Edith D. de Leeuw
Joop J. Hox
Don A. Dillman

International Handbook of Survey Methodology

EUROPEAN ASSOCIATION OF METHODOLOGY

International Handbook of Survey Methodology

EUROPEAN ASSOCIATION OF METHODOLOGY	The European Association of Methodology (EAM) serves to promote research and development of empirical research methods in the fields of the Behavioural, Social, Educational, Health and Economic Sciences as well as in the field of Evaluation Research. Homepage: http://www.eam-online.org

The purpose of the EAM book series is to advance the development and application of methodological and statistical research techniques in social and behavioral research. Each volume in the series presents cutting-edge methodological developments in a way that is accessible to a broad audience. Such books can be authored books, monographs, or edited volumes.

Sponsored by the European Association of Methodology, the EAM book series is open to contributions from the Behavioral, Social, Educational, Health and Economic Sciences. Proposals for volumes in the EAM series should include the following: (1) Title; (2) authors/editors; (3) a brief description of the volume's focus and intended audience; (4) a table of contents; (5) a timeline including planned completion date. Proposals are invited from all interested authors. Feel free to submit a proposal to one of the members of the EAM book series editorial board, by visiting the EAM website http://www.eam-online.org.

Volumes in the series include

De Leeuw/Hox/Dillman: International Handbook of Survey Methodology, 2008

Van Montfort/Oud/Satorra: Longitudinal Models in the Behavioral and Related Sciences, 2007

International Handbook of Survey Methodology

Edith D. de Leeuw
Joop J. Hox
Don A. Dillman

Psychology Press
Taylor & Francis Group
New York London

Psychology Press
Taylor & Francis Group
711 Third Avenue
New York, NY 10017

Published in Great Britain by
Psychology Press
Taylor & Francis Group
27 Church Road
Hove, East Sussex BN3 2FA

Psychology Press is an imprint of Taylor & Francis Group
First published by Lawrence Erlbaum Associates, Inc.

International Standard Book Number-13: 978-0-8058-5753-5 (Softcover) 978-0-858-5752-8 (Hardcover)
Library of Congress catalog number 2007047297

Library of Congress Cataloging-in-Publication Data

Catalog record is available from the Library of Congress

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the Psychology Press Web site at
http://www.psypress.com

Contents

Data Analysis

Special Issues

Preface

In August, 2003, two of us (De Leeuw and Dillman) met in Berlin at the International Statistical Institute meetings to teach a short course on survey design. The audience consisted of surveyors from most continents of the world. Our first impressions were how different the data collection and analysis problems were that people faced, for example doing face-to-face interviews in rural villages of Uganda and web surveys of the general population in The Netherlands and Denmark. Our second, and more lasting impression, was how much all of the participants had in common. Regardless of country, all of the surveyors in the room had to deal with sample designs, writing questions, turning those questions into meaningful questionnaires, locating sample units, processing data, and analyzing the results.

Procedures we originally thought to be of interest only to those from certain countries, such as visual design for mail and web questionnaires, turned out to be of equal interest to those from developing countries who were concerned with more effective visual layouts for interviewer questionnaires and instructions. The idea for this *International Handbook of Survey Methodology* originated from this experience of two fascinating days with this diverse audience with many common needs and interests.

Our experience there was bolstered further by observations of the difficulties being faced in mounting surveys across national borders, and increased concern that they have to be done. For example, expansion of the European Union from 6 countries in 1957 to 15 countries in 1995 (with 9 candidate-members in 2006), has increased interest in collecting cross-national statistical information, including information from sample surveys. We have also observed with much interest emergent efforts to regularly conduct polls and surveys across continents. These surveys aim to facilitate comparisons of responses across countries widely separated in space, as well as technological development, and economic well-being. All this survey effort has resulted in greater concern about how survey methods unique to one country compare to those used in other countries, and how well questionnaire formats and items translate across cultures. It is also difficult to maintain using the same survey mode in all countries.

Within many countries we have noticed the trend towards mixed-mode surveys that is now occurring. Concerns about coverage and nonresponse in telephone surveys, rising costs for conducting face-to-face interviews, and the emergence of web survey capabilities that only some households have, are all encouraging surveyors to mix modes

We are entering a new era in survey design, in which surveyors throughout the world must think about the fundamentals of survey data collection and methods of turning answers to questions into meaningful results. Increasingly it is a mixed-mode world. Whereas at one time it was possible to learn a single survey mode, e.g., face-to-face interviewing or telephone interviewing, and apply it to all survey situations, doing that is no longer possible. It is now imperative for students and practitioners of surveying to learn the procedures associated with multiple modes of collecting sample survey information and apply the method or combination of methods that fit their specific situation.

This handbook provides expert guidance from acknowledged survey methodologists and statisticians around the world, who bring their experiences to bear on issues faced in their own and other countries. It serves as an excellent text for courses and seminars on survey methodology at the masters and graduate level. It is a key reference for survey researchers and practitioners around the world. The book is also very useful for everyone who regularly collects or uses survey data, such as researchers in psychology, sociology, economics, education, epidemiology, and health studies and professionals in market and public opinion research.

The book consists of five parts: foundations, design, implementation, data analysis, and quality issues. The book begins by focusing on the foundations of all sample surveys, ranging from sources of survey error to ethical issues of design and implementation. It is followed by a design section, which gives building blocks for good survey design, from coverage and sampling to writing and testing questions for multiple survey modes. The third section focuses on five modes of data collection, from the oldest, face-to-face interviews, to the newest, interactive voice response, ending with the special challenges involved in mixing these modes within one survey. The fourth section turns to analyzing survey data, dealing with simple as well as complex surveys, and procedures for nonresponse adjustment through imputation and other means. The fifth and final section focuses on special issues of maintaining quality and of documenting the survey process for future reference. The first chapter of the book, *The cornerstones of survey research*, ends with a more detailed description of the structure and contents of this book. There is a companion website http://www.xs4all.nl/~edithl/surveyhandbook.

As we move further into the 21st century, surveys will become inherently more international in scope and in practice. It is our hope that this book will prove helpful for those who are learning the craft of surveying, which like other life skills, will increasingly be applied beyond one's country of origin.

We thank our colleagues across the world for many lively and stimulating discussions about survey methodology. We also thank our students who inspired us and especially the master class in survey methodology 2006 who enthusiastically and critically discussed the drafts. The final book has profited from close reading and copy-editing by Mallory McBride, Sophie van der Zee, Evert-Jan van Doorn, and Amaranta de Haan. We thank Allison O'Neill for her creative cover design. We also thank Emily Wilkinson and Debra Riegert of Lawrence Erlbaum Associates for their patience and careful prodding in getting this book done.

Edith de Leeuw
Joop Hox
Don Dillman

Chapter 1

The Cornerstones of Survey Research

Edith D. de Leeuw
Joop J. Hox
Department of Methodology & Statistics, Utrecht University

Don A. Dillman
Washington State University

1.1 INTRODUCTION

The idea of conducting a survey is deceptively simple. It involves identifying a specific group or category of people and collecting information from some of them in order to gain insight into what the entire group does or thinks; however, undertaking a survey inevitably raises questions that may be difficult to answer. How many people need to be surveyed in order to be able to describe fairly accurately the entire group? How should the people be selected? What questions should be asked and how should they be posed to respondents? In addition, what data collection methods should one consider using, and are some of those methods of collecting data better than others? And, once one has collected the information, how should it be analyzed and reported? Deciding to do a survey means committing oneself to work through a myriad of issues each of which is critical to the ultimate success of the survey.

Yet, each day, throughout the world, thousands of surveys are being undertaken. Some surveys involve years of planning, require arduous efforts to select and interview respondents in their home and take many months to complete and many more months to report results. Other surveys are conducted with seemingly lightning speed as web survey requests are transmitted simultaneously to people regardless of their location, and completed surveys start being returned a few minutes later; data collection is stopped in a few days and results are reported minutes afterwards. Whereas some surveys use only one mode of data collection such as the telephone, others may involve multiple modes, for example, starting with mail, switching to telephone, and finishing up with face-to-face interviews. In addition, some surveys are quite simple and inexpensive to do, such as a mail survey of members of a small professional association. Others are incredibly complex, such as a survey of the general public across all countries of the European Union in which the same questions need to be answered in multiple languages by people of all educational levels.

In the mid-twentieth century there was a remarkable similarity of survey procedures and methods. Most surveys of significance were done by face-to-face interviews in most countries in the world. Self-administered paper surveys, usually done by mail, were the only alternative. Yet, by the 1980s the telephone had replaced face-to-face interviews as the dominate survey mode in the United States, and in the next decade telephone surveys became the major data collection method in many countries. Yet other methods were emerging and in the 1990s two additional modes of surveying—the Internet and responding by telephone to prerecorded interview questions, known as Interactive Voice Response or IVR, emerged in some countries. Nevertheless, in some countries the face-to-face interview remained the reliable and predominantly used survey mode.

Never in the history of surveying have their been so many alternatives for collecting survey data, nor has there been so much heterogeneity in the use of survey methods across countries. Heterogeneity also exists within countries as surveyors attempt to match survey modes to the difficulties associated with finding and obtaining response to particular survey populations.

Yet, all surveys face a common challenge, which is how to produce precise estimates by surveying only a relatively small proportion of the larger population, within the limits of the social, economic and technological environments associated with countries and survey populations in countries. This chapter is about solving these common problems that we described as the cornerstones of surveying. When understood and responded to, the cornerstone challenges will assure precision in the pursuit of one's survey objectives.

1.2 WHAT IS A SURVEY?

A quick review of the literature will reveal many different definitions of what constitutes a survey. Some handbooks on survey methodology immediately describe the major components of surveys and of survey error instead of giving a definition (e.g., Fowler, Gallagher, Stringfellow, Zalavsky Thompson & Cleary, 2002, p. 4; Groves, 1989, p. 1), others provide definitions, ranging from concise definitions (e.g., Czaja & Blair, 2005, p. 3; Groves, Fowler, Couper, Lepkowski, Singer & Tourangeau, 2004, p. 2; Statistics Canada, 2003, p. 1) to elaborate descriptions of criteria (Biemer & Lyberg, 2003, Table 1.1). What have these definitions in common? The survey research methods section of the American Statistical Association provides on its website an introduction (Scheuren, 2004) that explains survey methodology for survey users, covering the major steps in the survey process and explaining the methodological issues. According to Scheuren (2004, p. 9) the word survey is used most often to describe a method of gathering information from a sample of individuals. Besides sample and gathering information, other recurring terms in definitions and descriptions are systematic or organized and quantitative. So, a survey can be seen as a research strategy in which quantitative information is systematically collected from a relatively large sample taken from a population.

Most books stress that survey methodology is a science and that there are scientific criteria for survey quality. As a result, criteria for survey quality have been widely discussed. One very general definition of quality is fitness for use. This definition was coined by Juran and Gryna in their 1980s book on quality planning and analysis, and has been widely quoted since. How this general definition is further specified depends on the product that is being evaluated and the user. For example, quality can be focusing on construction, on making sturdy and safe furniture, and on testing it. Like Ikea, the Swedish furniture chain, that advertised in its catalogs with production quality and gave examples on how a couch was tested on sturdiness. In survey statistics the main focus has been on accuracy, on reducing the mean squared error or MSE. This is based on the Hansen and Hurwitz model (Hansen, Hurwitz, & Madow, 1953; Hansen, Hurwitz, & Bershad, 1961) that differentiates between random error and systematic bias, and offers a concept of total error (see also Kish, 1965), which is still the basis of current survey error models. The statistical quality indicator is thus the MSE: the sum of all squared variable errors and all squared systematic errors. A more modern approach is total quality, which combines both ideas as Biemer and Lyberg (2003) do in their handbook on survey quality. They apply the concept of fitness for use to the survey process, which leads to the following quality requirements for survey data: accuracy as defined by the mean squared error, timeliness as defined by availability at the time it is needed, and accessibility, that is the data should be accessible to those for whom the survey was conducted.

There are many stages in designing a survey and each influences survey quality. Deming (1944) already gave an early warning of the complexity of the task facing the survey designer, when he listed no less than thirteen factors that affect the ultimate usefulness of a survey. Among those are the relatively well understood effects of sampling variability, but also more difficult to measure effects. Deming incorporates effects of the interviewer, method of data collection, nonresponse, questionnaire imperfections, processing errors and errors of interpretation. Other authors (e.g., Kish, 1965, see also Groves, 1989) basically classify threats to survey quality in two main categories, for instance differentiating between errors of nonobservation (e.g., nonresponse) and observation (e.g., in data collection and processing). Biemer and Lyberg (2003) group errors in sampling error and nonsampling error. Sampling error is due to selecting a sample instead of studying the whole population. Nonsampling errors are due to mistakes and/or system deficiencies, and include all errors that can be made during data collection and data processing, such as coverage, nonresponse, measurement, and coding error (see also Lyberg & Biemer, Chapter 22).

In the ensuing chapters of this handbook we provide concrete tools to incorporate quality when designing a survey. The purpose of this chapter is to sensitize the reader to the importance of designing for quality and to introduce the methodological and statistical principles that play a key role in designing sound quality surveys.

A useful metaphor is the design and construction of a house. When building a house, one carefully prepares the ground and places the cornerstones. This is the foundation on which the whole structure must rest. If this foundation is not designed with care, the house will collapse or sink in the unsafe, swampy underground as many Dutch builders have experienced in the past. In the same way, when designing and constructing a survey, one should also lay a well thought-out foundation. In surveys, one starts with preparing the underground by specifying the concepts to be measured. Then these clearly specified concepts have to be translated, or in technical terms, operationalized into measurable variables. Survey methodologists describe this process in terms of avoiding or reducing specification errors. Social scientists use the term construct validity: the extend to which a measurement method accurately represents the intended construct. This first step is conceptual rather than statistical; the concepts of concern must be defined and specified. On this foundation we place the four cornerstones of survey research: coverage, sampling, response, and measurement (Salant & Dillman, 1994; see also Groves, 1989).

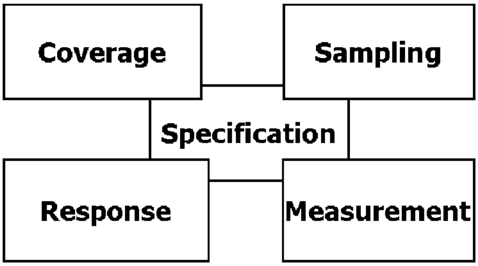

Figure 1.1 The cornerstones of survey research

Figure 1.1 provides a graphical picture of the cornerstone metaphor. Only when these cornerstones are solid, high quality data are collected, which can be used in further processing and analysis. In this chapter we introduce the reader to key issues in survey research.

1.3 BREAKING THE GROUND: SPECIFICATION OF THE RESEARCH AND THE SURVEY QUESTIONS

The first step in the survey process is to determine the research objectives. The researchers have to agree on a well-defined set of research objectives. These are then translated into a set of key research questions. For each research question one or more survey questions are then formulated, depending on the goal of the study. For example, in a general study of the population one or two general questions about well-being are enough to give a global indication of well-being. On the other hand, in a specific study of the influence of social networks on feelings of well-being among the elderly a far more detailed picture of well-being is needed and a series of questions has to be asked, each question measuring a specific aspect of well-being. These different approaches are illustrated in the text boxes noted later.

Example General Well-being Question (Hox, 1986)

Taking all things together, how satisfied or dissatisfied are you with life in general?

- ❑ VERY DISSATISFIED
- ❑ DISSATISFIED
- ❑ NEITHER DISSATISFIED, NOR SATISFIED
- ❑ SATISFIED
- ❑ VERY SATISFIED

Examples General + Specific Well-being Questions (Hox, 1986)

Taking all things together, how satisfied or dissatisfied are you with *life in general*?

- ❑ VERY DISSATISFIED
- ❑ DISSATISFIED
- ❑ NEITHER DISSATISFIED, NOR SATISFIED
- ❑ SATISFIED
- ❑ VERY SATISFIED

Taking all things together, how satisfied or dissatisfied are you with *the home in which you live*?

- ❑ VERY DISSATISFIED
- ❑ DISSATISFIED
- ❑ NEITHER DISSATISFIED, NOR SATISFIED
- ❑ SATISFIED
- ❑ VERY SATISFIED

Taking all things together, how satisfied or dissatisfied are you with *your health*?

Taking all things together, how satisfied or dissatisfied are you with *your social contacts*?

Survey methodologists have given much attention to the problems of formulating the actual questions that go into the survey questionnaire (cf. Fowler & Cosenza, Chapter 8). Problems of question wording, questionnaire flow, question context, and choice of response categories have been the focus of much attention. Much less attention has been directed at clarifying the problems that occur *before* the first survey question is committed to paper: the process that leads from the theoretical construct to the prototype survey item (cf. Hox, 1997). Schwarz (1997) notes that large-scale survey programs often involve a large and heterogeneous group of researchers, where the set of questions finally agreed upon is the result of complex negotiations. As a result, the concepts finally adopted for research are often vaguely defined.

