of the project are disclosed to the respondent or are obvious given the nature of the interview.

Direct method An approach to discriminant analysis that involves estimating the discriminant function so that all the predictors are included simultaneously.

Directory databases Databases that provide information on individuals, organisations and services.

Discriminant analysis A technique for analysing marketing research data when the criterion or dependent variable is categorical and the predictor or independent variables are interval in nature.

Discriminant analysis model The statistical model on which discriminant analysis is based.

Discriminant function The linear combination of independent variables developed by discriminant analysis that will best discriminate between the categories of the dependent variable.

Discriminant validity A type of construct validity that assesses the extent to which a measure does not correlate with other constructs from which it is supposed to differ.

Disordinal interaction The change in the rank order of the effects of one factor across the levels of another.

Divisive clustering A hierarchical clustering procedure where all objects start out in one giant cluster. Clusters are formed by dividing this cluster into smaller and smaller clusters.

Double cross-validation A special form of validation in which the sample is split into halves. One half serves as the estimation sample and the other as a validation sample. The roles of the estimation and validation halves are then reversed and the cross-validation process is repeated.

Double sampling A sampling technique in which certain population elements are sampled twice.

Double-barrelled question A single question that attempts to cover two issues. Such questions can be confusing to respondents and result in ambiguous responses.

Dummy variables A re-specification procedure using variables that take on only two values, usually 0 or 1.

Editing A review of the questionnaires with the objective of increasing accuracy and precision.

Elbow criterion A plot of stress versus dimensionality used in MDS. The point at which an elbow or a sharp bend occurs indicates an appropriate dimensionality.

Electronic observation An observational research strategy in which electronic devices, rather than human observers, record the phenomena being observed.

Element An object that possesses the information sought by the researcher and about which inferences are to be made.

Empiricism A theory of knowledge. A broad category of the philosophy of science that locates the source of all knowledge in experience.

Estimated or predicted value The value $Y_i = a + b_x$, where a and b are, respectively, estimators of β_0 and β_1, the corresponding population parameters.

Ethnography A research approach based upon the observation of the customs, habits and differences between people in everyday situations.

Euclidean distance The square root of the sum of the squared differences in values for each variable.

Evolving research design A research design where particular research techniques are chosen as the researcher develops an understanding of the issues and respondents.

Experiment The process of manipulating one or more independent variables and measuring their effect on one or more dependent variables, while controlling for the extraneous variables.

Experimental design The set of experimental procedures specifying (1) the test units and sampling procedures, (2) the independent variables, (3) the dependent variables, and (4) how to control the extraneous variables.

Experimental group An initial focus group run to test the setting of the interview, the opening question, the topic guide and the mix of respondents that make up a group.

Exploratory research A research design characterised by a flexible and evolving approach to understand marketing phenomena that are inherently difficult to measure.

Expressive techniques Projective techniques in which respondents are presented with a verbal or visual situation and are asked to relate the feelings and attitudes of other people to the situation.

External analysis of preferences A method of configuring a spatial map such that the ideal points or vectors based on preference data are fitted in a spatial map derived from the perception data.

External data Data that originate outside the organisation.

External suppliers Outside marketing research companies hired to supply marketing research services.

External validity A determination of whether the cause-and-effect relationships found in the experiment can be generalised.

Extraneous variables Variables other than dependent and independent variables which may influence the results of an experiment.

Eye tracking equipment Instruments that record the gaze movements of the eye.

F distribution A frequency distribution that depends upon two sets of degrees of freedom: the degrees of freedom in the numerator and the degrees of freedom in the denominator.

F statistic The ratio of two sample variances.

F test A statistical test of the equality of the variances of two populations.

Factor An underlying dimension that explains the correlations among a set of variables.

Factor analysis A class of procedures primarily used for data reduction and summarisation.

Factor scores Composite scores estimated for each respondent on the derived factors.

Factorial design A statistical experimental design used to measure the effects of two or more independent variables at various levels and to allow for interactions between variables.

Factors Categorical independent variables in ANOVA. The independent variables must all be categorical (non-metric) to use ANOVA.

Field environment An experimental location set in actual market conditions.

Field force Both the actual interviewers and the supervisors involved in data collection.

Field notes A log or diary of observations, events and reflections made by a researcher as a study is planned, implemented and analysed.

Field services Companies whose primary service is offering their expertise in collecting data for research projects.

Filter question An initial question in a questionnaire that screens potential respondents to ensure they meet the requirements of the sample.

Fixed-response alternative questions Questions that require respondents to choose from a set of predetermined answers.

Focus group A discussion conducted by a trained moderator among a small group of respondents in an unstructured and natural manner.

Forced rating scale A rating scale that forces respondents to express an opinion because a 'no opinion' or 'no knowledge' option is not provided.

Frequency distribution A mathematical distribution whose objective is to obtain a count of the number of responses associated with different values of one variable and to express these counts in percentage terms.

Friendship pair A technique used to interview children as two friends or classmates together.

Frugging The use of marketing research to deliberately disguise fundraising activities.

Full-service suppliers Companies that offer a full range of marketing research activities.

Full-text databases Databases that contain the complete text of secondary source documents comprising the database.

Functional equivalence A construct equivalence issue that deals specifically with whether a given concept or behaviour serves the same role or function in different countries.

Funnel approach A strategy for ordering questions in a questionnaire in which the sequence starts with the general questions, which are followed by progressively specific questions, to prevent specific questions from biasing general questions.

Galvanic skin response Changes in the electrical resistance of the skin that relate to a respondent's affective state.

Gamma A test statistic that measures the association between two ordinal-level variables. It does not make an adjustment for ties.

Generalisability The degree to which a study based on a sample applies to the population as a whole.

Geodemographic classification This groups consumers together based on the types of neighbourhood in which they live. If a set of neighbourhoods are similar across a wide range of demographic measures, they may also offer similar potential across most products, brands, services and media.

Geodemographic information system (GIS) At a base level, a GIS matches geographic information with demographic information. This match allows subsequent data analyses to be presented on maps.

Graphic rating scale See Continuous rating scale.

Graphical models Analytical models that provide a visual picture of the relationships between variables.

Grounded theory Theory derived from data, systematically gathered and analysed.

Hierarchical clustering A clustering procedure characterised by the development of a hierarchy or treelike structure.

Histogram A vertical bar chart in which the height of the bars represents the relative or cumulative frequency of occurrence.

History Specific events that are external to the experiment but that occur at the same time as the experiment.

Hit ratio The percentage of cases correctly classified by the discriminant analysis.

Hybrid conjoint analysis A form of conjoint analysis that can simplify the data collection task and estimate selected interactions as well as all main effects.

Hypertext markup language (HTML) The language of the Web.

Hypothesis An unproven statement or proposition about a factor or phenomenon that is of interest to the researcher.

Identification information A type of information obtained in a questionnaire that includes name, address and phone number.

Implicit alternative An alternative that is not explicitly expressed.

Implicit assumptions An assumption that is not explicitly stated in a question.

Imputation A method to adjust for non-response by assigning the characteristic of interest to the non-respondents based on the similarity of the variables available for both non-respondents and respondents.

Incidence rate The rate of occurrence of persons eligible to participate in the study expressed as a percentage.

Independent samples Two samples that are not experimentally related. The measurement of one sample has no effect on the values of the second sample.

Independent variables Variables that are manipulated by the researcher and whose effects are measured and compared.

Indirect approach A type of qualitative research in which the purposes of the project are disguised from the respondents.

Induction A form of reasoning that usually involves the inferences that an instance or repeated combination of events may be usually generalised.

Instrumentation An extraneous variable involving changes in the measuring instrument, in the observers, or in the scores themselves.

Interaction When assessing the relationship between two variables, an interaction occurs if the effect of X_1 depends on the level of X_2, and vice versa.

Interactive testing effect An effect in which a prior measurement affects the test unit's response to the independent variable.

Interdependence technique A multivariate statistical technique in which the whole set of inter-dependent relationships is examined.

Interdependence techniques Multivariate statistical techniques that attempt to group data based on underlying similarity and thus allow for interpretation of the data structures. No distinction is made as

to which variables are dependent and which are independent.

Internal analysis of preferences A method of configuring a spatial map such that the spatial map represents both brands or stimuli and respondent points or vectors and is derived solely from the preference data.

Internal consistency reliability An approach for assessing the internal consistency of the set of items, where several items are summated in order to form a total score for the scale.

Internal data Data available within the organisation for whom the research is being conducted.

Internal supplier Marketing research department located within a firm.

Internal validity A measure of accuracy of an experiment. It measures whether the manipulation of the independent variables, or treatments, actually caused the effects on the dependent variable(s).

Internet databases Databases that can be accessed, searched and analysed on the Internet. It is also possible to download data from the Internet and store it on the computer or an auxiliary device.

Internet services Companies which specialise in the use of the Internet to collect, analyse and distribute marketing research information.

Interquartile range The range of a distribution encompassing the middle 50% of the observations.

Interval scale A scale in which the numbers are used to rank objects such that numerically equal distances on the scale represent equal distances in the characteristic being measured.

Intra-cultural analysis Within-country analysis of international data.

Item equivalence Use of the same instrument in different countries.

Itemised rating scale A measurement scale having numbers or brief descriptions associated with each category. The categories are ordered in terms of scale position.

Judgemental sampling A form of convenience sampling in which the population elements are purposely selected based on the judgement of the researcher.

***k*-sample median test** A non-parametric test used to examine differences among more than two groups when the dependent variable is measured on an ordinal scale.

Kolmogorov-Smirnov (K-S) one-sample test A one-sample non-parametric goodness-of-fit test that compares the cumulative distribution function for a variable with a specified distribution.

Kolmogorov-Smirnov (K-S) two-sample test Non-parametric test statistic that determines whether two distributions are the same. It takes into account any differences in the two distributions, including median, dispersion and skewness.

Kruskal-Wallis one-way analysis of variance A non-metric ANOVA test that uses the rank value of each case, not merely its location relative to the median.

Kurtosis A measure of the relative peakedness of the curve defined by the frequency distribution.

Laboratory environment An artificial setting for experimentation in which the researcher constructs the desired conditions.

Laddering A technique for conducting depth interviews in which a line of questioning proceeds from product characteristics to user characteristics.

Latin square design A statistical design that allows for the statistical control of two non-interacting external variables in addition to the manipulation of the independent variable.

Leading question A question that gives the respondent a clue as to what the answer should be.

Least squares procedure A technique for fitting a straight line into a scattergram by minimising the vertical distances of all the points from the line.

Level of significance The probability of making a type I error.