When thinking about the process that leads from theoretical constructs to survey questions, it is useful to distinguish between conceptualization and operationalization. Before questions can be formulated, researchers must decide which concepts they wish to measure. They must define they intend to measure by naming the concept, describing its properties and its scope, and defining important subdomains of its meaning. The subsequent process of operationalization involves choosing empirical indicators for each concept or each subdomain. Theoretical concepts are often referred to as ‘constructs’ to emphasize that they are theoretical concepts that have been invented or adopted for a specific scientific purpose (Kerlinger, 1986). Fowler and Cosenza's (Chapter 8) discussion of the distinction between constructs and survey questions follows these line of reasoning.

To bridge the gap between theory and measurement, two distinct research strategies are advocated: a theory driven or top down strategy, which starts with constructs and works toward observable variables and a data driven or bottom up strategy, which starts with observations and works towards theoretical constructs (cf. Hox & De Jong-Gierveld, 1990). For examples of such strategies we refer to Hox (1997).

When a final survey question as posed to a respondent fails to ask about what is essential for the research question, we have a specification error. In other words, the construct implied in the survey question differs from the intended construct that should be measured. This is also referred to as a measurement that has low construct validity. As a result, the wrong parameter is estimated and the research objective is not met. A clear example of a specification error is given by Biemer and Lyberg (2003, p. 39). The intended concept to be measured was "…the value of a parcel of land if it were sold on a fair market today." A potential operationalization in a survey question would be "For what price would you sell this parcel of land?" Closer inspection of this question reveals that this question asks what the parcel of land is subjectively worth to the farmer. Perhaps it is worth so much to the farmer that she/he would never sell it at all.

There are several ways in which one can investigate whether specification errors occur. First of all, the questionnaire outline and the concept questionnaire should always be thoroughly discussed by the researchers, and with the client or information users, and explicit checks should be made whether the questions in the questionnaire reflect the study objectives. In the next step, the concept questionnaire should be pretested with a small group of real respondents, using so called cognitive lab methods. These are qualitative techniques to investigate whether and when errors occur in the question-answer process. The first step in the question answer process is understanding the question. Therefore, the first thing that is investigated in a pretest is if the respondents understand the question and the words used in the question as intended by the researcher. Usually questions are adapted and/or reformulated, based on the results of questionnaire pretests. For a good description of pretesting, methods, see Campanelli Chapter 10. Whenever a question is reformulated, there is the danger of changing its original (intended) meaning, and thus introducing a new specification error. Therefore, both the results of the pretests and the final adapted questionnaire should again be thoroughly discussed with the client.

1.4 PLACING THE CORNERSTONES: COVERAGE, SAMPLING, NONRESPONSE, AND MEASUREMENT

As noted earlier, specification of the research question and the drafting of prototype survey questions are conceptual rather than statistical; it concerns the

construct validity of the measurement. In other words, does the question measure what it is supposed to measure, does it measure the intended theoretical construct (Cronbach & Meehl, 1955). In contrast, the sources of data collection error summarized in our four cornerstones can be assessed statistically by examining the effect they have on the precision of the estimates.
Three of the four cornerstones refer explicitly to the fact that surveys typically collect data from a sample, a fraction of the population of interest. *Coverage error* occurs when some members of the population have a zero probability of being selected in the survey sample. For example, the sample list (frame) may fail to cover all elements of the population to which one wants to generalize results. *Sampling error* occurs because only a subset of all elements (people) in the population is actually surveyed. Sampling error is statistically well understood provided that probability samples are used: in general the amount of sampling error is a direct function of the number of units included the final sample. For a clear discussion of coverage and sampling, see Lohr (Chapter 6). *Nonresponse error* occurs when some of the sampled units do not respond and when these units differ from those who do and in a way relevant to the study. For an introduction into nonresponse and nonresponse error, see Lynn (Chapter 3). The last cornerstone is *measurement error*, which occurs when a respondent's answer to a question is inaccurate, departs from the "true" value (see also Hox, Chapter 20).

A perfect survey would minimize all four sources of errors. Coverage error is avoided when every member of the population has a known and nonzero chance of being selected into the survey. Sampling error is reduced simply by sampling enough randomly selected units to achieve the precision that is needed. Nonresponse error is avoided if everyone responds or if the respondents are just like the nonrespondents in terms of the things we are trying to measure. Measurement error can be prevented by asking clear questions; questions that respondents are capable and willing to answer correctly. In the survey design stage the methodological goal is to prevent or at least reduce potential errors; in the analysis stage the statistical goal is to adjust the analysis for errors in such a way that correct (i.e., unbiased and precise) results are produced. The methodological survey literature suggests a variety of methods for reducing the sources of survey error; however, one should keep in mind that there is more than one source of error and that one has to compromise and choose when attempting to reduce total survey error. And, do this all within a workable budget too; or as Lyberg and Biemer put it in Chapter 22: "the challenge in survey design is to achieve an optimal balance between survey errors and costs." In the remainder we discuss the four cornerstones in more detail and relate these to specific chapters in this book.

1.4.1 Coverage and Coverage Error

When doing a survey one has an intended population in mind: the target population. To draw a sample from the target population, a sample frame is needed. This can be a list of target population members, for instance, a list of all members of a certain organization, or the register of all inhabitants of a certain city. But it may also be a virtual list, or an algorithm, such as in area probability sampling or in Random Digit Dialing (RDD) sampling (cf. Lohr, Chapter 6 on coverage and sampling, and Steeh, Chapter 12 on RDD). In area probability sampling, the population is divided into clusters based on geographical proximity, and then specific areas are selected. In RDD, random telephone numbers are generated using an algorithm that conforms to properties of valid telephone numbers in the country that is being investigated. Frame coverage errors occur when there is a mismatch between the sampling frame and the target population. In other words when there is no one-to-one correspondence between the units in the frame and the units in the target population.

The most common form of coverage error is undercoverage, that is, not all units of the target population are included in the sampling frame. A clear example of undercoverage is persons with an unlisted phone number when the sampling frame is the telephone book. Another form of coverage error is overcoverage; here a unit from the target population appears more than once in the sampling frame. Duplications like this can occur when a sampling frame results from the combination of several lists. For example, on one list a woman is listed under her maiden name, and on a second list under her married name. If these lists are combined, the same person is listed under two different entries. Another example is surveys that use mobile (cell) telephones; these overcover persons who own more than one phone. A third type of coverage error is caused by erroneous inclusions in the frame. For example, a business number is included on a list with household phone numbers.

As a final example, consider the case of web surveys. A common way to attract respondents to a web survey is placing a link to the survey on a popular web site. Basically, this means that the researcher has no control over who responds to the questionnaire. Coverage error for web surveys is related to two different causes (cf. Ramos, Sevedi, & Sweet, 1998). First, it is the respondent who has to make contact with the data collection program. In a web survey, this requires access to a computer and the Internet, plus some degree of computer skill. Individuals who lack these are not covered. In addition, interviewing software is in general not hardware or software independent. Screens look differently in different resolutions, or when different browsers are used to access the survey website, and some combinations of hardware and software may make the survey website inaccessible to some users, resulting in coverage error. For an overview of different types of web surveys and their potential for errors, see Lozar Manfreda and Vehovar (Chapter 14).

The availability of comprehensive lists or algorithms that cover the population differs widely depending on the target population, but also on the country. For instance, in countries like Denmark and The Netherlands the national statistical agency has access to the population registry (see also Bethlehem Chapter 26). This makes it possible for the national statistical agency to draw a probability sample not only of the general population, but also to draw specific subsamples. Some countries have good lists of mobile phone users, whereas others do not. In some areas, the telephone system has a well-defined structure of used and unused number banks, which makes it possible to generate random telephone numbers with good coverage properties. In most areas, the telephone system does not have such a structure or several competing telephone systems are in use, which makes generating random telephone numbers more difficult (cf. Steeh, Chapter 12).

Web surveys are a special challenge to survey methodologists, because the coverage problem is large and difficult to solve. There are no lists of the population that can be used to draw samples with known properties. Email addresses have no common structure that can be used to generate random addresses similar to the way random telephone numbers are generated in RDD. Finally, the often-used volunteer samples are convenience samples, for which coverage cannot be determined (cf. Lozar Manfreda & Vehovar, Chapter 14).

1.4.2 Sampling and Sampling Error

Sampling error occurs because only a sample of the population is investigated instead of the whole population. Sampling and sampling error is treated by Lohr (Chapter 6). Based on the values for the variables in the *probability* sample, the value for the population is estimated using statistical theory. When simple random sampling is used, standard statistical techniques can be used; however, when more complicated sampling schemes are used, such as cluster sampling or stratification, the standard statistical techniques do not provide accurate *p*-values and confidence intervals and more complicated statistical techniques should be used. Methods for analyzing complex survey designs are discussed by Stapleton in Chapter 18.

Sampling error can be controlled by drawing samples that are large enough to produce the precision wanted. Table 1.1 gives an indication of the number of respondents needed for estimated percentages with a specified precision (e.g., Devore & Peck, 2005, pp. 377–378).

Table 1.1 Precision: Number of respondents needed for percentage estimates within 95 percent Confidence Interval (C.I.).

Number of respondents	Width of 95% C.I.
96	± 10%
384	± 5%
1537	± 2.5%
9604	± 1%

Base percentage 50%, 95% Confidence Interval based on normal approximation

The main point of Table 1.1 is that a large precision requires very large samples. The rule of thumb is that to decrease the sampling errors by half we need a completed sample that is four times as large.

The most important issue about sampling is that if our sample is *not* a probability sample, statistical inference is not appropriate. The difference between probability and nonprobability sampling is that nonprobability sampling does *not* use a *random* selection procedure. This does not necessarily mean that nonprobability samples are unrepresentative of the population; however, it does mean that nonprobability samples cannot depend upon statistical probability theory. With a probabilistic sample, we know the probability that we represent the population well and therefore we can estimate confidence intervals and significance tests. With a nonprobability sample, we may or may not represent the population well, but it is not appropriate to apply statistical inference to generalize to a general population. At best, we can use statistical inference to assess the precision with which we can generalize to a population consisting of whoever responded. Whether this is representative for any general population is beyond statistical inference.

1.4.3 Response and Nonresponse Error

Nonresponse is the inability to obtain data for all sampled units on all questions. There are two types of nonresponse in surveys: *unit nonresponse* and *item nonresponse*. Unit nonresponse is the failure to obtain any information from an eligible sample unit. Unit nonresponse can be the result of noncontact or refusal. Lynn (Chapter 3) provides an extensive overview on nonresponse and nonresponse error; for a discussion of nonresponse error in cross-cultural studies, see Couper and de Leeuw (2003); for statistical adjustment and weighting see Biemer and Christ (Chapter 16). Item-nonresponse or item missing data refers to the failure to obtain information for one or more questions in a survey, given that the other questions are completed. For an introduction see de Leeuw, Hox, and Huisman (2003), for statistical approaches to deal with missing data see Chapter 18 by Rässler, Rubin, and Schenker.

Nonresponse error is a function of the response rate and the differences between respondents and nonrespondents. If nonresponse is the result of a pure chance process, in other words if nonresponse is completely at random, then there is no real problem. Of course, the realized sample is smaller, resulting in larger confidence intervals around estimators. But the conclusions will not be biased due to nonresponse. Only when respondents and nonrespondents do differ from each other on the variables of interest in the study, will there be a serious nonresponse problem. The nonresponse is then *selective* nonresponse and certain groups may be underrepresented. In the worst case, there is a substantial association between the nonresponse and an important variable of the study causing biased results. A classic example comes from mobility studies: people who travel a lot are more difficult to contact for an interview on mobility than people who travel rarely. Thus, selective nonresponse caused by specific noncontacts leads to an underestimate of mobility. For more examples, see Lynn (Chapter 3).

Two main approaches are used to cope with nonresponse: *reducing* and *adjusting*. Nonresponse reduction applies strategies that, in general, reduce the number of noncontacts and refusals. Causes of noncontact depend on the specific survey design. For instance, in face-to-face surveys, noncontact can be the result of the inability of the interviewer to reach the respondent within the allotted number of contact attempts. Increasing the number of contact attempts not only increases the number of contacted and thus the response rate, but also the costs. Varying the days and times at which contact is attempted also increases the response rate, without affecting the cost as much. In mail and Internet surveys, noncontacts can be the result of undeliverable mailings due to errors in the address list. Tools to reduce refusals also depend on the data collection mode used. For instance, interview surveys may use specially trained interviewers to convert refusals, while mail and Internet surveys have to rely on incentives or special contacts to counteract explicit refusals. For more detail, see Lynn (Chapter 3).

Nonresponse adjustment refers to statistical adjustments that are applied after the data are collected. If the difference between the respondents and the nonrespondents is known, for instance because we can compare certain characteristics of the respondents to known population values, statistical weighting can be used to make the sample resemble the population with respect to these characteristics. The problem with statistical adjustment is that usually only simple respondent attributes such as age, sex, and education can be used to weigh the sample. This improves the representativeness of the sample with respect to the variables of central substantive interest only if these variables are related to the attributes used in the weighting scheme. Biemer and Christ discuss weighting for survey data in detail in Chapter 17.

Finally, nonresponse figures should be clearly reported in surveys. This often takes the form of a response rate figure. When reporting response rates it is important to state how the response rate was calculated. For details of response rate calculation and a description of sources of nonresponse, see the brochure on standard definitions of the American Association for Public Opinion Research (AAPOR). A regularly updated version and an online response rate calculator can be found on the AAPOR website (www.aapor.org).

1.4.4 Measurement and Measurement Error

Measurement error is also called error of observation. Measurement errors are associated with the data collection process itself. There are three main sources of measurement error: the questionnaire, the respondent, and the method of data collection. When interviewers are used for data collection, the interviewer is a fourth source of error.

A well-designed and well-tested questionnaire is the basis for reducing measurement error. The questions in the questionnaire must be clear, and all respondents must be able to understand the terms used in the same way. With closed questions, the response categories should be well defined, and exhaustive. When a question is not clear, or when the response categories are not clearly defined, respondents will make errors while answering the question or they do not know what to answer. When the data are collected through interviews, interviewers will then try to help out, but in doing this they can make errors too and introduce additional interviewer error (Fowler, 1995). Therefore, improving the *questionnaire* is a good start to improve the total survey quality. For a good introduction into designing and writing effective questions, see Fowler and Cosenza (Chapter 8). It should be emphasized that even carefully designed questionnaires may contain errors and that a questionnaire should always be evaluated and pretested before it may be used in a survey. In Chapter 10 Campanelli provides the reader with information about the different methods for testing survey questions and gives practical guidelines on the implementation of each of the methods.

Respondents can be a source of error in their own right when they provide incorrect information. This may be unintentional, for instance when a respondent does not understand the question or when a respondent has difficulty remembering an event. But a respondent can also give incorrect information on purpose, for instance when sensitive questions are asked (see also Lensvelt-Mulders, Chapter 24). Measurement errors that originate from the respondent are beyond the control of the researcher. A researcher can only try to minimize respondent errors by making the respondent's task as easy and as pleasant as possible. In other words, by writing clear questions that respondents are willing to answer. In Chapter 2, Schwarz, Knäuper, Oyserman, and Stich describe how respondents come up with an answer and review the cognitive and communicative processes underlying survey responses.

The *method of data collection* can be a third source of measurement error. In Chapter 7 of this book, de Leeuw describes the advantages and disadvantages of major data collection techniques. One of the key differences between survey modes is the way in which certain questions can be asked. For instance, in a telephone interview respondents have to rely on auditive cues only: they only hear the question and the response categories. This may cause problems when a long list of potential answers has to be presented. Dillman, in Chapter 9 on the logic and psychology of questionnaire design, describes mode differences in questionnaire design and proposes a unified or uni mode design to overcome differences between modes. This is of major importance when mixed-mode designs are used, either within one survey, or in longitudinal studies (e.g., panel surveys see also Chapter 25 by Sikkel & Hoogendoorn), or between surveys as can be the case in cross-national and comparative studies in which one mode (e.g., telephone) is used in one country an another mode (e.g., face-to-face interviews) is used in another. For important issues in comparative survey research, see Harkness (Chapter 4); for more detail on the challenges of mixed mode surveys, see De Leeuw, Dillman, and Hox (Chapter 16).

A second major difference between modes is the presence versus the absence of an interviewer. There may be very good reasons to choose a method without interviewers and leave the locus of control with the respondents, such as ensuring more privacy and more time to reflect for respondents. Self-administered questionnaires in general are described by De Leeuw and Hox in Chapter 13; technological innovations are described by Lozar Manfreda and Vehovar in Chapter 14 on Internet Surveys and by Miller Steiger and Conroy in Chapter 15 on Interactive Voice Response. On the other hand, using interviewers also has many positive points, especially when very complex questionnaires are used or when special tasks have to be performed. As Loosveldt states in Chapter 11: "…the task of the interviewer is more comprehensive and complex than merely asking questions and recording the respondent's answer. Interviewers implement the contact procedure, persuade the respondents to participate, clarify the respondent's role during the interview and collect information about the respondent."

However, when an interviewer is present, the interviewer can be a source of error too. Interviewers may misinterpret a question, may make errors in administering a questionnaire, or in registering the answers. When posing the question, interviewers may unintentionally change its meaning. By giving additional information or explaining a misunderstood word, they may inappropriately influence a respondent. Even the way interviewers look and dress may influence a respondent in a face-to-face interview. Selecting and training interviewers carefully helps reducing interviewer related errors. For more details, see Chapter 23 on interviewer training by Lessler, Eyerman, and Wang. Interviewers can make genuine mistakes, but they also may intentionally cheat. Interviewers have been known to falsify data, or skip questions to shorten tedious interviews. Monitoring interviewers helps to reduce this. Having a quality controller listening in on telephone interviewers is a widely used method. In face-to-face interviews, recordings can be made and selected tapes can be checked afterwards. Special verification contacts or re-interviews may be used to evaluate interviewer performance in large-scale face-to-face surveys (cf. Lyberg & Biemer, Chapter 22; Japec, 2005, p. 24).

1.5 FROM DATA COLLECTION TO ANALYSIS: HOW THE FOUNDATION AFFECTS THE STRUCTURE

There are several ways in which the design of a survey and the precise data collection procedure affects the subsequent data analysis stage. These also involve the four cornerstones. The most direct influence is the actual *sampling* procedure that is used. As mentioned earlier, standard statistical procedures assume that the data are a simple random sample from the population. In most surveys, other sampling schemes are used because these are more efficient or less expensive, for instance cluster sampling or stratification. When these sampling schemes are used, the analysis must employ special statistical methods (see also Stapleton, Chapter 18). Similarly, when weighting (cf. Biemer & Christ, Chapter 17) is used to compensate for different inclusion probabilities, either by design or because of nonresponse problems, special statistical methods must be used. Standard statistical packages may or may not include these methods. For instance, the package SPSS (version 15 and higher) can analyze complex survey data with weights and complicated sampling schemes, but it includes only selected statistical analyses for such data. The other procedures in SPSS can include weighting, but do not correct the standard errors for the effects of weighting, which produces incorrect statistical tests.

A less obvious way in which the survey design affects the data analysis lies in the adjustment for the combination of coverage error and nonresponse. These may result in data that are not representative for the population, and the most often-used adjustment method is weighting on respondent characteristics for which the population values are known. For more detail, see Biemer and Christ (Chapter 17). Two issues are important here. First, statistical adjustment aims at producing unbiased estimates of population parameters when selection probabilities are not equal; however, no amount of statistical cleverness restores information that we have failed to collect. So, prevention by reducing the problem in the data collection phase is important. Second, the quality of the adjustment depends strongly on the amount and quality of background information that we have available to construct the weights. Collecting this information requires careful planning in the design phase. Auxiliary variables must be included for which the population values are known, for instance for a sample from the general population via the national statistical agency, or for samples from a special population via an existing registry. Because the use of registries is regulated by privacy concerns, in the latter case it may be necessary to obtain prior permission. For more on privacy and ethics in survey research, see Singer (Chapter 5). Finally, to be able to use the information, it is crucial that the data collection procedure uses the same wording and response categories that were used to collect the known population data (cf. Dillman, Chapter 9). Preferably, the same method of data collection should be used, to prevent confounding of selection and measurement errors.

A special case of nonresponse is the failure to obtain information on some of the questions, which leads to incomplete data for some of the respondents. Just as is the case with unit-nonresponse discussed earlier, prevention and the collection of auxiliary information is important with item missing data too (see also de Leeuw, Hox, & Huisman, 2003). The next step is statistical adjustment. In Chapter 19, Rässler, Rubin, and Schenker discuss concepts regarding mechanisms that create missing data, as well as four commonly used approaches to deal with (item) missing data.

Measurement errors, that is discrepancies between the measurement and the true value, influence the analysis in more subtle ways. Again, prevention is the best medicine. Measurement errors originate from the question wording and the questionnaire, from the survey method and the interviewer, from the respondents and from complex interactions between these. Many decisions in the survey design phase have the potential to affect measurement error (cf. Biemer & Lyberg, Chapter 22). Prevention rest on the application of known best practices in survey design; this assumes that these are well documented (cf. Mohler, Pennel, & Hubbard, Chapter 21). Another important step in reducing measurement error as far as possible is thorough pretesting of the survey instrument before it is actually used (cf. Campanelli, Chapter 10). In the analysis phase, some adjustments for the effect of measurement errors can be made; Hox discusses this in Chapter 20. Adjustments for measurement errors can be made when multi-item scales are used, or if auxiliary information is available about the amount of measurement error in specific variables. Again, to be able to adjust in the analysis phase, the design of the survey must make sure that the necessary information is available.

1.6 CAN WE AFFORD IT? BALANCING DESIGN FEATURES AND SURVEY QUALITY

Earlier we discussed the foundation of survey research: breaking the ground (specification) and placing the four cornerstones (coverage, sampling, nonresponse, and measurement). The same fundamental quality criteria are discussed in quality handbooks. For instance, in Eurostat's 2000 publication on the assessment of quality in statistics, the first quality criterion is the relevance of the statistical concept. A statistical product is relevant if it meets user's needs. This implies that user's needs must be established at the start. The concept of relevance is closely related to the specification problem and the construct validity of measurement. Did we correctly translate the substantive research question into a survey question? If not, we have made a specification error, and the statistical product does not meet the needs of the users. Almost all handbooks on survey statistics mention *accuracy* of the estimate as quality criterion. Accuracy depends on all four cornerstones and is discussed at length earlier in this chapter. But, there are additional criteria for quality as well. Biemer and Lyberg (2003) stress the importance of timeliness defined as available at the time it is needed, and accessibility, that is the data should be accessible to those for whom the survey was conducted. Eurostat (2000) distinguishes seven distinct dimensions of statistical quality, adding a.o. comparability, meaning that it should be possible to make reliable comparisons across time and across space. Comparability is extremely important in cross-cultural and cross-national studies (see also Harkness, Chapter 4). For a discussion of quality and procedures for quality assurance and quality control, see Lyberg and Biemer (Chapter 22).

Both Biemer and Lyberg's (2003) quality concepts and Eurostat's (2000) dimensions go beyond the foundation and cornerstones described earlier in this chapter, and are relevant for the quality of the entire survey process and the data it produces. Their criteria were developed mainly for use in large scale survey organizations and governmental statistical offices, but survey quality and quality assurance is an issue that also applies to smaller scale surveys, where the survey researcher is also the survey user. It does not matter if it is a small scale survey or a large survey, whether the survey is using paper and pencil or high technology, quality can and should be built into all surveys. For procedures for quality assessment, see Lyberg and Biemer (Chapter 22).

To come back to the metaphor of building a house: there are many different ways to build a good, quality house. But, there is also a large variety in types of houses, ranging from a simple summer cottage to a luxurious villa, from a houseboat to a monumental 17th century house at a canal, from a working farm to a dream palace. What is a good house depends on the needs of the resident, what is a good survey depends on the survey user (cf. Dippo, 1997). The research objectives determine the population under study and the types of questions that should be asked. Privacy regulations and ethics may restrict the design; other practical restriction may be caused by available time and funds. Countries and survey organizations may differ in available resources, such as skilled labor, administrative capacities, experience with certain procedures or methods, computer hardware and software. It is clear that survey methodologists must balance survey costs and available resources against survey errors, and that any actual survey will be the result of methodological compromises. Surveys are a complex enterprise and many aspects must be considered when the goal is to maximize data quality with the available resources and within a reasonable budget of time and costs.

Finally, surveys are carried out in a specific cultural context, which may also affect the way these aspects influence the survey quality. Survey methodologists need to take this into account when designing a survey. For instance, when a telephone (or Internet) survey is contemplated for an international study, it is important to understand how telephones and Internet are viewed in the different cultures included in the survey. Is it a personal device, such as mobile telephones? Is it a household device, as landline telephones mostly are? Or is it a community device, with one (mobile) telephone or Internet connection shared by an entire village? Survey design means that costs and quality must be optimized, and in a global world this means that they must be optimized within the bounds of cultural and technological resources and differences.

1.7 CONTENTS OF THIS BOOK

The goal of this book is to introduce the readers to the central issues that are important for survey quality, to discuss the decisions that must be made in designing and carrying out a survey, and to present the current methodological and statistical knowledge about the consequences of these decisions for the survey data quality.

The first section of the book, *Foundations*, is a broad introduction in survey methodology. In addition to this introduction, it contains chapters on the psychology of asking questions, the problem of nonresponse, issues and challenges in international surveys, and ethical issues in surveys.

The second section, *Design*, presents a number of issues that are vital in designing a quality survey. It includes chapters on coverage and sampling, choosing the method of data collection, writing effective questions, constructing the questionnaire, and testing survey questions.

The third major section, *Implementation*, discusses the details of a number of procedures to carry out a survey. There are chapters on face-to-face interviews, telephone interviews, self-administered questionnaires, Internet surveys and Interactive Voice Response surveys. Finally, there is a chapter on the challenges that result when different data collection modes are mixed within a survey.

The fourth section, *Data analysis*, discusses a number of statistical subjects that are especially important in analyzing survey data. These include chapters on constructing adjustment weights, analyzing data from complex surveys, coping with incomplete data (item nonresponse), and accommodating measurement errors. The final section, *Special issues*, contains a number of special interest topics for quality surveys. It includes chapters on survey documentation, quality assurance and quality control, interviewer training, collecting data on sensitive topics, and panel surveys including access panels. The final chapter introduces collecting survey-type data without asking questions of respondents, by combining and integrating existing information.

GLOSSARY OF KEY CONCEPTS

Construct validity. The extend to which a measurement instrument measures the intended construct and produces an observation distinct from that produced by a measure of a different construct.

Coverage error. Coverage errors occur when the operational definition of the population includes an omission, duplication, or wrongful inclusion of an element in the population. Omissions lead to undercoverage, and duplications and wrongful inclusions lead to overcoverage.

Measurement error. The extent to which there are discrepancies between a measurement and the true value, that the measurement instrument is designed to measure. Measurement error refers to both variance and bias, where variance is random variation of a measurement and bias is systematic error. There are a number of potential sources; for example, measurement error can arise from the respondent, questionnaire, mode of data collection, interviewer, and interactions between these.

Nonresponse error. Nonresponse is the failure to collect information from sampled respondents. There are two types of nonresponse: unit nonresponse and item nonresponse. Unit nonresponse occurs when the survey fails to obtain any data from a unit in the selected sample. Item nonresponse (incomplete data) occurs when the unit participates but data on particular items are missing. Nonresponse leads to nonresponse error if the respondents differ from the nonrespondents on the variables of interest.

Sampling error. Error in estimation due to taking a sample instead of measuring every unit in the sampling frame. If probability sampling is used then the amount of sampling error can be estimated from the sample.

Specification error. Specification error occurs when the concept measured by a survey question and the concept that should be measured with that question differ. When this occurs, there is low construct validity.

References

AAPOR (2007). The problem with reporting margins of error and sampling error in online and other survey sof self-selected individuals. Retrieved June, 2007 , from www.aapor.org.

AAPOR . (2005). Standard definitions: Final dispositions of case codes and outcome rates for surveys, online edition 3.1. Retrieved November 2006 , from: http://www.aapor.org/pdfs/standarddefs_3.1.pdf.

Adorno, T. W. (1950). Some aspects of religious ideology as revealed in the interview material. In T. W. Adorno , E. Frenkel-Brunswick , D. Levinson , & R. N. Sanford (Eds.), The Authoritarian Personality. New York: Harper.

Ahart, A. M. , & Sacket, P. R. (2004). A new method of examining relationships between individual difference measures and sensitive behavior criteria: Evaluating the Unmatched Count Technique. Organisational Research Methods, 7, 101–114.

Alemagno, S.A. , Cochran, D. , Feucht, T.E. , Stephens, R.C. , & Wolfe, S.A. (1996). Assessing substance abuse treatment needs among the homeless: A telephone-based interactive voice response system. American Journal of Public Health, 86, 1626–1628.

American Psychological Association . (2003). Psychological research online: Opportunities and challenges (Working Paper Version 3/31/03). Washington, DC.: American Psychological Association.

Andary, L. , Stolk, Y. , & Klimidid, S. (2003). Assessing mental health across cultures. Bowen Hills, Australia: Australian Academic Press.

Anderson, J. R. (2000). Learning and memory: An integrated approach. New York: Wiley.

Anderson, J.R. (1995). Cognitive psychology and its implications. New York: Freeman.

Anderson, M. , & Seltzer, W. (2004, March). The challenges of taxation, investigation, and regulation: Statistical confidentiality and U.S. federal statistics, 1910–1965. Paper prepared for Census Bureau Symposium, America's Scorecard: The Historic Role of the Census in an Ever-Changing Nation, Woodrow Wilson International Center for Scholars, Washington, DC.

Anderson, R. T. , Aaronson, N. K. , & Wilkin, D. (1993). Critical review of the international assessment of health-related quality of life. Quality of Life Research, 2, 369–395.

Andrews, F. M. (1984). Construct validity and error components of survey measures: A structural modeling approach. Public Opinion Quarterly, 48, 409–422.

Angus, V. C. , Entwistle, V. A. , Emslie, M. J. , Walker, K. A. , & Andrew, J. E. (2003). The requirement for prior consent to participate on survey response rates. BMC Health Services Research, 3, 21–30.

Aoki, K. , & Elasmar, M. (2000, May). Opportunities and challenges of a web survey: A field experiment. Paper presented at the 55th Annual Conference of American Association for Public Opinion Research, Portland, OR.

Argyle, M. (1973). Social interaction. London: Tavistock

Argyle, M. (1994). The psychology of social class. London: Routledge.

Armacost R. L. , Hosseini J. C. , Morris S. A. , & Rehbein K. A. (1991). An empirical comparison of direct questioning, scenario, and randomized response methods. Decision Sciences, 22, 1037–1060.

Armer, M. J. , & Grimshaw, A. D. (1973). Methodological problems and possibilities in comparative research. In M. J. Armer & A. D. Grimshaw (Eds.), Comparative social research: Methodological problems and strategies. New York: Wiley.

Asch, S. E. (1951). The effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), Groups, leadership, and men. Pittsburgh, PA: Carnegie.

Bailar, B. A. (1989). Information needs, surveys, and measurement error. In D. Kasprzyk , G. Duncan , G. Kalton , & M. P. Singh (Eds.), Panel Surveys. New York: John Wiley and Sons.

Bailar, B. A. , & Dalenius, T. (1969). Estimating the response variance components of the U.S. Bureau of the Census' survey model. Sankhya, B, 341–360.

Baker, R. (1999). Codes of ethics: Some history. Perspectives on the Profession, 19.

Balden, W. (2004, March). Multi-mode data collection: Benefits and downsides. Paper presented at the 2004 conference of the Great Lakes chapter of the Marketing Research Association, Cancun, Mexico. Retrieved April 15, 2004 , from: http://glcmra.org/cancun. (full paper upon request from the author).