Lifestyles Distinctive patterns of living described by the activities people engage in, the interests they have, and the opinions they hold of themselves and the world around them.

Likert scale A measurement scale with five response categories ranging from 'strongly disagree' to 'strongly agree' that requires respondents to indicate a degree of agreement or disagreement with each of a series of statements related to the stimulus objects.

Limited-service suppliers Companies that specialise in one or a few phases of a marketing research project.

Line chart A chart that connects a series of data points using continuous lines.

Linguistic equivalence The equivalence of both spoken and written language forms used in scales and questionnaires.

Linkage methods Agglomerative methods of hierarchical clustering that cluster objects based on a computation of the distance between them.

Longitudinal design A type of research design involving a fixed sample of population elements measured repeatedly. The sample remains the same over time, thus providing a series of pictures that, when viewed together, portray a vivid illustration of the situation and the changes that are taking place.

Loyalty card At face value, a sales promotion device used by supermarkets, pharmacists, department stores, petrol stations and even whole shopping centres and towns to encourage repeat purchases. For the marketing researcher, the loyalty card is a device that can link customer characteristics to actual product purchases.

Mahalanobis procedure A stepwise procedure used in discriminant analysis to maximise a generalised measure of the distance between the two closest groups.

Mail panel A large and nationally representative sample of households that have agreed to participate in periodic mail questionnaires, product tests and telephone surveys.

Main testing effect An effect of testing occurring when a prior observation affects a later observation.

Mann-Whitney *U* test A statistical test for a variable measured on an ordinal scale, comparing the differences in the location of two populations based on observations from two independent samples.

Marketing decision problem The problem confronting the marketing decision-maker, which asks what the decision-maker has to do.

Marketing information systems (MkIS) A formalised set of procedures for generating, analysing, sorting and distributing pertinent information to marketing decision-makers on an ongoing basis.

Marketing intelligence Qualified observations of events and developments in the marketing environment.

Marketing research A key element within the total field of marketing information. It links the consumer, customer and public to the market through information which is used to identify and define marketing opportunities and problems; to generate, refine and evaluate marketing actions; and to improve understanding of marketing as a process and of the ways in which specific marketing activities can be made more effective.

Marketing research problem A problem that entails determining what information is needed and how it can be obtained in the most feasible way.

Marketing research process A set of six steps which define the tasks to be accomplished in conducting a marketing research study. These include problem definition, developing an approach to the problem, research design formulation, fieldwork, data preparation and analysis, and report generation and presentation.

Matching A method of controlling extraneous variables that involves matching test units on a set of key background variables before assigning them to the treatment conditions.

Mathematical models Analytical models that explicitly describe the relationship between variables, usually in equation form.

Maturation An extraneous variable attributable to changes in the test units themselves that occur with the passage of time.

Mean The average; that value obtained by summing all elements in a set and dividing by the number of elements.

Measure of location A statistic that describes a location within a data set. Measures of central tendency describe the centre of the distribution.

Measure of variability A statistic that indicates the distribution's dispersion.

Measurement The assignment of numbers or other symbols to characteristics of objects according to certain pre-specified rules.

Measurement error The variation in the information sought by the researcher and the information generated by the measurement process employed.

Median A measure of central tendency given as the value above which half of the values fall and below which half of the values fall.

Metric data Data that are interval or ratio in nature.

Metric equivalence See Scalar equivalence.

Metric MDS A multidimensional scaling method that assumes that input data are metric.

Metric scale A scale that is either interval or ratio in nature.

Missing responses Values of a variable that are unknown because the respondents concerned provided ambiguous answers to the question or because their answers were not properly recorded.

Mode A measure of central tendency given as the value that occurs with the most frequency in a sample distribution.

Moderator An individual who conducts a focus group interview, by setting the purpose of the interview, questioning, probing and handling the process of discussion.

Monadic scale See Non-comparative scale.

Mood board A collage created in a focus group setting. Focus group respondents are asked to snip words and pictures from magazines that they see as representing the values a particular brand is perceived to have. In some circumstances, collages can also be made up from audio and video tapes.

Mortality An extraneous variable attributable to the loss of test units while the experiment is in progress.

Multicollinearity A state of high intercorrelations among independent variables.

Multidimensional scaling (MDS) A class of procedures for representing perceptions and preferences of respondents spatially by means of a visual display.

Multiple comparison tests *A posteriori* contrasts that enable the researcher to construct generalised confidence intervals that can be used to make pairwise comparisons of all treatment means.

Multiple cross-sectional design A cross-sectional design in which there are two or more samples of respondents, and information from each sample is obtained only once.

Multiple discriminant analysis Discriminant analysis technique where the criterion variable involves three or more categories.

Multiple regression A statistical technique that simultaneously develops a mathematical relationship between two or more independent variables and an interval-scaled dependent variable.

Multiple regression model An equation used to explain the results of multiple regression analysis.

Multiple time series design A time series design that includes another group of test units to serve as a control group.

Multiple η^2 The strength of the joint effect of two (or more) factors, or the overall effect.

Multivariate analysis of variance (MANOVA) An ANOVA technique using two or more metric dependent variables.

Multivariate techniques Statistical techniques suitable for analysing data when there are two or more measurements on each element and the variables are analysed simultaneously. Multivariate techniques are concerned with the simultaneous relationships among two or more phenomena.

Mystery shopper An observer visiting providers of goods and services as if they were really a customer, and recording characteristics of the service delivery.