Baldrige National Quality Program (2006). Retrieved from: http://www.quality.nist.gov/
Bargmeyer, B. , & Gillman, D. (2003). Metadata Standards and Metadata Registries: An overview. Retrieved April 16, 2003 , from: http://www.bls.gov/ore/pdf/st000010.pdf
Barnard, J. , & Rubin, D. B. (1999). Small-sample degrees of freedom with multiple imputation. Biometrika, 86, 948–955.
Batagelj, Z. , & Vehovar, V. (1998). WWW Surveys. In A. Ferligoj (Ed.), Advances in methodology, data analysis, and statistics. Ljubljana: Faculty of Social Sciences.
Bauman, S. , Jobity, N. , Airey, J. & Atak, H. (2000). Invites, intros and incentives: Lessons from a web survey. Paper presented at the 55th annual conference of American Association for Public Opinion Research, Portland, OR.
Baumgartner, H. , & Steenkamp, J. E. M. (2001). Response styles in marketing research: A cross-national investigation. Journal of Marketing Research, 38, 143–156.
Beatty, P. (2004). The dynamics of cognitive interviewing. In S. Presser , J. M. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , et al. (Eds.), Methods for testing and evaluating survey questionnaires. New York: Wiley.
Beebe, T. J. , Davern, M. E. , McAlpine, D. D. , Call, K. T. , & Rockwood, T. H. (2005). Increasing response rates in a survey of Medicaid enrolees: The effect of a prepaid monetary incentive and mixed modes (mail and telephone). Medical Care, 43, 411–414.
Begeer, W. , De Vries, W. F. M. , & Dukker, H. D. (1986). Statistics and administration. Netherlands Official Statistics, 1, 7–17.
Belli, R. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. Memory, 6, 383–406.
Belli, R. , Schwarz, N. , Singer, E. , & Talarico, J. (2000). Decomposition can harm the accuracy of retrospective behavioral reports. Applied Cognitive Psychology, 14, 295–308.
Belson, W. A. (1981). The design and understanding of survey questions. Aldershot: Gower.
Bennett, M. , & Trussel, N. (2001). Return postage type as cooperation stimulant for a mail mode portion of a multi-mode survey. Paper, 56th Annual AAPOR conference, Montreal, Canada.
Berlim, M. T. , & Fleck, M. P. A. (2003). ‘Quality of Life’: a brand new concept for research and practice in psychiatry. Revista Brasileira de Psiquiatria [Brazilian Journal of Psychiatry], 25, 249–252. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516444620030004000013&lng=en&nrm=iso, September 2006.
Berry, J. (1969). On cross-cultural comparability. International Journal of Psychology, 4, 119–128.
Berscheid, E. , Baron, R. , Dermer, M. , & Libman, M. (1973). Anticipating informed consent: An empirical approach. American Psychologist, 28, 913–925.
Best, S. , & Krueger, B. (2004). Internet data collection. Thousand Oaks, CA: Sage.
Bethlehem, J. G. (2002). Weighting nonresponse adjustments based on auxiliary information. In R. M. Groves , D. A. Dillman , J. L. Eltinge , & R. J. A. Little (Eds.), Survey nonresponse. New York: Wiley.
Bethlehem, J. G. , Keller, W. J. , & Pannekoek, J. (1990). Disclosure control of microdata. Journal of the American Statistical Association, 85, 38–45.
Biemer, P. P. , & Caspar, R. A. (1994). Continuous quality improvement for survey operations: Some general principles and applications. Journal of official statistics, 10, 307–326.
Biemer, P. P. , & Lyberg, L. E. (2003). Introduction to survey quality. New York: Wiley.
Biemer, P. P. , & Wiesen . (2002). Latent class analysis of embedded repeated measurements: An application to the national household survey on drug abuse. Journal of the Royal Statistical Society series A, 165, 97–120.
Billiet, J. (2003). Cross-cultural equivalence with Structural Equation Modeling. In J. A. Harkness , F. J. R. Van De Vijver , & P. Mohler (Eds.), Cross-cultural survey methods. Hoboken, NJ: Wiley.
Billiet, J. , & Loosveldt, G. (1988). Improvement of the quality of responses to factual survey questions by interviewer training. Public Opinion Quarterly, 52, 190–211.

Blair, E. , & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. Journal of Consumer Research, 14, 280–288.
Blazeby, J. , Sprangers, M. , Cull, A. , Groenvold, M. , & Bottomley, A. (2002). Guidelines for developing questionnaire modules. Retrieved August 28, 2006 from: http://www.eortc.be/home/qualityoflife/Manuals/Moduledevelopment%20guidelines-2002.pdf
Bless, H. , Igou, E. R. , Schwarz, N. , & Wänke, M. (2000). Reducing context effects by adding context information: The direction and size of context effects in political judgment. Personality and Social Psychology Bulletin, 26, 1036–1045.
Bless, H. , Schwarz, N. , & Wänke, M. (2003). The size of context effects in social judgment. In J. P. Forgas , K. D. Williams , & W. von Hippel (Eds.), Social judgments: Implicit and explicit processes. Cambridge, MA: Cambridge University Press.
Blohm, M. , & Koch, A. (2004, August). First contacts by phone or in person: Some evidence from ALLBUS 2000. Paper presented at the 15th Workshop on Household Survey Nonresponse, Maastricht, Netherlands.
Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.
Bolton, P. (2001). Cross-cultural validity and reliability testing of a standard psychiatric assessment instrument without a Gold Standard. Journal of Nervous & Mental Disease, 189, 238–242.
Borgers, N. , de Leeuw, E. , & Hox, J. (2000). Children as respondents in survey research: Cognitive development and response quality. Bulletin de Méthodologie Sociologique, 66, 60–75.
Boruch, R. F. (1971). Assuring confidentiality of responses in social research: A note on strategies. The American Sociologist, 6, 308–311.
Botman, S. L. , Moore, T. F. , Moriarity, C. L. , & Parsons, V. L. (2000). Design and estimation for the National Health Interview Survey, 1995–2004. Vital and Health Statistics, 2, 130.
Brackstone, G. (1999). Managing data quality in a statistical agency. Survey Methodology, 25, 139–149.
Bradburn, N. M. , Sudman, S. , Blair, E. , & Locander, W. (Eds.). (1979). Improving interview method and questionnaire. San Francisco: Jossey Bass.
Bradburn, N. M. , Sudman, S. , & Wansink, B. (2004). Asking questions (2nd ed.). San Francisco: Jossey Bass.
Brambilla, D. J. , & McKinlay, S. M. (1987). A comparison of responses to mailed questionnaires and telephone interviews in a mixed mode health survey. American Journal of Epidemiology, 126, 962–971.
Braun, M. , & Harkness, J. A. (2005). Text and context: Challenges to comparability in survey questions. In J. H. P. Hoffmeyer-Zlotnik , & J. A. Harkness (Eds.), ZUMA-Nachrichten spezial No. 11. Methodological aspects in cross-national research. Mannheim, BRD: ZUMA.
Bremer, J. , Terhanian, G. , & Strange, P. (2004, August). Propensity score matching as a bias correction method for internet-based studies. Paper presented at the Joint Statistical Meeting, Toronto, Canada.
Brenner, M. (1981). Patterns of social structure in the research interview. In Brenner (Ed.), Social method and social life. London: Academic Press.
Bretschneider, M. , & Schumacher, J. (1996). DEMOS eine Datenbank zum Nachweis Kommunaler Umfragen auf dem Weg zum Analyse instrument. ZA-Information, 38, 59–81.
Brewer, K. R. W. (1981). Estimating marijuana usage using randomized response: Some paradoxical findings. Australian Journal of Statistics, 23, 139–148.
Brick, M. J. , Waksberg, J. , Kulp, D. , & Starer, A. (1995). Bias in list-assisted telephone samples. Public Opinion Quarterly, 59, 218–235.
Brogan, D. (1998). Software for sample survey data, misuse of standard packages. In P. Armitage & T. Colton (Eds.), Encyclopedia of biostatistics, 5 (pp. 4167–4174). New York: Wiley.
Bruce, R. , & Anderson, W. (1967). On the comparability of meaningful stimuli in cross-cultural research. Sociometry, 30, 124–136.
Buckingham, A. , & Saunders, P. (2004). The survey methods workbook. Cambridge: Polity Press.
Bullinger, M. (2004). The challenge of cross-cultural quality of life assessment. Paper for the international workshop on researching well-being in developing countries. Retrieved June 20,

2006 , from: http://www.welldev.org.uk/news/hanse-pdfs/bullinger-hanse-full.pdf.
Bushery, J. , Reichert, J. W. , & Blass, R. F. (2003). U.S. census 2010 quality assurance strategy. ASA Proceedings of the Section of Government Statistics.
Callegaro, M. , Buskirk, T. , Piekarski, L. , Kuusela, V. , Vehovar, V. , & Steeh, C. (2004, August). Calculating outcome rates for mobile phone surveys: A proposal for a modified AAPOR standard and its application to three case studies. Paper presented at the RC33 Sixth Annual Conference on Social Science Methodology, Amsterdam, The Netherlands.
Campanelli, P. C. (1997). Testing survey questions: New directions in cognitive interviewing. Bulletin de Méthodologie Sociologique, 55, 5–17.
Campanelli, P. C. , Martin, E. A. , & Rothgeb, J. M. (1991). The use of respondent and interviewer debriefing studies as a way to study response error in survey data. The Statistician, 40, 253–264.
Campanelli, P. C. , Rothgeb, J. , Esposito, J. , & Polivka, A. (1991, May). Methodologies for evaluating survey questions: An illustration from a CPS CATI/RDD test. Paper presented at the annual meeting of the American Association for Public Opinion Research, Phoenix, AZ.
Cannell, C. F. , Fisher, G. , & Bakker, T. (1965). Reporting on hospitalization in the Health Interview Survey (PHS Publication No. 1000, Series 2, No. 6). Washington, D.C.: US Government Printing Office.
Cannell, C. F. , Fowler F. J. , & Marquis, K. H. (1968). The influence of interviewer and respondent psychological and behavioural variables on the reporting in household interviews. Vital and Health Statistics, series 14, 26.
Cannell, C. F. , & Kahn, R. (1968). Interviewing. In G. Lindzey , & E. Aronson (Eds.), The handbook of social psychology, Vol 2. Reading, MA: Addison-Wesley.
Cannell, C. F. , Marquis, K. , & Laurent, A. (1977). A summary of studies. Vital & Health Statistics, Series 2, 69. Washington, DC: Government Printing Office.
Cannell, C. F. , Miller, P. V. , & Oksenberg, L. (1981). Research on interviewing techniques. In S. Leinhardt (Ed.), Sociological methodology 1981 (389–437). San Francisco: Jossey-Bass.
Cannell, C. F. , Oksenberg, L. , Kalton, G. , Bischoping, K. , & Fowler, F. J. (1989). New techniques for pretesting survey questions (Final Report). Bethesda, MD: National Information Center for Health Services Research and Health Care Technology Assessment (NICHSR).
Cannell, C.F. , Oksenberg, L. , & Converse, J.M. (1977). Strivin for response accuracy: experiments in new interviewing techniques. Journal of Marketing Research, 14, 306–315.
Cantor, D. , Allen, B. , Schneider, S. , Hagerty-Heller, T. , & Yuan, A. (2004). Testing an automated refusal avoidance training methodology. ASA Proceedings of the Annual Meeting of the American Statistical Association. Arlington, VA: ASA.
Casady, R. J. , & Lepkowski J. M. (1991). Optimal allocation for stratified telephone survey design. ASA Proceedings of the section on survey research methods, 111–116.
Casady, R. J. , & Lepkowski, J. M. (1993). Stratified telephone survey designs. Survey Methodology, 19, 103–113.
Caspar, R. (2004). Material prepared by R. Caspar for Joint Program in Survey Methods (JPSM) course 2004 on “Questionnaire Design” conducted by Campanelli and Caspar. College Park, MD: Joint Programme in Survey Methodology, University of Maryland, University of Michigan, Westat.
CBS . (2005). Enquête-onderzoek onder allochtonen. Problemen en oplossingen. [Surveying alochtonous respondets; problems and solutions] Heerlen/Voorburg, the Netherlands: Centraal Bureau voor de Statistiek.
Chakrabarty, R. P. , & Torres, G. (1996). American housing survey: A quality profile. Washington, DC: U.S. Department of Housing and Urban Development and U.S. Department of Commerce.
Chambers, R. L. , & Skinner, C. J. (Eds.) (2003). Analysis of survey data. Chichester, UK: Wiley.
Chen, C. , Lee, S. , & Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. Psychological Science, 6, 170–175.

Cheng, A. T. A. (2001). Case definition and culture: Are people all the same? The British Journal of Psychiatry, 179, 1–3.
Cheung, F. M. (2004). Use of western and indigenously developed personality tests in Asia. Applied Psychology: An International Review, 53, 173–191.
Christian, L. M. , & Dillman, D. A. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questionnaires. Public Opinion Quarterly, 68, 57–80.
Christian, L. M. , Dillman, D. A. , and Smyth, J. D. (2007). The effects of aural vs. visual communication on answers to scalar questions in mixed-mode surveys. In James M Lepkowski , Clyde Tucker , J. Michael Brick , Edith D. de Leeuw , Lilli Japec , Paul J. Lavrakas , Michael W. Link , Roberta L. Sangster . Advances in Telephone Survey Methodology. New York: Wiley.
Christian, L. M. , Dillman, D. A. , & Smyth, J. D. (2005). Instructing web and telephone respondents to report date answers in format desired by the surveyor (Social and Economic Sciences Research Center Technical Report 05–067). Washington, DC: Washington State University.
Chrostowski, S. J. , & Malak, B. (2003). Translation and cultural adaptation of the TIMSS 2003 Instruments. In M. O. Martin , I. V. S. Mullis , & S. J. Chrostowski (Eds.), TIMSS 2003 Technical Report. Chestnut Hill, MA: TIMSS & PIRLS International Study Center. Retrieved July 20, 2006 from: http://timss.bc.edu/PDF/t03_download/T03_TR_Front.pdf.
Church, A. H. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. Public Opinion Quarterly, 57, 62–79.
Church, A. H. (2001). Is there a method to our madness? The impact of data collection methodology on organizational survey results. Personnel Psychology, 54, 937–969.
Citro, C. F. , Cork, D. L. , & Norwood, J. L. (Eds.). (2004). The 2000 census: Counting under adversity. Washington, D.C.: National Academies Press.
Clark, H. H. , & Clark, E. V. (1977). Psychology and language. New York: Harcourt, Brace, Jovanovich.
Clark, H. H. , & Schober, M. F. (1992). Asking questions and influencing answers. In J. M. Tanur (Ed.), Questions about questions: Inquiries into the cognitive bases of surveys. New York: Russell Sage Foundation.
Coates, D. (2004, August). On-line surveys: Does one size fit all? Paper presented at the RC-33 6th International Conference on Social Methodology, Amsterdam, The Netherlands.
Cochran, W. G. (1977). Sampling techniques (3rd ed). New York: Wiley.
Coleman, M. P. , Evans, B. G. , & Barrett, G. (2005). Confidentiality and the public interest in medical research: Will we ever get it right? Clinical Medicine, 3, 219–228.
Collins, D. (2002). Material prepared by D. Collins for Center for Applied Social Surveys (CASS) course 2002 on “Pretesting Survey Questionnaires” conducted by Campanelli, Collins, & Thomas. Southampton, UK: University of Southampton, Southampton Statistical Science Research Institute, Center for Applied Social Surveys.
Collins, J. (2001). Good to great. New York: Harper Business.
Collins, M. , & Sykes, W. (1999). Extending the definition of survey quality. Journal of Official Statistics, 15, 57–66.
Comley, P. (2000). Pop-up surveys. What works, what doesn’t work and what will work in the future. Proceedings of the ESOMAR Worldwide Internet Conference Net Effects, 237. Retrieved November 21, 2001 , from:
http://www.virtualsurveys.com/papers/popup_paper.htm
Conner, D. R. (1993). Managing at the speed of change. New York: Villard Books.
Conrad, F. G. , & Schober, M. F. (2000). Clarifying question meaning in a household telephone survey. Public Opinion Quarterly, 64, 1–28.
Conrad, F. G. , Schober, M. F. , Dijkstra, W. (in press). Cues of communication difficulty in telephone interviews. In J. Lepkowski , C. Tucker , M. Brick , E. de Leeuw , L. Japec , P. Lavrakas , et al. (Eds.), Advances in telephone survey methodology. New York: Wiley.
Converse, J. M. (1987). Survey research in the united states. Berkeley, CA: University of California Press.
Converse, J. M. , & Presser, S. (1986). Survey questions: Handcrafting the standardized questionnaire. Beverly Hills, CA: Sage.

Cook, C. , Heath, F. , & Thompson, R. L. (2000). A meta-analysis of response rates in web- or Internet-based surveys. Educational and Psychological Measurement, 60, 821–836.
Cook, L. L. , Schmitt-Cascallar, A. P. , & Brown, C. (2005). Adapting achievement and aptitude tests: A review of methodological issues. In R. K. Hambleton , P. F. Merenda , & C. D. Spielberger (Eds.), Adapting educational and psychological tests for cross-cultural assessment. Mahwah, NJ: Erlbaum.
Cooley, P. C. , Miller, H. G. , Gribble, J. N. , & Turner, C. F. (2000). Automating telephone surveys: Using T-ACASI to obtain data on sensitive topics. Computers in Human Behavior, 16, 1–11.
Council for Marketing and Opinion Research . (2001). Highlights from CMOR's cooperation tracking system data. Retrieved November 14, 2005 , from: http://www.cmor.org/rc/rcnews.cfm?aID=0501
Couper, M. P. (1998, August). Measuring survey quality in a CASIC environment. Paper presented at the Joint Statistical Meetings of the American Statistical Association, Dallas, TA.
Couper, M. P. (2000). The good, the bad, and the ugly (Working paper series No. 077). Michigan, MI: University of Michigan, Institute for Social Research, Survey Methodology Program.
Couper, M. P. (2000). Web surveys: A review of issues and approaches. Public Opinion Quarterly, 64, 464–494.
Couper, M. P. (2001). The promises and perils of web surveys. Paper presented at the ASC-conference on the challenge of the Internet. Retrieved January 2006 , from: http://www.asc.org.uk.
Couper, M. P. (2002). New technologies and survey data collection: Challenges and opportunities. Invited paper presented at the International Conference on Improving Surveys. Copenhagen, August 2002. Retrieved January 2006 , from: http://www.icis.dk/ICIS_papers/Keynote1_0_3.pdf.
Couper, M. P. , & de Leeuw, E. D. (2003). Nonresponse in cross-cultural and cross-national surveys. In J. A. Harkness , F. J. R. van de Vijver , & P. Mohler (Eds.), Cross-cultural survey methods (157–177). New York: Wiley, 2003.
Couper, M. P. , Hansen, S. E. , & Sadosky, S. A. (1997). Evaluating interviewer performance in a CAPI survey. In L. Lyberg , P. Biemer , M. Collins , E. de Leeuw , C. Dippo , N. Schwarz ., et al. (Eds.), Survey measurement and process quality. New York: Wiley.
Couper, M. P. , & Lyberg, L. (2005, April). The use of paradata in survey research. Proceedings of the 54th Session of the International Statistical Institute, Sydney, Australia.
Couper, M. P. , & Nichols, J. D. I. . (1998). The history and development of computer-assisted survey information collection methods. In J. M. O'Reilly (Ed.), Computer-assisted survey information collection. New York: Wiley.
Couper, M. P. , Singer, E. , & Tourangeau, R. (2004). Does voice matter? An Interactive Voice Response (IVR) experiment. Journal of Official Statistics, 20, 551–570.
Couper, M. P. , Singer, E. , & Tourangeau, R. (2003). Understanding the effects of audio-CASI on self-reports of sensitive behavior. Public Opinion Quarterly, 67, 385–395.
Couper, M. P. , Tourangeau, R. , & Kenyon, K. (2004). Picture this! Exploring visual effects in web surveys. Public Opinion Quarterly, 68, 255–266.
Couper, M. P. , Tourangeau, R. , Conrad, F. G. , & Singer, E. (2006). Evaluating the effectiveness of visual analog scales: A web experiment. Social Science Computer Review, 24, 227–245.
Couper, M. P. , Traugott, M. , & Lamias, M. (2001). Web survey design and administration. Public Opinion Quarterly, 65, 230–253.
Crawford, S. , Couper, M. P. , & Lamias, M. J. (2001). Web surveys: Perceptions of burden. Social Science Computer Review, 19, 146–162.
Crawford, S. , McCabe, S. E. , & Pope, D. (2005). Applying web-based survey design standards. Journal of Prevention and Intervention in the Community, 29, 43–66.
Cronbach, L. , & Meehl, P. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281–302.
Currivan, D. B. , Nyman, A. L. , Turner, C. F. , & Biener, L. (2004). Does telephone audio computer-assisted self-interviewing improve the accuracy of prevalence estimates of youth smoking? Evidence from the Umass tobacco study. Public Opinion Quarterly, 68, 542–564.

Curtin, R. , Presser, S. & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. Public Opinion Quarterly, 69, 87–98.
Czaja, R. , & Blair, J. (2005). Designing surveys: A guide to decisions and procedures (2nd ed.). Thousand Oaks, CA: Sage.
Czaja, R. , & Blair, J. (1996) Designing surveys: A guide to decisions and procedures. Thousand Oaks, CA: Sage, Pine Forge Press.
Dalenius, T. (1977). Strain at a gnat and swallow a camel: Or, the problem of measuring sampling and non-sampling errors. ASA Proceedings of the Social Statistics Section, 20, 21–25.
Daling J.R. , Malone K. E. , Voigt L.F. , White E. , Weiss N.S. (1994). Risk of breast cancer among young women: relationship to induced abortion. Journal of the National Cancer Institute, 86, 1584–1592.
Dalton, D. R. , Wimbush, J. C. , & Daily, C. M. (1994). Using the Unmatched Count Technique (UCT) to estimate base rates for sensitive behavior. Personnel Psychology, 47, 817–828.
Davis, D. , & Silver, B. (2003). Stereotype threat and race of interviewer effects in a survey on political knowledge. American Journal of Political Science, 47, 33–45.
De Heer, W. (1999). International response trends: Results of an international survey. Journal of Official Statistics, 15, 129–142.
De Jong, W. A. M. (1991). Technieken voor het koppelen van bestanden. Statistische Onderzoekingen, M41. Voorburg: Statistics Netherlands.
De Leeuw, E. D. (2005). Surveying Children. In S.J. Best and B. Radcliff (Eds). Polling America: An Encyclopedia of Public Opinion. p. 831–835, Westport. CT. Greenwood Press.
De Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. Journal of Official Statistics, 21, 233–255.
De Leeuw, E. D. (1999). De vraag naar gevoelige informatie: Een overzicht. [Questions on sensitive information: A Review] Pedagogische Studiën, 76, 92–103.
De Leeuw, E. D. (1992). Data quality in mail, telephone and face-to-face surveys. Amsterdam: TT-Publikaties. Retrieved from: http://www.xs4all.nl/~edithl.
De Leeuw, E. , Borgers, N. , & Smits, A. (2004). Pretesting questionnaires for children and adolescents. In S. Presser , J. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , & E. Singer (Eds.), Methods for testing and evaluating survey questionnaires. Hoboken, NJ: Wiley.
De Leeuw, E. , Callegaro, M. , Hox, J. J. , Korendijk, E. , & Lensvelt-Mulders, G. (2007). The influence of advance letters on response in telephone surveys: A meta-analysis. Public Opinion Quarterly, 71, 1–31.
De Leeuw, E. D. , & De Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves , D. A. Dillman , J. L. Eltinghe , & R. J. A. Little (Eds.), Survey nonresponse. New York: Wiley.
De Leeuw, E. D. & Hox, J.J. (1988). The effects of response stimulating factors on response rates and data quality in mail surveys. Journal of Official Statistics, 4, 241–249. (also at www.jos.nu)
De Leeuw, E. D. & Hox, J. J. (2004). I am not selling anything: 29 Experiments in telephone introductions. International Journal of Public Opinion Research, 16, 464–473.
De Leeuw, E. D. , Hox, J. J. , & Huisman, M. (2003). Prevention and treatment of item nonresponse. Journal of Official Statistics (JOS) 19, 153–176.
De Leeuw, E. D. , Hox, J. J. , & Kef, S. (2003). Computer-Assisted Self-Interviewing tailored for special populations and topics. Field Methods, 15, 223–251.
De Leeuw, E. D. , Lepkowski, J. , & Kim, S. W. (2002, August). Have telephone surveys a future in the twenty-first century? Paper presented at the 2002 International Conference on Information Systems, Copenhagen, Denmark.
De Leeuw, E. D. , & Nichols, J. D. I. (1996). Technological innovations in data collection: Acceptance, data quality and costs. Sociological Research on Line, 1.
De Leeuw, E. D. , & Van der Zouwen, J. (1988). Data quality in telephone and face-to-face surveys: A comparative meta-analysis. In R. Groves , P. Biemer , L. Lybert , J. Massey , W. Nicholls II , & J. Waksberg (Eds.), Telephone survey methodology. New York: Wiley.

DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner & E. Martin (Eds.), Surveying subjective phenomena, volume two. New York: Russell Sage.
DeMaio, T. J. , & Landreth, A. (2004). Do different cognitive interview techniques produce different results? In S. Presser , J. M. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , et al. (Eds.), Methods for testing and evaluating survey questionnaires. New York: Wiley.
DeMaio, T. J. & Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), Answering questions: Methodology for determining cognitive and communicative processes in survey research (p. 177–195). San Francisco: Jossey Bass.
Deming, W. E. (1986). Out of the crisis. Cambridge, MA: Cambridge University Press.
Deming, W. E. (1944). On errors in surveys. American Sociological Review, 9, 359–369.
Deming, W. E. , & Stephan, F. F. (1940). On a least squares adjustment of a sample frequency table when the expected marginal totals are known. Annals of Mathematical Statistics, 11, 427–444.
Dempster, A. P. , Laird, N. M. , & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society B, 39, 1–38.
Denton, C. (2005). National representative report—Costa Rica/Central America. First Quarter Newsletter of the World Association for Public Opinion Research, 3–4.
de Rada, V.D. (2001). Mail Surveys Using Dillman’s TDM in a Southern European Country: Spain. International Journal of Public Opinion Research, 13, 159–172
De Rouvray, C. , & Couper, M. P. (2002). Designing a strategy for reducing ‘no opinion’ responses in web-based surveys. Social Science Computer Review SSCREH, 20, 3–9.
DeVellis, R. F. (2003). Scale development: Theory and applications. Thousand Oaks, CA: Sage.
Devore, J. , & Peck R. (2005). Statistics: The exploration and analysis of data. Belmont: Brooks/Cole-Thomson Learning.
Dewar, A. (2006). Maximizing response to a postal survey on social capital. Office for national Statistics Survey Methodology Bulletin, 58, 29–33. (also at www.statistics.gov.uk/about/services/dcm/reports_publications.asp)
Dillman, D. A. (2007). Mail and internet surveys: The tailored design method (2nd ed.). New York: Wiley.
Dillman, D. A. (2000). Mail and internet surveys: The tailored design method. New York: Wiley.
Dillman, D. A. (1991). The design and administration of mail surveys. Annual Review of Sociology, 17, 225–249.
Dillman, D. A. (1978). Mail and telephone surveys. The Total Design Method. New York: Wiley.
Dillman, D. A. , & Bowker, D. K. (2001). The web questionnaire challenge to survey methodologists. In: U. D. Reips , & M. Bosnjak (Eds.), Dimensions of internet science. Lengerich: Pabst Science Publishers.
Dillman, D. A. , Clark, J. R. , & West, K. K. (1995). Influence of an invitation to answer by telephone on response to census questionnaires. Public Opinion Quarterly, 51, 201–219.
Dillman, D. A. , & Christian, L. M. (2005). Survey mode as a source of instability across surveys. Field Methods, 17, 30–52.
Dillman, D. A. , Dolson, D. E. , Machlis, G. E. (1995). Increasing response to personally-delivered mail-back questionnaires. Journal of Official Statistics, 11, 129–139. (Also at www.jos.nu)
Dillman, D. A. , Gertseva, A. , & Mahon-Haft, T. (2005). Achieving usability in establishment surveys through the application of visual design principles. Journal of Official Statistics, 21, 183–214. (Also at www.jos.nu)
Dillman, D. A. , Phelps, G. , Tortorra, R. , Swift, K. , Kohrell, J. , & Berck, J. (2002). Response rate and measurement differences in mixed mode surveys: Using mail, telephone, Interactive Voice Response and the Internet. Retrieved April 14, 2005 , from: http://www.sesrc.wsu.edu/dillman/papers/Mixed%20Mode%20ppr_with%20Gallup_POQ.pdf.

Dillman, D. A. , & Redline, C. (2004). Testing paper self-administered questionnaires: Cognitive interview and field test comparisons. In S. Presser , J. M. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , et al. (Eds.), Methods for testing and evaluating survey questionnaires. New York: Wiley.
Dillman, D. A. , & Tarnai, J. (1988). Administrative issues in mixed mode surveys. In R. M. Groves , P. P. Biemer , L. E. Lyberg , J. T. Massey , W. L. Nicholls II , & J. Wakesberg (Eds.), Telephone survey methodology. New York: Wiley.
Dillman, D. A. , Tortora, R. D. , & Bowker, D. (1998). Influence of plain vs. fancy design on response rates for web surveys. Proceedings of the section on survey methods research, 21. Retrieved November 26, 2001 , from:
http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf
Dillman, D. A. , Tortora, R. D. , & Bowker, D. (1998). Principles for constructing Web surveys. Retrieved from: http://survey.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf.
Dippo, C. S. (1997). Survey measurement and process improvement: Concepts and integration. In L. Lyberg , P. Biemer , M. Collins , E. de Leeuw , C. Dippo , N. Schwarz , et al. (Eds.), Survey measurement and process quality. New York: Wiley.
Dippo, C. S. , & Sundgren, B. (2000). The role of metadata in statistics. Proceedings of ICES-2, the Second International Conference on Establishment Surveys. Buffalo, NY.
Dorey, F. J. , Little, R. J. A. , & Schenker, N. (1993). Multiple imputation for threshold-crossing data with interval censoring. Statistics in Medicine, 12, 1589–1603.
Dornyei, Z. (2003). Questionnaires in second language research: Construction, administration, & processing. Mahwah, N.J.: Lawrence Erlbaum Associates.
Dowd, K. , Kinsey, S. , Wheeless, S. , Suresh, S. , & The NSCAW Research Group . (2002). National survey of child and adolescent well-being: Wave I data file user's manual. Research Triangle Park, NC: RTI International.
Doyle, P. , Lane, J. , Theeuwes, J. & Zayatz, L. (2001). Confidentiality, disclosure, and data access: Theory and practical applications for statistical agencies. Amsterdam etc.: Elsevier.
DuMouchel, W. H. , & Duncan, G. I. (1983). Using sample survey weights in multiple regression analyses of stratified samples. Journal of the American Statistical Association, 78, 535–543.
Duncan, G. J. , & Kalton, G. (1987). Issues of design and analysis of surveys across time. International Statistical Review, 55, 97–117.
Dunn, L. B. , & Gordon, N. E. (2005). Improving informed consent and enhancing recruitment for research by understanding economic behavior. Journal of the American Medical Association, 293, 609–612.
Edwards, B. , Schneider, C. , & Dean Brick, P. (in press). Visual elements of questionnaire design: Experiments with a CATI establishment survey. In J. Lepkowski , C. Tucker , M. Brick , E. de Leeuw , L. Japec , P. Lavrakas , et al. (Eds.), Advances in telephone survey methodology. New York: Wiley.
Efron, B. (1994). Missing data, imputation, and the bootstrap. Journal of the American Statistical Association, 89, 463–479.
Efron, R. , & Tibshirani, R. J. (1993). An introduction to the bootstrap. London: Chapman & Hall.
Eichman, C. (1999). Research methods on the web: Experiences and implications for market research. Proceedings of the ESOMAR Worldwide Internet Conference Net Effects, 69–76.
Enander, J. , & Sajti, A. (1999, February). Online survey of online customers: Value-added market research through data collection on the Internet. Proceedings of the ESOMAR Worldwide Internet Conference Net Effects, 35–51.
Ericsson K. A. , & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.
Erikson, B. H. , & Nosanchuk, T. A. (1983). Applied network sampling. Social Networks, 5, 367–382.
Esposito, J. L. , Campanelli, P. C. , Rothgeb, J. M. , & Polivka, A. E. (1991). Determining which questions are best: Methodologies for evaluating survey questions. Proceedings of the Section of Survey Methods Research, 14.
Eurostat . (2005). European Statistics Code of Practice. For the national and community statistical authorities. Retrieved from:

http://epp.eurostat.cec.eu.int/pls/portal/docs/page/pgp_ds_quality/-tab47141301/versione_inglese_web.pdf.
Eurostat (2000, April). Assessment of the Quality in Statistics (Eurostat/A4/Quality/00/General/Standard report). Luxembourg. Retrieved August, 2006 , from: http://www.unece.org/stats/documents/2000/11/metis/crp.2.e.pdf.
Ewing, M. , Caruana, A. , & Zinkhan, G. M. (2002). On the cross-national generalisability and equivalence of advertising response scales developed in the USA. International Journal of Advertising, 21, 323–343.
Faden, R.R. & Beauchamp, T.L. (1986). A history and theory of nformed consent. New York: Oxford University Press.
Fellegi, I. , & Ryten, J. (2005). The effectiveness of a supranational statistical office: Pluses, minuses, and challenges viewed from the outside. Journal of Official Statistics, 21, 145–170.
Fisher, R.A. (1938). Quotation from address to the First Indian Statistical Congress.
Flaherty, D. H. (1989). Protecting privacy in surveillance societies: the federal republic of Germany, Sweden, France, Canada, and the US. Chapel Hill, NC: University of North Carolina Press.
Fienberg, S.E. & Willenborg, L.C.R.J. (1998). Special issue on confidentiality of statistical data. Journal of Official Statistics, 14, 4.
Folsom, R. E. , & Singh, A. C. (2000). The general exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification. Proceedings of the Survey Research Methods Section, 23, 598–603.
Forsyth, B. H. , & Hubbard, M. L. (1992, August). A method for identifying cognitive properties of survey items. Paper presented at the annual meeting of the American Statistical Association, Boston, MA.
Forsyth, B. H. , & Lessler, J. T. (1991). Cognitive laboratory methods: A taxonomy. In P. P. Biemer , R. M. Groves , & L. E. Lyberg (Eds.), Measurement errors in surveys. New York: Wiley.
Forsyth, B. H. , Rothgeb, J. , & Willis, G. (2004). Does pretesting make a difference? In S. Presser , J. M. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , et al. (Eds.), Methods for testing and evaluating survey questionnaires. New York: Wiley.
Fowler, F. J. (2004). Getting beyond pretesting and cognitive interviews: The case for more experimental pilot studies. In S. Presser , J. M. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , et al. (Eds.), Methods for testing and evaluating survey questionnaires. New York: Wiley.
Fowler, F. J. (2001). Why it is easy to write bad questions. ZUMA-Nachrichten, 48, 49–66.
Fowler, F. J. (1997). Choosing questions to measure the quality of experience with medical providers and health plans. ASA Proceedings of the section on Survey Research Methods, 20, 51–54.
Fowler, F. J. (1995). Improving Survey Questions: Design and evaluation. Thousands Oaks, CA: Sage.
Fowler, F. J. (1992). How unclear terms affect survey data. Public Opinion Quarterly, 56, 218–231.
Fowler, F. J. , & Cannell, C. F. (1996). Using behavioral coding to identify cognitive problems with survey questions. In N. Schwarz , & S. Sudman (Eds.), Answering questions: Methodology for determining cognitive and communicative processes in survey research (pp. 15–36). San Francisco: Jossey-Bass.
Fowler, F. J. , Gallagher, P. M. , Stringfellow, V. L. , Zalavsky, A. M. , Thompson, J. W. , & Cleary, P. D. (2002). Using telephone interviews to reduce nonresponse bias to mail surveys of health plan members. Medical Care, 40, 190–200.
Fowler, F. J. , & Mangione, T. W. (1990). Standardized survey interviewing: Minimizing interviewer-related error. Beverly Hills, CA: Sage Publications.
Fowler, F. J. , & Mangione, T. W. (1986). Reducing interviewer effects on health survey data. Washington, DC: National Center for Health Services Research.
Fowler, F. J. , & Mangione, T. W. (1984). Standardized survey interviewing. ASA Proceedings of the Annual Meeting of the American Statistical Association. Arlington, VA: ASA.