***n*-way analysis of variance** An ANOVA model where two or more factors are involved.

Natural observation Observing behaviour as it takes place in the environment.

Nominal scale A scale whose numbers serve only as labels or tags for identifying and classifying objects with a strict one-to-one correspondence between the numbers and the objects.

Nomological validity A type of validity that assesses the relationship between theoretical constructs. It seeks to confirm significant correlations between the constructs as predicted by a theory.

Non-comparative scale One of two types of scaling techniques in which each stimulus object is scaled independently of the other objects in the stimulus set. Also called monadic scale.

Non-hierarchical clustering A procedure that first assigns or determines a cluster centre and then groups all objects within a pre-specified threshold value from the centre.

Non-metric analysis of variance An ANOVA technique for examining the difference in the central tendencies of more than two groups when the dependent variable is measured on an ordinal scale.

Non-metric correlation A correlation measure for two non-metric variables that relies on rankings to compute the correlation.

Non-metric data Data derived from a nominal or ordinal scale.

Non-metric MDS A type of multidimensional scaling which assumes that the input data are ordinal.

Non-metric scale A scale that is either nominal or ordinal in nature.

Non-parametric tests Hypothesis testing procedures that assume that the variables are measured on a nominal or ordinal scale.

Non-probability sampling Sampling techniques that do not use chance selection procedures but rather rely on the personal judgement of the researcher.

Non-response bias Bias caused when actual respondents differ from those who refuse to participate.

Non-response error A type of non-sampling error that occurs when some of the respondents included in the sample do not respond. This error may be defined as the variation between the true mean value of the variable in the original sample and the true mean value in the net sample.

Non-sampling error An error that can be attributed to sources other than sampling and that can be random or non-random.

Null hypothesis A statement in which no difference or effect is expected. If the null hypothesis is not rejected, no changes will be made.

Numeric databases Databases containing numerical and statistical information that may be important sources of secondary data.

Objective evidence Perceived to be unbiased evidence, supported by empirical findings.

Oblique rotation Rotation of factors when the axes are not maintained at right angles.

Offline databases Databases that are available on diskette or CD-ROM.

Omega squared (ω^2) A measure indicating the proportion of the variation in the dependent variable that is related to a particular independent variable or factor.

Omnibus survey A distinctive form of survey that serves the needs of a syndicated group. The omnibus survey targets particular types of respondents such as those in specific geographic locations, e.g. Luxembourg residents, or consumers of particular types of products, e.g. business air travellers. With that target group of respondents, a core set of questions can be asked, with other questions added as syndicate members wish.

One-group pre-test–post-test design A pre-experimental design in which a group of test units is measured twice.

One-shot case study A pre-experimental design in which a single group of test units is exposed to a treatment **X**, and then a single measurement of the dependent variable is taken.

One-tailed test A test of the null hypothesis where the alternative hypothesis is expressed directionally.

One-way analysis of variance An ANOVA technique in which there is only one factor.

Online databases Databases, stored in computers, that require a telecommunications network to access.

Operational data Data generated about an organisation's customers, through day-to-day transactions.

Operational equivalence A type of equivalence that measures how theoretical constructs are operationalised in different countries to measure marketing variables.

Operationalised The derivation of measurable characteristics to encapsulate marketing phenomena, e.g. the concept of 'customer loyalty' can be operationalised through measurements such as frequency of repeat purchases or the number of years that a business relationship has existed.

Optimising partitioning method A non-hierarchical clustering method that allows for later reassignment of objects to clusters to optimise an overall criterion.

Order bias (position bias) A respondent's tendency to choose an alternative merely because it occupies a certain position or is listed in a certain order.

Ordinal interaction An interaction where the rank order of the effects attributable to one factor does not change across the levels of the second factor.

Ordinal scale A ranking scale in which numbers are assigned to objects to indicate the relative extent to which some characteristic is possessed. Thus, it is possible to determine whether an object has more or less of a characteristic than some other object.

Orthogonal rotation Rotation of factors in which the axes are maintained at right angles.

Paired comparison scaling A comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object in the pair according to some criterion. The data obtained are ordinal in nature.

Paired samples In hypothesis testing, the observations are paired so that the two sets of observations relate to the same respondents.

Paired samples *t* test A test for differences in the means of paired samples.

Pairwise deletion A method for handling missing responses in which all cases or respondents with any missing responses are not automatically discarded; rather, for each calculation, only the cases or respondents with complete responses are considered.

Pan-cultural analysis Across-countries analysis in which the data for all respondents from all the countries are pooled and analysed.

Panel A sample of respondents who have agreed to provide information at specified intervals over an extended period.

Pantry audit A type of audit where the researcher inventories the brands, quantities and package sizes of products in a consumer's home.

Paradigm A set of assumptions consisting of agreed-upon knowledge, criteria of judgement, problem fields and ways to consider them.

Parallel threshold method A non-hierarchical clustering method that specifies several cluster centres at once. All objects that are within a pre-specified threshold value from the centre are grouped together.

Parallel translation A translation method in which a committee of translators, each of whom is fluent in at least two languages, discusses alternative versions of a questionnaire and makes modifications until consensus is reached

Parametric tests Hypothesis testing procedures that assume that the variables of interest are measured on at least an interval scale.

Part correlation coefficient A measure of the correlation between *Y* and *X* when the linear effects of the other independent variables have been removed from *X* (but not from *Y*).