Fowler, F. J. , & Mangione, T. W. (1982). The effect of training and supervision on common measures of field interviewer performance. NCHSR Research Proceedings Series Health Survey Research Methods. Arlington, VA: ASA.
Fox, J. A. , & Tracy, P. E. (1986). Randomized response: A method for sensitive surveys. Newbury Park, CA: Sage.
Frey, J. H. (1989). Survey research by telephone (2nd ed.). Thousand Oaks, CA: Sage.
Friedrich-Freksa, M. , & Liebelt, M. (2005, March). From short message service (SMS) questionnaires to mobile internet surveys: Integrating the mobile phone into market research. Paper presented at the Seventh International General Online Research Conference (GOR '05), Zurich, Switzerland.
Fuchs, M. (2005). Children and adolescents as respondents: Experiments on question order, response order, scale effects and the effect of numeric values associated to response options. Journal of Official Statistics, 21, 701–725.
Fuller, W. A. (1987). Measurement error models. New York: Wiley.
Gabler, S. & Häder, S. (2000). Überlegungen zur Anwendung von RLD-Verfahren bei Telefonumfragen in Deutschland. Pp. 33–47 in: V. Hüfgen , (Ed.), Methoden in Telefonumfragen. Opladen, FRG: Westdeutscher Verlag GmbH,
Gagné R. , Briggs, L. J. , & Wager, W. W. (1989). Principles of instructional design. New York: Holt, Rinehart, & Winston.
Gagné, R. (1962). The acquisition of knowledge. Psychological Review, 69, 355–365.
Gagné, R. , Yekovich, C. W. , & Yekovich, F. R. (1993). The cognitive psychology of school learning. New York: Harper Collins.
Gauthier, A. H. (2000). The promises of comparative research. Paper prepared for the European Panel Analysis Group. Retrieved June 23, 2006 , from: http://www.iser.essex.ac.uk/-epag/pubs/workpaps/pdf/2000-16.pdf.
Gaziano, C. (2005). Comparative analysis of within-household respondent selection techniques. Public Opinion Quarterly, 69, 124–157.
Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychological Assessment, 6, 304–312.
Gelfand, A. E. & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association, 85, 398–409.
Gelman, A. , & Carlin, J. B. (2002). Poststratification and weighting adjustment. In R. M. Groves , D. A. Dillman , J. L. Eltinge , & R. J. A. Little (Eds.), Survey nonresponse. New York: Wiley.
Glynn, R. , Laird, N. M. , & Rubin, D. B. (1986). Selection modeling versus mixture modeling with nonignorable nonresponse. In H. Wainer (Ed.), Drawing inferences from self-selected samples. New York: Springer.
Goerman, P. L. (2006). Adapting cognitive interview techniques for use in pretesting Spanish language survey instruments (U.S. Census Bureau Research Report Series, Survey Methodology No. 2006–3). Retrieved June 20, 2006 , from: http://www.census.gov/srd/papers/pdf/rsm2006-03.pdf
Gonier, D. E. (1999). The emperor gets new clothes. In: Towards validation: Online research day: An ARF emerging issue workshop (pp. 8–13). New York: Advertising Research Foundation. Retrieved January 9, 2001 , from: http://www.dmsdallas.com/emporere/emporer.html
Goode, W. J. (1973). Explorations in social theory. New York, etc.: Oxford University Press.
Goodwin, C. J. (2005). Research in psychology: Methods and design (4th ed). New York: Wiley.
Göritz, A. (2005). Incentives in web-based studies: What to consider and how to decide (WebSM Guide No. 2). Web Survey Methodology Site. Retrieved December 29, 2005 , from: http://www.websm.org/uploadi/editor/goeritz2005-incentives.pdf
Graesser, A. , Bommareddy, S. , Swamer, S. , & Golding, J. (1996). Integrating questionnaires design with a cognitive computational model of human question answering. In N. Schwarz & S. Sudman (Eds.), Answering questions: Methodology for determining cognitive and communicative processes in survey research. San Francisco: Jossey Bass.

Graesser, A. , Kennedy, T. , Wiemer-Hastings, P. , & Ottati, V. (1999). The use of computational cognitive models to improve questions on surveys and questionnaires. In M. Sirken , D. Herrmann , S. Schechter , N. Schwarz , J. Tanur , and R. Tourangeau (Eds.), Cognition and survey research. New York: Wiley.
Gräf, L. (2002). Assessing internet questionnaires: The online pretest lab. In: B. Batinic , U. D. Reips , M. Bosnjak , & A. Werner (Eds.), Online social sciences (73–93). Hogrefe & Huber, Seattle.
Gray, P. G. (1955). The memory factor in social surveys. Journal of the American Statistical Association, 50, 344–363.
Greenfield, P. M. (1997). You can’t take it with you: Why ability assessments don’t cross cultures. American Psychologist, 52, 1115–1124.
Grembowski, D. , & Phillips, D. (2005). Linking mother and child access to dental care: A multimode survey. (NIH-funded study, grant # DE 14400. SESRC data report No. 05–023). Pullman, WA: WSU.
Gribble, J. N. , Miller, H. G. , Cooley, P. C. , Catania, J. A. , Pollack, L. & Turner, C. F. (2000). The impact of T-ACASI interviewing on reporting drug use among men who have sex with men. Substance Use and Misuse, 80, 869–890.
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and semantics, vol. 3: Speech acts. New York: Academic Press.
Grice, P. (1989). Studies in the way of words. Cambridge, MA: Harvard University Press.
Grimshaw, A. D. (1973). Comparative Sociology: In what ways different from other Sociologies? In M. Armer , & A. D. Grimshaw (Eds.), Comparative social research: Methodological problems and strategies. New York: Wiley.
Groves, R. M. (in press). Research synthesis: Nonresponse bias. Public Opinion Quarterly.
Groves, R. M. (1989). Survey errors and survey costs: An introduction to survey errors. New York: Wiley.
Groves, R.M. , Cialdini, R.B. & Couper, M.P. (1992). Understanding the decision to participate in a survey. Public Opinion Quarterly, 56, 475–495.
Groves, R. M. , & Couper, M. P. (1998). Nonresponse in household interview surveys. New York: Wiley.
Groves, R. M. , Dillman, D. A. , Eltinge, J. L. , & Little, R. J. A. (2002). Survey nonresponse. New York: Wiley.
Groves, R. M. , Fowler, F. J. , Couper, M. P. , Lepkowski, J. M. , Singer, E. , & Tourangeau, R. (2004). Survey methodology. New York: Wiley.
Groves, R. M. , & Kahn, R. (1979). Surveys by telephone: A national comparison with personal interviews. New York: Academic Press.
Groves, R. M. , & Lepkowski, J. M. (1985). Dual frame, mixed mode survey designs. Journal of Official Statistics, 1, 263–286.
Groves, R. M. , & McGonagle, K. (2001). A theory-guided interviewer training protocol regarding survey participation. Journal of Official Statistics, 17, 249–265.
Groves, R. M. , Singer, E. , & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. Public Opinion Quarterly, 64, 299–308.
Guest, L. (1954). A new training method for opinion interviewers. Public Opinion Quarterly, 18, 287–299.
Haberstroh, S. , Oyserman, D. , Schwarz, N. , Kühnen, U. , & Ji, L. J. (2001). Is the interdependent self more sensitive to question context than the independent self? Self-construal and the observation of conversational norms. Journal of Experimental Social Psychology, 38, 323–329.
Häder, S. , & Gabler, S. (2003). Sampling and estimation. In J. A. Harkness , F. J. R. van de Vijver , & P. Mohler (Eds.), Cross-cultural survey methods. Hoboken, NJ: Wiley.
Hall, T.E. , & Roggenbuck, J.W. (2002). Response format effects in questions about norms: Implications for the reliability and validity of the normative approach. Leisure Sciences, 24, 325–338.
Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton , P. F. Merenda , & C. D. Spielberger (Eds.), Adapting educational and psychological tests for cross-cultural assessment. Mahwah, NJ: Erlbaum.

Hambleton, R. K. , Merenda, P. F. & Spielberger, C. D. (Eds.) (2005). Adapting educational and psychological tests for cross-cultural assessment. Mahwah, NJ: Erlbaum.
Haney, C. , Banks, C. , & Zimbardo, P. (1973). Interpersonal dynamics in a simulated prison. International Journal of Criminology and Penology, 1, 69–97.
Hanh, V. T. X. , Guillemin, F. , Cong, D. D. , Parkerson, G. R. Jr. , Thu, P. B. , Quynh, P. T. , & Briançon, S. (2005). Health related quality of life of adolescents in Vietnam: Cross-cultural adaptation and validation of the Adolescent Duke Health Profile. Journal of Adolescence, 28, 127–146.
Hansen, M. H. , Hurwitz, W. N. , & Bershad, M. A. (2001). Measurement error in censuses and surveys. Landmark papers in survey statistics. The IASS Jubilee Commemorative Volume. International association of survey statisticians of the International Statistical Institute. (Original work published 1961)
Hansen, M. H. , Hurwitz, W. N. , & Bershad, M. A. (1961). Measurement errors in censuses and surveys. Bulletin of the International Statistical Institute, 38, 359–374.
Hansen, M. H. , Madow, W. G. , & Tepping, B. J. (1983). An evaluation of model-dependent and probability-sampling inferences in sample surveys. Journal of the American Statistical Association, 78, 776–793.
Hantrais, L. , & Mangen, S. (1996). Cross-national research methods in the social sciences. London: Pinter.
Harkness, J.A. (2002/2007). ESS Translation Strategies and Procedures. Retrieved July 25, 2007 , from European Social Survey site: www.europeansocialsurvey.org
Harkness, J. A. (2006a). Round 3 ESS translation strategies and procedures. Retrieved June 16, 2006 , from:
http://naticent02.uuhost.uk.uu.net/ess_docs/R3/Methodology/r3_translation_guidelines.doc.
Harkness, J. A. (2006b, July). What happens when you start working on guidelines. Keynote presentation at the Fifth International Test Commission Conference, Brussels, Belgium.
Harkness, J. A. (1999). In pursuit of quality: Issues for cross-national survey research. International Journal of Social Research Methodology, 2, 125–140.
Harkness, J. A. , Langfeldt, B. , & Scholz, E. (2001). International social survey programme study monitoring 1996–1998. Reports to the International Social Survey Programme General Assembly on Monitoring Work Undertaken for the International Social Survey Programme by Zentrum für Umfragen, Methoden und Analysen, Germany.
Harkness, J.A. , Mohler, P.Ph. , and van de Vijver, F.J.R. (2003) in Harkness, J.A. , van de Vijver, F.J.R. & Mohler, P.Ph. (Eds.). Cross-Cultural Survey Methods. Hoboken, New Jersey: Wiley. p. 3–16.
Harkness, J. A. , Pennell, B. E. , & Schoua-Glusberg, A. (2004). Survey questionnaire translation and assessment. In S. Presser , J. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , & E. Singer (Eds.), Methods for testing and evaluating survey questionnaires. Hoboken, NJ: Wiley.
Harkness, J. A. , Schoebi, N. , Joye, D. , Mohler, P. , Faass, T. , Behr, D. (2007) Oral translation in telephone surveys (in press). J. Lepkowski , C. Tucker , M. Brick , E. de Leeuw , L. Japec , P. Lavrakas , et al. (Eds.), Advances in telephone survey methodology. New York: Wiley.
Harkness, J. A. , & Schoua-Glusberg, A. (1998). Questionnaires in translation. In J. A. Harkness (Ed.), ZUMA-Nachrichten Spezial No. 3. Cross-Cultural Survey Equivalence. Mannheim: ZUMA.
Harkness, J. A. , Van de Vijver, F. J. R. , & Johnson, T. P. (2003). Questionnaire design in comparative research. In J. A. Harkness , F. J. R. van de Vijver & P. Mohler (Eds.), Cross-cultural survey methods. Hoboken, NJ: Wiley.
Harkness, J. A. , Van de Vijver, F. J. R. , & Mohler, P. (2003). Cross-cultural survey methods. New York: Wiley.
Harmon, M. , Westin, E. , & Levin, K. (2005, May). Does type of pre-notification affect web survey response rates? Paper presented at the American Association for Public Opinion Research 60th Annual Conference, Miami Beach, FL.
Harris-Kojetin, L. D. , Fowler, F. J. , Brown, J. A. , Schnaier, J. A. , & Sweeny, S. F. (1999). The use of cognitive testing to develop and evaluate CAHPS 1.0 core survey items. Medical Care, 37, Supplement, MS10–MS21.

Haworth, M. , & Signore, M. (2005, April). Quality measurement and reporting: Frameworks, guidelines and user needs. Invited paper for the ISI meeting, Sydney, Australia.
Headland, T. N. , Pike, K. L. , & Harris, M. (1990). Emics and etics: The insider / outsider debate. Newbury Park, CA.: Sage.
Heath, A. , & Martin, J. (1977). Why are there so few formal measuring instruments in social and political research?. In L. Lyberg , P., Biemer , M. Collins , E. de Leeuw , C. Dippo , N. Schwarz , & D. Trewin (Eds.), Survey measurement and process quality (71–86). New York: Wiley.
Heberlein, T. A. , & Baumgartner, R. M. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. American Sociological Review, 43, 447–462.
Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement, 5, 475–492.
Heerwegh, D. , & Loosveldt, G. (2002). Web surveys: The effect of controlling survey access using PIN numbers. Social Science Computer Review, 20, 10–21.
Heitjan, D. F. , & Little, R. J. A. (1991). Multiple imputation for the fatal accident reporting system. Journal of the Royal Statistical Society C, 40, 13–29.
Herdman, M. , Fox-Rushby, J. , & Badia, X. (1997). 'Equivalenc?' and the translation and adaptation of health-related quality of life questionnaires. Quality of Life Research, 6, 237–247.
Herr, P. M. (1986). Consequences of priming: Judgment and behavior. Journal of Personality and Social Psychology, 51, 1106–1115.
Hert, C. , Denn, S. , & Haas, S. W. (2004). The role of metadata in the statistical knowledge network: An emerging research agenda. Social Science Computer Review, 22, 92–99.
Hess, J. , Singer, E. , & Bushery, J. (1999). Predicting test-retest reliability from behavior coding. International Journal of Public Opinion Research, 11, 346–360.
Hill, T. P. (1998). The first digit phenomena. Retrieved from: http://www.math.gatech.edu/~hill/-publications/cv.dir/1st-dig.pdf.
Himmelfarb, S. , & Lickteig, C. (1982). Social desirability and the randomized response technique. Journal of Personality and Social Psychology, 43, 710–717.
Hippler, H.-J. , & Seidel, K. (1985). Schriftliche Befragung bei algemeine Bevolkerungsstichproben-Untersuchungen zur Dillman'schen 'Total Design Method'. ZUMA-Nachrichten, 16. Mannheim: ZUMA
Ho, D. (2003). "Do-not-call" still a big hit. Retrieved November 14, 2005 , from: http://www.cbsnews.com/stories/2003/03/11/politics/main543573.shtml.
Hochstim, J. R. (1967). A critical comparison of three strategies of collecting data from households. Journal of the American Statistical Association, 62, 976–989.
Hoffmeyer-Zlotnik, J. H. P. (1997). Random Route Stichproben nach ADM. In S. Gabler & J. H. P. Hoffmeyer-Zlotnik (Eds.), Stichproben in derUmfragepraxis. Opladen: Westdeutscher Verlag.
Hoffmeyer-Zlotnik, J. H. P. , & Harkness, J. A. (Eds.), (2005) ZUMA-Nachrichten Spezial No. 11. Methodological aspects in cross-national research. Mannheim: ZUMA.
Hoffmeyer-Zlotnik, J. H. P. , & Wolf, C. (Eds.) (2003). Advances in cross-national comparison. A European working book for demographic and socio-economic variables. New York: Kluwer Academic/Plenum Publishers.
Holt, D. , & Jones, T. (1998). Quality work and conflicting quality objectives. Paper prepared for DGINS meeting, Eurostat., May 2005.
Hoogendoorn, A. W. & Sikkel, D. (1998). Response burden and panel attrition. Journal of Official Statistics, 14, 189–205.
Hox, J. J. (2002). Multilevel analysis: Techniques and applications. Mahwah, N.J.: Erlbaum.
Hox, J. J. (1997). From theoretical concept to survey item. In: L. Lyberg , P. Biemer , M. Collins , E. de Leeuw , C. Dippo , N. Schwarz , & D. Trewin (Eds), Survey Measurement and Process Quality (47–71). New York: Wiley.
Hox, J. J. , & De Jong-Gierveld, J. J. (Eds.) (1990). Operationalization and research strategy. Lisse, NL: Swets & Zeitlinger.