Partial correlation coefficient A measure of the association between two variables after controlling or adjusting for the effects of one or more additional variables.

Perceived respondent anonymity The respondents' perceptions that their identities will not be discerned by the interviewer or researcher.

Personal observation An observational research strategy in which human observers record the phenomenon being observed as it occurs.

Phi coefficient (Φ) A measure of the strength of association in the special case of a table with two rows and two columns (a 2×2 table).

Picture response technique A projective technique in which respondents are shown a picture and are asked to tell a story describing it.

Pie chart A round chart divided into sections.

Pilot-testing Testing the questionnaire on a small sample of respondents for the purpose of improving the questionnaire by identifying and eliminating potential problems.

Population The aggregate of all the elements, sharing some common set of characteristics, that comprise the universe for the purpose of the marketing research problem.

Position bias See Order bias.

Positivism A philosophy of language and logic consistent with an empiricist philosophy of science.

Post-test-only control group design Experimental design in which the experimental group is exposed to the treatment but the control group is not and no pre-test measure is taken.

Power of a statistical test The probability of rejecting the null hypothesis when it is in fact false and should be rejected.

Pre-coding In questionnaire design, assigning a code to every conceivable response before data collection.

Predictive validity A type of validity that is concerned with how well a scale can forecast a future criterion.

Pre-experimental designs Designs that do not control for extraneous factors by randomisation.

Pre-test–post-test control group design An experimental design in which the experimental group is exposed to the treatment but the control group is not. Pre-test and post-test measures are taken on both groups.

Primary data Data originated by the researcher specifically to address the research problem.

Principal axis factoring See Common factor analysis.

Principal components analysis An approach to factor analysis that considers the total variance in the data.

Probability proportionate to size (PPS) A selection method where the probability of selecting a sampling unit in a selected cluster varies inversely with the size of the cluster. Therefore, the size of all the resulting clusters is approximately equal.

Probability sampling A sampling procedure in which each element of the population has a fixed probabilistic chance of being selected for the sample.

Probing A motivational technique used when asking survey questions to induce the respondents to enlarge on, clarify or explain their answers and to help the respondents focus on the specific content of the interview.

Problem audit A comparative examination of a marketing problem to understand its origin and nature.

Problem definition A broad statement of the general problem and identification of the specific components of the marketing research problem.

Problem identification research Research undertaken to help identify problems that are not necessarily apparent on the surface, yet exist or are likely to arise in the future.

Problem-solving research Research undertaken to help solve marketing problems.

Product moment correlation (r) A statistic summarising the strength of association between two metric variables.

Projective technique An unstructured and indirect form of questioning that encourages respondents to project their underlying motivations, beliefs, attitudes or feelings regarding the issues of concern.

Psycho-galvanometer An instrument that measures a respondent's galvanic skin response.

Psychographics Quantified profiles of individuals based upon lifestyle characteristics.

Pupilometer An instrument that measures changes in the eye pupil diameter.

Q-sort scaling A comparative scaling technique that uses a rank order procedure to sort objects based on similarity with respect to some criterion.

Qualitative research An unstructured, primarily exploratory design based on small samples, intended to provide insight and understanding.

Quantitative observation The recording and counting of behavioural patterns of people, objects and events in a systematic manner to obtain information about the phenomenon of interest.

Quantitative research Research techniques that seek to quantify data and, typically, apply some form of statistical analysis.

Quasi-experimental designs Designs that apply part of the procedures of true experimentation yet lack full experimental control.

Questionnaire A structured technique for data collection consisting of a series of questions, written or verbal, that a respondent answers.

Quota sampling A non-probability sampling technique that is two-stage restricted judgemental sampling. The first stage consists of developing control categories or quotas of population elements. In the second stage, sample elements are selected based on convenience or judgement.

Random error An error that arises from random changes or differences in respondents or measurement situations.

Random sampling error The error because the particular sample selected is an imperfect repesentation of the population of interest. It may be defined as the variation between the true mean value for the sample and the true mean value of the population.

Randomisation A method of controlling extraneous variables that involves randomly assigning test units to experimental groups by using random numbers. Treatment conditions are also randomly assigned to experimental groups.

Randomised block design A statistical design in which the test units are blocked on the basis of an external variable to ensure that the various experimental and control groups are matched closely on that variable.

Range The difference between the smallest and largest values of a distribution.

Rank order scaling A comparative scaling technique in which respondents are presented with several objects simultaneously and asked to order or rank them according to some criterion.

Ratio scale The highest scale. This scale allows the researcher to identify or classify objects, rank order the objects, and compare intervals or differences. It is also meaningful to compute ratios of scale values.

Regression analysis A statistical procedure for analysing associative relationships between a metric-dependent variable and one or more independent variables.

Reliability The extent to which a scale produces consistent results if repeated measurements are made on the characteristic.

Repeated measures analysis of variance An ANOVA technique used when respondents are exposed to more than one treatment condition and repeated measurements are obtained.

Research brief A document produced by the users of research findings or the buyers of a piece of marketing research. The brief is used to communicate the perceived requirements of a marketing research project.

Research design A framework or blueprint for conducting the marketing research project. It specifies the details of the procedures necessary for obtaining the information needed to structure or solve marketing research problems.

Research proposal The official layout of the planned marketing research activity.

Research questions Refined statements of the specific components of the problem.

Residual The difference between the observed value of Y_i and the value predicted by the regression equation $\hat{Y}_i$.

Response error A type of non-sampling error arising from respondents who do respond but who give inaccurate answers or whose answers are mis-recorded or mis-analysed. It may be defined as a variation between the true mean value of the variable in the net sample and the observed mean value obtained in the market research project.

Response latency The amount of time it takes to respond to a question.

Response rate The percentage of the total attempted interviews that are completed.

Role playing Respondents are asked to assume the behaviour of someone else.

Runs test A test of randomness for a dichotomous variable.

Sample A subgroup of the elements of the population selected for participation in the study.

Sample control The ability of the survey mode to reach the units specified in the sample effectively and efficiently.

Sample size The number of elements to be included in a study.

Sampling control An aspect of supervising that ensures that the interviewers strictly follow the sampling plan rather than select sampling units based on convenience or accessibility.

Sampling distribution The distribution of the values of a sample statistic computed for each possible sample that could be drawn from the target population under a specified sampling plan.

Sampling frame A representation of the elements of the target population that consists of a list or set of directions for identifying the target population.

Sampling unit An element, or a unit containing the element, that is available for selection at some stage of the sampling process.

Sampling with replacement A sampling technique in which an element can be included in the sample more than once.

Sampling without replacement A sampling technique in which an element cannot be included in the sample more than once.

Scalar equivalence The demonstration that two individuals from different countries with the same value on some variable will score at the same level on the same test. Also called metric equivalence

Scale transformation A manipulation of scale values to ensure compatibility with other scales or otherwise to make the data suitable for analysis.

Scaling The generation of a continuum upon which measured objects are located.

Scanner data Data obtained by passing merchandise over a laser scanner that reads the UPC code from the packages.

Scanner diary panels Scanner data where panel members are identified by an ID card, allowing information about each panel member's purchases to be stored with respect to the the individual shopper.

Scanner diary panels with cable TV The combination of a scanner diary panel with manipulations of the advertising that is being broadcast by cable television companies.

Scanning device Technology that reads the UPC code from merchandise by passing it over a laser scanner.

Secondary data Data collected for some purpose other than the problem at hand.

Selection bias An extraneous variable attributable to the improper assignment of test units to treatment conditions.

Semantic differential A seven-point rating scale with end points associated with bipolar labels.

Sentence completion A projective technique in which respondents are presented with a number of incomplete sentences and are asked to complete them.

Sequential sampling A probability sampling technique in which the population elements are sampled sequentially, data collection and analysis are done at each stage, and a decision is made as to whether additional population elements should be sampled.

Sign test A non-parametric test for examining differences in the location of two populations, based on paired populations, that compares only the signs of the differences between pairs of variables without taking into account the magnitude of the differences.

Significance of the interaction effect A test of the significance of the interaction between two or more independent variables.

Significance of the main effect of each factor A test of the significance of the main effect for each individual factor.

Significance of the overall effect A test that some differences exist between some of the treatment groups.

Simple random sampling (SRS) A probability sampling technique in which each element has a known and equal probability of selection. Every element is selected independently of every other element, and the sample is drawn by a random procedure from a sampling frame.

Simulated test market A quasi-test market in which respondents are preselected; they are then interviewed and observed on their purchases and attitudes towards the product.

Single cross-sectional design A cross-sectional design in which one sample of respondents is drawn from the target population and information is obtained from this sample once.

Single linkage A linkage method based on minimum distance or the nearest neighbour rule.

Skewness A characteristic of a distribution that assesses its symmetry about the mean.

Snowball sampling A non-probability sampling technique in which an initial group of respondents is selected randomly. Subsequent respondents are selected based on the referrals or information provided by the initial respondents. By obtaining referrals from referrals, this process may be carried out in waves.

Social desirability The tendency of respondents to give answers that may not be accurate but may be desirable from a social standpoint.

Solomon four-group design An experimental design that explicitly controls for interactive testing effects, in addition to controlling for all the other extraneous variables.

Special-purpose databases Databases that contain information of a specific nature, e.g. data on a specialised industry.

Specific components of the problem The second part of the marketing research problem definition that focuses on the key aspects of the problem and provides clear guidelines on how to proceed further.

Split-half reliability A form of internal consistency reliability in which the items constituting the scale are divided into two halves and the resulting half scores are correlated.

Standard deviation The square root of the variance.

Standard error The standard deviation of the sampling distribution of the mean or proportion.

Standard test market A test market in which the product is sold through regular distribution channels. For example, no special considerations are given to products simply because they are being test-marketed.

Standardisation The process of correcting data to reduce them to the same scale by subtracting the sample mean and dividing by the standard deviation.

Standardised services Companies that use standardised procedures to provide marketing research to various clients.

Stapel scale A scale for measuring attitudes that consists of a single adjective in the middle of an even-numbered range of values.

Static group A pre-experimental design in which there are two groups: the experimental group (EG), which is exposed to the treatment, and the control group (CG). Measurements on both groups are made only after the treatment, and test units are not assigned at random.

Statistical control A method of controlling extraneous variables by measuring the extraneous variables and adjusting for their effects through statistical methods.

Statistical designs Designs that allow for the statistical control and analysis of external variables.

Statistical inference The process of generalising the sample results to a target population.

Statistical regression An extraneous variable that occurs when test units with extreme scores move closer to the average score during the course of the experiment.