Hox, J. J. & De Leeuw, E. D. (1994). A comparison of nonresponse in mail, telephone, and face-to-face surveys. Applying multilevel modeling to meta-analysis. Quality & Quantity, 329–344.
Hox, J. J. , De Leeuw, E. D. , & Kreft, G. G. (1991). The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In P. P. Biemer , R. M. Groves , L. E. Lyberg , N. A. Mathiowetz , & S. Sudman (Eds.), Measurement Errors in Surveys. New York: Wiley.
Hoyle, R. H. , Harris, M. J. , & Judd, C. M. (2002). Research methods in social relations. London: Wadsworth.
http://news.com.com/Start-up+aims+to+join+telephone%2C+wireless+calls/2100-1033_3-5950752.html.
Hughes, L. C. , & Preski, S. (1997). Using key informant methods in organizational survey research: Assessing for informant bias. Research in Nursing and Health, 20, 81–92.
Humphreys, L. (1970). Tearoom trade: Impersonal sex in public places. Chicago: Aldine.
Hunt, S. D. , Sparkman, R. D. Jr. , & Wilcox, J. B. (1982). The pretest in survey research: Issues and preliminary findings. Journal of Marketing Research, 19, 269–273.
Hyman, H. H. (1955). Survey design and analysis. Glencoe: Free Press.
International Standards Organization . (2006). Standards for Marketing, Opinion and Social Research (ISO Standard 20252: 2006). Retrieved from: http://www.iso.org/iso/-en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=39339&ICS1=1&ICS2=40&ICS3=3.
International Working Group for Disease Monitoring and Forecasting , (1995). Capture-recapture and multiple-record systems estimation. American Journal of Epidemiology, 142, 1047–1058.
Jabine, T. , King, K. , & Petroni, R. (1990). Quality profile for the Survey of Income and Program Participation (SIPP). Washington, DC: U.S. Bureau of the Census.
Jabine, T. , Straf, M. , Tanur, J. , & Tourangeau, R. (Eds.). (1984). Cognitive aspects of survey methodology: Building a bridge between disciplines. Washington, DC: National Academy Press.
Jackson, C. P. , & Boyle, J. M. (1991). Mail response rate improvement in a mixed-mode survey. Proceedings ASA-SMRS: American Statistical Association. Retrieved April 14, 2005 , from: www.amstat.org/sections/srms/proceedings/.
Japec, L. (2005). Quality issues in interview surveys: some contributions. Stockholm: Department of Statistics, University of Stockholm.
Javeline, D. (1999). Response effects in polite cultures: A test of acquiescence in Kazakhstan. Public Opinion Quarterly, 63, 1–28.
Jenkins, C. , & Dillman, D. A. (1997). Towards a theory of self-administered questionnaire design. In L. Lyberg , P. Biemer , M. Collins , E. de Leeuw , K. Dippo , N. Schwarz , et al. (Eds.), Survey measurement and process quality. New York: Wiley.
Jensen, U. , & Rässler, S. (2006). Stochastic production frontiers with multiply imputed German establishment data. Zeitschrift für Arbeitsmarktforschung, 39, 277–295.
Ji, L. , Schwarz, N. , & Nisbett, R. E. (2000). Culture, autobiographical memory, and behavioral frequency reports: Measurement issues in cross-cultural studies. Personality and Social Psychology Bulletin, 26, 586–594.
Johnson, T. P. (1998). Approaches to equivalence in cross-cultural and cross-national survey research. In J. A. Harkness (Ed.), ZUMA-Nachrichten Spezial No. 3.: Cross-cultural survey equivalence. Mannheim: ZUMA.
Johnson, T. P. , & Van de Vijver, F. J. R. (2003). Social desirability in cross-cultural research. In J. A. Harkness , F. J. R. van de Vijver , & P. Mohler (Eds.), Cross-cultural survey methods. Hoboken, NJ: Wiley.
Johnson, T. P. , Kulesa, P. , Cho, Y. I. , & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. Journal of Cross-Cultural Psychology, 36, 264–277.
Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. Psychometrika, 34, 183–202.
Jowell, R. (1998). How comparative is comparative research? American Behavioral Scientist, 42, 168–177.

Juran, J. M. , & Gryna, F. M. Jr. (1980). Quality planning and analysis: From product development through use. New York,: McGraw-Hill.

Jussaume, R. A. & Yamada, Y. (1990). A comparison of the viability of mail surveys in Japan and the United States. Public Opinion Quarterly, 54, 219–228.

Kaczmirek, L. , Neubarth, W. , Bosnjak, M. , & Bandilla, W. (2005, March). Progress indicators in filter based surveys. Individual and dynamic calculation methods. Paper presented at General Online Research Conference 2005, Zurich, Switzerland.

Kalton, G. (1977). Practical methods for estimating survey sampling errors. Bulletin of the International Statistical Institute, 47, 495–514.

Kalton, G. (1983). Introduction to survey sampling. Thousand Oaks, CA: Sage.

Kalton, G. (1986). Handling wave nonresponse in panel surveys. Journal of Official Statistics, 2, 303–314.

Kalton, G. (2000). Developments in survey research in the past 25 years. Survey Methodology, 26, 3–10.

Kalton, G. , & Anderson, D. W. (1986). Sampling rare populations. Journal of the Royal Statistical Society, Series A, 149, 65–82.

Kalton, G. , Winglee, M. , Krawchuk, S. , & Levine, D. (2000). Quality Profile for SASS: Rounds 1–3: 1987–1995 (NCES 2000–308). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Kanso, A. (2000). Mail surveys: key factors affecting response rates. Journal of Promotion Management, 5, 2, 3–16.

Kaplowitz, M. D. , Hadlock, T. D. , & Levine, R. (2004). A comparison of web and mail survey response rates. Public Opinion Quarterly, 68, 94–101.

Kasprzyk, D. , Duncan, G. , Kalton, G. , & Singh, M. P. (1989). Panel surveys. New York: John Wiley and Sons.

Kasprzyk, D. , & Kalton, G. (2001, May). Quality profiles in U.S. statistical agencies. Paper presented at the International Conference on Quality in Official Statistics, Stockholm, Sweden.

Katz, J. (1972). Experimenting with human beings. New York: Russell Sage.

Kennickell, A. B. (1991). Imputation of the 1989 survey of consumer finances: Stochastic relaxation and multiple imputation (with discussion). Proceedings of the Section on Survey Research Methods, 14, 1–10.

Kerlinger, F. N. (1986). Foundations of behavioral research. New York: Holt, Rinehart & Winston.

Killworth, P. D. , Johnsen, E. C. , McCarthy, C. , Shelley, G. A. , & Bernard, H. R. (1998). A social network approach to estimating seropositive prevalence in the US. Social Networks, 18, 289–312.

Kim, M. (1994). Cross-cultural comparisons of the perceived importance of conversational constraints. Human communication Research, 21, 128–151.

Kish, L (1995). Survey sampling. New York: Classic Wiley Paperback, Wiley (original work published 1965).

Kish, L. (1992). Weighting for unequal P *i* . Journal of Official Statistics, 8, 183–200.

Kish, L. (1987). Statistical design for research. New York: Wiley.

Kish, L. (1965). Survey sampling. New York: Wiley.

Kish, L. (1962). Studies of interviewer variance for attitudinal variables. Journal of the American Statistical Association, 57, 92–115.

Kish, L. & Frankel, M.R. (1974). Inferences from complex samples. Journal of the Royal Statistical Society (B), 36, 1–37.

Kittleson, M. J. (1997). Determining effective follow-up of e-mail surveys. American Journal of Health Behavior, 21, 193–196.

Kline, R. B. (2005). Principles and practice of Structural Equation Modeling. New York/London: The Guildford Press.

Knapp, F. , & Heidingsfelder, M. (2001). Drop-out analysis: Effects of the survey design. In U. D. Reips , & M. Bosnjak (Eds.), Dimensions of internet science. Lengerich: Pabst Science Publishers.

Knäuper, B. (1999a). Age differences in question and response order effects. In N. Schwarz , D. Park , B. Knäuper , & S. Sudman (Eds.), Cognition, aging, and self-reports. Philadelphia: Psychology Press.
Knäuper, B. (1999b). The impact of age and education on response order effects in attitude measurement. Public Opinion Quarterly, 63, 347–370.
Knäuper, B. , Schwarz, N. , & Park, D. (2004). Frequency reports across age groups. Journal of Official Statistics, 20, 91–96.
Kohn, M. L. (1987). Cross-national research as an analytic strategy: American Sociological Association, 1987 Presidential Address. American Sociological Review, 52, 713–731.
Kripke, D. F. , Garfinkel, L. , Wingard, D. L. , Klauber, M. R. , & Marler, M. R. (2002). Mortality associated with sleep duration and insomnia. Archives of General Psychiatry, 59, 131–136.
Krosnick, J. A. (1991). Response strategies for coping with cognitive elements of attitude measures in surveys. Applied Cognitive Psychology, 5, 213–236.
Krosnick, J. A. , & Alwin, D. F. (1987). An evaluation of a cognitive theory of response order effects in survey measurement. Public Opinion Quarterly, 51, 201–219.
Krosnick, J. A. , & Fabrigar, L. (1997). Comparisons of party identification and policy preferences: The impact of survey question format. American Journal of Political Science, 37, 941–964.
Krotki, K. P. (2001, August). Web-based surveys for a nationally representative sample. Paper presented at Joint Statistical Meeting, Atlanta, GA.
Kuk, A.Y. S. (1990). Asking sensitive questions indirectly. Biometrica, 77, 436–438.
Kuusela, V. (2003, September). Mobile phones and telephone survey methods. In R. Banks , J. Currall , J. Francis , L. Gerrard , R. Kahn , T. Macer , et al. (2003). Proceedings of the fourth ASC international conference. Chesham, UK: ASC.
Kuusela, V. , & Simpanen, M. (2002). Effects of mobile phones on telephone survey practices and results. Paper presented at the International Conference on Improving Surveys, Copenhagen, Denmark. Retrieved November 16, 2005 , from: http://www.icis.dk/ICIS_papers/A_2_3.pdf.
Kvale, S. (1996). Interviews: An introduction to qualitative research interviewing. Thousand Oaks, CA: Sage.
Kwak, N. , & Radler, B. T. (2002). A comparison between mail and web surveys: response pattern, respondent profile, and data quality. Journal of Official Statistics, 18, 257–274.
Kwak, N. , & Radler, B. T. (1999, November). A comparison between mail and web-based surveys: Response pattern, data quality, and characteristics of respondents. Paper presented at 1999 Annual Research Conference, organised by Midwest Association for Public Opinion Research, Chicago, IL.
Lahiri, S.N. (2003). Resampling methods for dependent data. New York: Springer-Verlag.
Laiho, J. (2005, April). Discussion of session on quality measurement and reporting for surveys. Paper presented at the ISI meeting, Sydney, Australia.
Laiho, J. , & Lynn, P. (1999, October). Separating the stages of the survey co-operation process. Paper presented at the International Conference on Household Survey Nonresponse, Portland Oregon.
Lang, S. (2004). Randomized response: Befragungstechniken zur vermeidung von verzerrungen by sensitiven fragen. Munchen: Universität München.
LaPlant, W. P. Jr. , Lestina, G. J. Jr. , Gillman, D. W. , & Appel, M. V. (1996, March). Proposal for a Statistical Metadata Standard. Paper presented at the Census Annual Research Conference, Arlington, VA.
Lau, L. K. P. (2004, August). Mobile phone surveys in Hong Kong: Methodological issues and comparisons with conventional telephone surveys. Paper presented at the RC33 Sixth Annual Conference on Social Science Methodology, Amsterdam, The Netherlands.
Lavrakas, P. J. (2004, May). Will a ‘perfect storm’ of cellular-linked forces sink RDD sampling? Paper presented at the Fifty-Ninth Annual Conference of the American Association for Public Opinion Research, Phoenix, AZ.
Lavrakas, P. J. (1987). Telephone survey methods: Sampling, selection and supervision. Newbury Park, CA: Sage Publications.
Lee, E. S. , Forthofer, R. N. , & Lorimor, R. J. (1989). Analyzing complex survey data. Newbury Park, CA: Sage Publications.

Lee, H. , Rancourt, E , & Särndal, C. E. (2002). Variance estimation for survey data under single imputation. In R. M. Groves , D. A. Dillman , J. L. Eltinge , & R. J. A. Little (Eds.), Survey nonresponse. New York: Wiley.
Lee, R. M. (1993). Doing research on sensitive topics. London: Sage Publications.
Lee, R. M. , & Renzetti, C. M. (1993). The problems of researching sensitive topics: An overview and introduction. In C. M. Renzetti & R. M. Lee (Eds.), Researching sensitive topics. Newbury Park, CA: Sage.
Lee, S. (2006). Propensity score adjustment as a weighting scheme for volunteer panel web surveys. Journal of Official Statistics, 22, 2, 329–349. Also available at www.jos.nu.
Lehtonen, R. , & Pahkinen, E. (2004). Practical methods for design and analysis of complex surveys (2nd ed.). New York: Wiley.
Lensvelt-Mulders, G. J. L. M. , Hox, J. J. , & Van der Heijden, P. G. M. (2005). How to improve the efficiency of randomized response designs. Quality and Quantitiy, 39, 253–265.
Lensvelt-Mulders, G. J. L. M. , Hox, J. J. , Van der Heijden, P. G. M. , & Maas, C. J. M. (2005). Meta-analysis of randomized response research: 35 Years of validation studies. Sociological methods and research, 33, 319–348.
Lepkowski, J. M. (2005). Non-observation error in household surveys in developing countries. In United Nations , Household sample surveys in developing transition countries (pp. 140–170). New York: United Nations. Retrieved July 21, 2006 from: http://unstats.un.org/unsd/hhsurveys/pdf/-Household_surveys.pdf.
Lepkowski, J. M. (1988). Telephone sampling methods in the United States. In R. M. Groves , P. P. Biemer , L. E. Lyberg , J. T. Massey , W. L. Nicholls II , & J. Waksberg (Eds.), Telephone survey methodology. New York: John Wiley and Sons.
Lesser, V. M. , & Newton, L. (2001, May). Mail, e-mail and web surveys: A cost and response rate comparison in a study of undergraduate research activity. Paper presented at 2001 Association for Public Opinion Research Annual Conference, Montreal, Quebec, Canada.
Lessler, J. T. , & Forsyth, B. H. (1996). A coding system for appraising questionnaires. In N. Schwarz & S. Sudman (Eds.), Answering questions: Methodology for determining cognitive and communicative processes in survey research. San Francisco: Jossey Bass.
Lessler, J. T. , & Kalsbeek, W. (1992). Nonsampling errors in surveys. New York: Wiley.
Linacre, S. J. , & Trewin, D. J. (1993). Total survey design: Application to a collection of the construction industry. Journal of Official Statistics, 9, 611–621.
Link, M. W. , & Mokdad, A. H. (2004, May). Responding to the national do not call registry: Evaluation of call attempt protocol changes in the BRFSS. Paper presented at the fifty-ninth annual conference of the American Association for Public Opinion Research, Phoenix, AZ.
Link, M. W. , & Oldendick, R. (1999). Call screening. Public Opinion Quarterly, 63, 577–589.
Linton, M. (1982). Transformations of memory in everyday life. In U. Neisser (Ed.), Memory observed: Remembering in natural contexts. San Francisco: Freeman.
Lipset, S. M. (1986). Historical traditions and national characteristics: A comparative analysis of Canada and the United States. Canadian Journal of Sociology, 11, 113–155.
Little, R. J. A. (2004). To model or not to model? Competing modes of inference for finite population sampling. Journal of the American Statistical Association, 99, 546–556.
Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data. Journal of the American Statistical Association, 88, 125–134.
Little, R. J. A. (1988). Missing data adjustments in large surveys. Journal of Business and Economic Statistics, 6, 287–301.
Little, R. J. A. , & Rubin, D. B. (2002). Statistical analysis with missing data (2nd ed.). New York: Wiley.
Little, R. J. A. , & Rubin, D. B. (1987). Statistical analysis with missing data. New York: Wiley.
Lobel, T. E. , & Teiber, A. (1994). Effect of self-esteem and need for approval on affective and cognitive reactions: Defense and true self-esteem. Personality and Individual Differences, 16, 315–321.
Loftus, E. , Smith, K. D. , Klinger, M. R. , & Fiedler, J. (1991). Memory and mismemory for health events. In J. Tanur (Ed.), Questions and questions (102–137). New York: Russell Sage Foundation.