Stepwise discriminant analysis Discriminant analysis in which the predictors are entered sequentially based on their ability to discriminate between the groups.

Stepwise regression A regression procedure in which the predictor variables enter or leave the regression equation one at a time.

Story completion A projective technique in which respondents are provided with part of a story and are required to give the conclusion in their own words.

Stratified sampling A probability sampling technique that uses a two-step process to partition the population into subsequent sub-populations, or strata. Elements are selected from each stratum by a random procedure.

Structured data collection Use of a formal questionnaire that presents questions in a prearranged order.

Structured observation Observation where the researcher clearly defines the behaviours to be observed and the techniques by which they will be measured.

Structured questions Questions that pre-specify the set of response alternatives and the response format. A structured question could be multiple-choice, dichotomous or a scale.

Substitution A procedure that substitutes for non-respondents other elements from the sampling frame who are expected to respond.

Sugging The use of marketing research to deliberately disguise a sales effort.

Surrogate variables A subset of original variables selected for use in subsequent analysis.

Survey techniques Technniques based upon the use of structured questionnaires given to a sample of a population.

Surveys Interviews with a large number of people using a questionnaire.

Symmetric lambda The symmetric lambda does not make an assumption about which variable is dependent. It measures the overall improvement when prediction is done in both directions.

Syndicated services Companies that collect and sell common pools of data designed to serve information needs shared by a number of clients.

Syndicated sources (services) Information services offered by marketing research organisations that provide information from a common database to different firms that subscribe to their services

Systematic error An error that affects the measurement in a constant way and represents stable factors that affect the observed score in the same way each time the measurement is made.

Systematic sampling A probability sampling technique in which the sample is chosen by selecting a random starting point and then picking every ith element in succession from the sampling frame.

t distribution A symmetrical bell-shaped distribution that is useful for sample testing ($n < 30$). It is similar to the normal distribution in appearance.

t statistic A statistic that assumes that the variable has a symmetric bell-shaped distribution, that the mean is known (or assumed to be known), and that the population variance is estimated from the sample.

t test A univariate hypothesis test using the t distribution, which is used when the standard deviation is unknown and the sample size is small.

Target population The collection of elements or objects that possess the information sought by the researcher and about which inferences are to be made.

tau b A test statistic that measures the association between two ordinal-level variables. It makes an adjustment for ties and is the most appropriate when the table of variables is square.

tau c A test statistic that measures the association between two ordinal-level variables. It makes an adjustment for ties and is most appropriate when the table of variables is not square but a rectangle.

Telescoping A psychological phenomenon that takes place when an individual telescopes or compresses time by remembering an event as occurring more recently than it actually occurred.

Territorial map A tool for assessing discriminant analysis results by plotting the group membership of each case on a graph.

Test market A carefully selected part of the marketplace particularly suitable for test marketing.

Test marketing An application of a controlled experiment done in limited, but carefully selected, test markets. It involves a replication of the planned national marketing programme for a product in test markets.

Test statistic A measure of how close the sample has come to the null hypothesis. It often follows a well-known distribution, such as the normal, t, or chi-square distribution.

Test units Individuals, organisations or other entities whose response to independent variables or treatments is being studied.

Testing effects Effects caused by the process of experimentation.

Test–re-test reliability An approach for assessing reliability, in which respondents are administered identical sets of scale items at two different times, under as nearly equivalent conditions as possible.

Thematic maps Maps that solve marketing problems. They combine geography with demographic information and a company's sales data or other proprietary information and are generated by a computer.

Theoretical sampling Data gathering driven by concepts derived from evolving theory and based on the concept of 'making comparisons'.

Theory A conceptual scheme based on foundational statements, or axioms, that are assumed to be true.

Third-person technique A projective technique in which respondents are presented with a verbal or visual situation and are asked to relate the beliefs and attitudes of a third person in that situation.

Time series design A quasi-experimental design that involves periodic measurements of the dependent variable for a group of test units. Then the treatment is administered by the researcher or occurs naturally. After the treatment, periodic measurements are continued to determine the treatment effect.

Topic guide A list of topics, questions and probes that are used by a moderator to help them manage a focus group discussion.

Total error The variation between the true mean value in the population of the variable of interest and the observed mean value obtained in the marketing research project

Trace analysis An approach in which data collection is based on physical traces, or evidence, of past behaviour.

Transcripts 'Hard copies' of the questions and probes and the corresponding answers and responses in focus group or depth interviews.

Transitivity of preference An assumption made to convert paired comparison data with rank order data. It implies that if brand A is preferred to brand B, and brand B is preferred to brand C, then brand A is preferred to brand C.

Treatment In ANOVA, a particular combination of factor levels or categories.

Trend analysis A method of adjusting for non-response in which the researcher tries to discern a trend between early and late respondents. This trend is projected to non-respondents to estimate their characteristic of interest.

True experimental designs Experimental designs distinguished by the fact that the researcher can randomly assign test units to experimental groups and also randomly assign treatments to experimental groups.

True score model A mathematical model that provides a framework for understanding the accuracy of measurement.

Two-group discriminant analysis Discriminant analysis technique where the criterion variable has two categories.

Two-sample median test Non-parametric test statistic that determines whether two groups are drawn from populations with the same median. This test is not as powerful as the Mann-Whitney *U*.

Two-tailed test A test of the null hypothesis where the alternative hypothesis is not expressed directionally.

Type I error An error that occurs when the sample results lead to the rejection of a null hypothesis that is in fact true. Also known as alpha error (α).

Type II error An error that occurs when the sample results lead to acceptance of a null hypothesis that is in fact false. Also known as beta error (β).

Univariate techniques Statistical techniques appropriate for analysing data when there is a single measurement of each element in the sample, or, if there are several measurements on each element, when each variable is analysed in isolation.

Unstructured observation Observation that involves a researcher monitoring all relevant phenomena, without specifying the details in advance.

Unstructured questions Open-ended questions that respondents answer in their own words.

Validation sample That part of the total sample used to check the results of the estimation sample.

Validity The extent to which a measurement represents characteristics that exist in the phenomenon under investigation.

Variable re-specification The transformation of data to create new variables or the modification of existing variables so that they are more consistent with the objectives of the study.

Variance The mean squared deviation of all the values of the mean.

Variance method An agglomerative method of hierarchical clustering in which clusters are generated to minimise the within-cluster variance.

Varimax procedure An orthogonal method of factor rotation that minimises the number of variables with high loadings on a factor, thereby enhancing the interpretability of the factors.

Verbal models Analytical models that provide a written representation of the relationships between variables.

Verbal protocol A technique used to understand respondents' cognitive responses or thought processes by having them think aloud while completing a task or making a decision.

Viewing laboratory A room where a focus group may be conducted and simultaneously observed, usually by using a two-way mirror.

Voice pitch analysis Measurement of emotional reactions through changes in the respondent's voice.

Volume tracking data Scanner data that provide information on purchases by brand, size, price, and flavour or formulation.

Ward's procedure A variance method in which the squared euclidean distance to the cluster means is minimised.

Weighting A statistical procedure that attempts to account for non-response by assigning differential weights to the data depending on the response rates.

Wilcoxon matched-pairs signed-ranks test A non-parametric test that analyses the differences between the paired observations, taking into account the magnitude of the differences.

Word association A projective technique in which respondents are presented with a list of words, one at a time. After each word, they are asked to give the first word that comes to mind.

z test A univariate hypothesis test using the standard normal distribution.

z value The number of standard errors a point is away from the mean.

SUBJECT INDEX

Figures in **bold** indicate glossary entries

IMPORTANT: READ CAREFULLY

WARNING: BY OPENING THE PACKAGE YOU AGREE TO BE BOUND BY THE TERMS OF THE LICENCE AGREEMENT BELOW.

This is a legally binding agreement between You (the user or purchaser) and Pearson Education Limited. By retaining this licence, any software media or accompanying written materials or carrying out any of the permitted activities You agree to be bound by the terms of the licence agreement below.

If You do not agree to these terms then promptly return the entire publication (this licence and all software, written materials, packaging and any other components received with it) with Your sales receipt to Your supplier for a full refund.

SINGLE USER LICENCE AGREEMENT

❏ YOU ARE PERMITTED TO:

- Use (load into temporary memory or permanent storage) a single copy of the software on only one computer at a time. If this computer is linked to a network then the software may only be installed in a manner such that it is not accessible to other machines on the network.

- Make one copy of the software solely for backup purposes or copy it to a single hard disk, provided you keep the original solely for back up purposes.

- Transfer the software from one computer to another provided that you only use it on one computer at a time.

❏ YOU MAY NOT:

- Rent or lease the software or any part of the publication.

- Copy any part of the documentation, except where specifically indicated otherwise.

- Make copies of the software, other than for backup purposes.

- Reverse engineer, decompile or disassemble the software.

- Use the software on more than one computer at a time.

- Install the software on any networked computer in a way that could allow access to it from more than one machine on the network.

- Use the software in any way not specified above without the prior written consent of Pearson Education Limited.

ONE COPY ONLY

This licence is for a single user copy of the software
PEARSON EDUCATION LIMITED RESERVES THE RIGHT TO TERMINATE THIS LICENCE BY WRITTEN NOTICE AND TO TAKE ACTION TO RECOVER ANY DAMAGES SUFFERED BY PEARSON EDUCATION LIMITED IF YOU BREACH ANY PROVISION OF THIS AGREEMENT.

Pearson Education Limited owns the software You only own the disk on which the software is supplied.

LIMITED WARRANTY
Pearson Education Limited warrants that the diskette or CD rom on which the software is supplied are free from defects in materials and workmanship under normal use for ninety (90) days from the date You receive them. This warranty is limited to You and is not transferable. Pearson Education Limited does not warrant that the functions of the software meet Your requirements or that the media is compatible with any computer system on which it is used or that the operation of the software will be unlimited or error free.

You assume responsibility for selecting the software to achieve Your intended results and for the installation of, the use of and the results obtained from the software. The entire liability of Pearson Education Limited and its suppliers and your only remedy shall be replacement of the components that do not meet this warranty free of charge.

This limited warranty is void if any damage has resulted from accident, abuse, misapplication, service or modification by someone other than Pearson Education Limited. In no event shall Pearson Education Limited or its suppliers be liable for any damages whatsoever arising out of installation of the software, even if advised of the possibility of such damages. Pearson Education Limited will not be liable for any loss or damage of any nature suffered by any party as a result of reliance upon or reproduction of or any errors in the content of the publication.

Pearson Education Limited does not limit its liability for death or personal injury caused by its negligence.

This licence agreement shall be governed by and interpreted and construed in accordance with English law.