Lohr, S. L. (1999). Sampling: Design and analysis. Pacific Grove, CA: Duxbury Press.
Lohr, S. L. , & Rao, J. N. K. (2000). Inference in dual frame surveys. Journal of the American Statistical Association, 95, 271–280.
Lozar Manfreda, K. (1999, December). Participation in web surveys. Paper presented at 9th international meeting dissertation research in psychometrics and sociometrics. Oegstgeest, The Netherlands.
Lozar Manfreda, K. , Bosjnak, M. , Berzelak, J. , Haas. I. , Vehovar, V. (in press). Web surveys versus other survey modes: A meta-analysis comparing response rates. International Journal of Market Research.
Lozar Manfreda, K. , Vehovar, V. , & Batagelj, Z. (2001). Web versus mail questionnaire for an institutional survey. In A. Westlake , W. Sykes , T. Manners , & M. Rigg . (Eds.), The Challenge of the Internet: Proceedings of the 2nd ASC International Conference on Survey Research Methods. Berkeley, UK: Association for Survey Computing.
Luppes, M. (1995). A content analysis of advance letters from expenditure surveys of seven countries. Journal of Official Statistics, 11, 461–480.
Lyberg, L. E. (2006, August). Quality assurance in comparative studies. Paper presented at the 2006 Joint Statistical Meetings, Seattle, VA.
Lyberg, L. E. , Biemer, P. P. , Collins, M. , De Leeuw, E. D. , Dippo, C. , Schwarz, N. , & Trewin, D. (Eds) (1997). Survey measurement and process quality. New York: Wiley.
Lyberg, L. E. , & Couper, M. (2005, April). The use of paradata in survey research. Invited paper presented at the ISI meeting in Sydney, Australia.
Lyberg, L. E. , & Dean, P. (1992, May). Methods of reducing nonresponse rates: A review. Paper presented at the American Association for Public Opinion Research Conference (AAPOR), St. Petersburg, FL.
Lynn, P. (2003). Developing quality standards for cross-national survey research: Five approaches. International Journal of Social Research Methodology, 6, 323–336.
Lynn, P. (2003). PEDAKSI: Methodology for collecting data about survey non-respondents. Quality and Quantity, 37, 239–262.
Lynn, P. (1996). Weighting for nonresponse. In Association for Statistical Computing, Survey and Statistical Computing 1996. Berkeley, U.K.: Association for Statistical Computing.
Lynn, P. , Beerten, R. , Laiho, J. , & Martin, J. (2002). Towards standardization of survey outcome categories and response rate calculations. Research in Official Statistics, 9, 61–85.
Lynn, P. , Beerten, R. , Laiho, J. , & Martin, J. (2001). Recommended standard final outcome categories and standard definitions of response rate for social surveys (Working paper of the Institute for Social and Economic Research, 2001–23). Colchester: University of Essex. Retrieved from: http://www.iser.essex.ac.uk/pubs/workpaps/pdf/2001-23.pdf.
Lynn, P. , Häder, S. , Gabler, S. , & Laaksonen, S. (2007). Methods for achieving equivalence of samples in cross-national surveys. Journal of Official Statistics, 23, 107–124.
Lynn, P. , Lyberg, L. , & Japec, L. (2006). What's so special about cross-national surveys? In J. A. Harkness (Ed.), ZUMA-Nachrichten Spezial No. 12. Conducting cross-national and cross-cultural surveys. Mannheim: ZUMA
Macer, T. (2003). We seek them here, we seek them there. How technological innovation in mixed mode survey software is responding to the challenges of finding elusive respondents. In R. Banks (Ed.), Survey and statistical computing IV. The impact of technology on the survey process. Association for Survey Computing. See also: Macer, T. (2004). CAWI and CATI: Software reviews. In *Quirks Marketing Review*. Retrieved April 15, 2005 , from: www.meaning.uk.com.quirks/09.html.
Macer, T. (2001). Net development. Research, the Magazine of the Market Research Society, 422. Retrieved August 2006 , from: http://www.meaning.uk.com/arts/55.html.
Maddala, G. S. (1983) Limited dependent and qualitative variables in econometrics. Cambridge, MA: Cambridge University Press.
Madow, W. G. , Nisselson, H. , & Olkin, I. (1983). Incomplete data in sample surveys, volume 1: Report and case studies. New York: Academic Press.
Madow, W. G. & Olkin, I. (1983). Incomplete data in sample surveys, volume 3: Proceedings of the symposium. New York: Academic Press.
Madow, W. G. , Olkin, I. , & Rubin, D. B. (1983). Incomplete data in sample surveys, volume 2: Theory and bibliographies. New York: Academic Press.

Malhotra, Y. (2001). Expert systems for knowledge management: Crossing the chasm between information processing and sense making. Expert Systems with Applications, 20, 7–16.
Mangione, T. W. , Fowler, F. J. , & Louis, T. A. (1992). Question characteristics and interviewer effects. Journal of Official Statistics, 8, 293–307.
Marquis, K. H. , & Cannell, C. F. (1969). A study of interviewer-respondent interaction in the urban employment survey. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan.
Marshall G. , Rose D. , Vogler C. , & Newby H. (1988). Social class in modern Britain. London: Hutchinson.
Mason, R. , Carlson, J. , & Tourangeau, R. (1994). Contrast effects and subtraction in part-whole questions. Public Opinion Quarterly, 58, 569–578.
Massey, J. T. (1995). Estimating the response rate in a telephone survey with screening. Proceedings of the section on survey research methods, 18, 673–677.
Massey, J. T. , & Botman, S. (1988). Weighting adjustments for random digit dialed surveys. In R. Groves , P. P. Biemer , L. E. Lyberg , J. T. Massey , W. L. Nichols II , & J. Waksberg (Eds.), Telephone survey methodology (143–160). New York: Wiley.
Mathiowetz, N. A. , & McGonagle, K. A. (2000). An assessment of the current state of dependent interviewing. Journal of Official Statistics, 16, 401–418.
Matsuo, H. , McIntyre, K. P. , Tomazic, T. , & Katz, B. (2004). The online survey: Its contributions and potential problems. ASA Proceedings of the Section on Survey Rersearch Methods, 27, 3998–4000.
Mayer, C. E. (2004). In 1 year, do-not-call list passes 62 million: Complaints about telemarketers pile up. Washington Post, June 24, E05. Retrieved November 14, 2005 , from: http://www.washingtonpost.com/wp-dyn/articles/A767-2004Jun23.html.
Mayer, T. S. , & O'Brien, E. (2001). Interviewer refusal aversion training to increase survey participation. ASA Proceedings of the Annual Meeting of the American Statistical Association. Arlington, VA: ASA
McConaghy, M. , & Carey, S (2004, May). Training to help interviewers avoid refusals: Results of stage 1 of a pilot using Avoiding Refusal Training (ART) with interviewers at the Office for National Statistics, UK. Paper presented at the 2004 Annual Meeting of the American Association for Public Opinion Research, Phoenix, AZ.
McKenna, S. P. , Doward, L. C. (2005). The translation and cultural adaptation of patient-reported outcome measures. Value in Health, 8(2), 89–104.
McLachlan, G. J. , & Krishnan, T. (1997). The EM algorithm and extensions. New York: Wiley.
McMahon, S. R. , Iwamoto, M. , Massoudi, M. S. , Yusuf, H. R. , Stevenson, J. M. , Davod, F. , Chu, S. Y. , & Pickering, L. K. (2003). Comparison of e-mail, fax, and postal surveys of pediatricians. Pediatrics, 111, 299–303.
Medical Outcomes Trust Bulletin . (1997). Approaches to instrument translation: Issues to consider. Medical Outcomes Trust Bulletin, 5(4). Retrieved June 25, 2006 from: http://www.outcomes-trust.org/bulletin/0797blltn.htm.
Mehta, R. , & Sivadas, E. (1995). Comparing response rates and response content in mail versus electronic mail surveys. Journal of the Market Research Society, 37, 429–439.
Mejer, L. (2003). Harmonization of socio-economic variables in EU statistics. In J. H. P. Hoffmeyer-Zlotnik , & C. Wolf (Eds.), Advances in cross-national comparison. A European working book for demographic and socio-economic variables. New York: Kluwer Academic / Plenum Publishers.
Meng, X. L. (1994). Multiple-imputation inferences with uncongenial sources of input (with discussion). Statistical Science, 9, 538–573.
Menon, G. (1994). Judgments of behavioral frequencies: Memory search and retrieval strategies. In N. Schwarz & S. Sudman S. , (Eds.), Autobiographical memory and the validity of retrospective reports. New York: Springer Verlag.
Menon, G. , Rhagubir, P. , & Schwarz, N. (1995). Behavioral frequency judgments: An accessibility-diagnosticity framework. Journal of Consumer Research, 22, 212–228.
Milgram, S. (1963). Behavioral study of obedience. Journal of Abnormal and Social Psychology, 67,: 371–378.

Miller, K. (2003). Conducting cognitive interviews to understand question-response limitations among poorer and less educated respondents. American Journal of Health Behavior, 27, 264–272.
Mingay, D. J. , & Kim, R. (1998). Using cognitive and traditional methods to develop an automated telephone administered questionnaire. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, St. Louis, MO.
Mingay, D. M. (2000). Is telephone audio computer-assisted self-interviewing (T-ACASI) a method whose time has come? Proceedings of the section on survey research methods, 23, 1062–1067.
Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. Psychometrika, 56, 177–196.
Mislevy, R. J. , Johnson, E. G. , & Muraki, E. (1992). Scaling procedures in NAEP. Journal of Educational Statistics, 17, 131–154.
Mohler, P. , & Uher, R. (2003). Documenting comparative surveys for secondary analysis. In J. A. Harkness , F. J. R. van de Vijver & P. Mohler (Eds.), Cross-cultural survey methods. Hoboken, NJ: Wiley.
Molitor, F. , Kravitz, R.L. , To, Y.-Y. , & Fink, A. (2001). Methods in survey research: Evidence for the reliability of group administration vs. personal interviews. American Journal of Public Health, 91, 826–827.
Montgomery, D. (2005). Introduction to statistical quality control. New York: Wiley.
Moon, Y. (1998). Impression management in computer-based interviews: The effects of input modality, output modality and distance. Public Opinion Quarterly, 62, 610–622
Moran, J. , & Biemer, P. (2004, December). Quality assurance and quality management for survey research. Proceedings of the Federal Committee on Survey Methodology Conference, Washington, D.C.
Morgan, D. L. (1988). Focus groups as qualitative research. Thousand Oaks, CA: Sage.
Morganstein, D. R. , & Marker, D. A. (1997). Continuous quality improvement in statistical agencies. In L. Lyberg , P. Biemer , M. Collins , E. de Leeuw , C. Dippo , N. Schwarz ., et al. (Eds.), Survey measurement and process quality. New York: Wiley.
Morton-Williams, J. (1979). The use of “verbal interaction coding” for evaluating a questionnaire. Quality and Quantity, 13, 59–75.
Morton-Williams, J. (1993). Interviewer approaches. Aldershot, U.K.: Dartmouth Publishing.
Moskowitz, J. M. (2004). Assessment of cigarette smoking and smoking susceptibility among youth: Telephone computer-assisted self-interviews versus computer-assisted telephone interviews. Public Opinion Quarterly, 68, 565–587.
Mudryk, W. , Bougie, B. , Xiao, P. , & Yeung, A. (2001). Statistical methods in quality control at Statistics Canada. Course Reference Materials, Statistics Canada, Ottawa.
Münnich, R. , & Rässler, S. (2005). PRIMA: A new multiple imputation procedure for binary variables. Journal of Official Statistics, 21, 325–341.
Nathan, G. (2001). Telesurvey methodologies for household surveys—a review and some thoughts for the future? Survey Methodology, 27, 7–31.
National Center for Health Statistics . (1989). Questionnaire design in the cognitive research laboratory, Series 6: Cognition and survey measurement, no 1 (DHHS Publication No. PHS 89–1076). Hyattsville, MD: US Department of Health and Human Services.
National Center for Social Research . (1999). How to improve survey response rates: A guide for interviewers on the doorstep. London, New Delhi: Sage Publications, Thousand Oaks.
National Research Council . (1979). Privacy and confidentiality as factors in survey response. Washington DC: National Academy Press.
National Research Council . (2005). Expanding access to research data: Reconciling risks and opportunities. Washington DC: National Academy.
Nichols, J. D. (1992). Capture-recapture models. Bioscience, 42, 94–102.
Nicholls, W. L. II , Baker, R. P. Martin, J. (1997). The effect of new data collection technology on survey data quality. In: Lyberg, L. , Biemer, P. , Collins, M. , De Leeuw, E. , Trewin, D. , Dippo, C. , & Schwarz, N. (Eds) Survey Measurement and Process Quality. New York: Wiley.
Norenzayan, A. , & Schwarz, N. (1999). Telling what they want to know: Participants tailor causal attributions to researcher?’ interests. European Journal of Social Psychology, 29,

1011–1020.
Norman, D. A. (1982). Learning and memory. New York: W. H. Freeman & Company.
NSDUH (2005). Frequently Asked Questions. Accessed 28 April 2005 at https://nsduhweb.rti.org.
Nunnally, J. C. , & Bernstein, J. H. (1994). Psychometric Theory. New York: McGraw-Hill.
Nyman, A. M. , Roman, A. M. , & Turner, C. F. (2001, May). Comparison of Computer-Assisted Telephone Survey Methodologies: CATI vs. T-ACASI. Paper presented at annual meeting of American Association for Public Opinion Research, Montreal, Quebec, Canada.
O'Brien, E. M. , Mayer, T. S. , Groves, R. M. , & O'Neill, G. E. (2002). Interviewer training to increase survey participation. ASA Proceedings of the Annual Meeting of the American Statistical Association. Arlington, VA: ASA
O'Muirchearteigh, C. (1991). Simple response variance: Estimation and determinants. In P. N. Beimer , R. M. Groves , L. E. Lyberg , N. A. Mathiewetz , & S. Sudman (Eds.), Measurement errors in surveys (pp. 287–310). New York: Wiley.
Oksenberg, L. , & Cannell, C. (1989). New methods for pretesting survey questionnaires. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, St. Petersburg, FL.
Oksenberg, L. , Cannell, C. , & Kalton, G. (1991). New strategies for pretesting survey questions. Journal of Official Statistics, 7, 349–365.
Ongena, Y. (2005). Interviewer and respondent interaction in survey interviews. Amsterdam: Vrije Universiteit.
Ostrom, T. M. , & Upshaw, H. S. (1968). Psychological perspective and attitude change. In A. C. Greenwald , T. C. Brock , & T. M. Ostrom (Eds.), Psychological foundations of attitudes. New York: Academic Press.
Øyen, E. (Ed.). (1990). Comparative methodology: Theory and practice in international social research, London: Sage.
Oyserman, D. , Coon, H. , & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. Psychological Bulletin, 128, 3–73.
Paasche-Orlow, M. K. , Taylor, H. A. , & Brancati, F. L. (2003). Readability standards for informed consent forms as compared with actual readability. New England Journal of Medicine, 348, 721–726.
Pan, Y. , & De la Puente, M. (2005). Census Bureau guideline for the translation of data collection instruments and supporting materials: Documentation on how the guideline was developed (U.S. Census Bureau Research Report Series, Survey Methodology #2005–06). Retrieved June 19, 2006 , from: http://www.census.gov/srd/papers/pdf/rsm2005-06.pdf
Parackel, M. (2003). Internet-based and mail survey: A hybrid probabilistic survey approach. Proceedings of the 2003 Australian Web Conference. Retrieved April 15, 2005 , from: http://ausweb.scu.edu.au/aw03/papers/parackal.html.
Parameswaran, R. , & Yaprak, A. (1987). A cross-national comparison of consumer research measures. Journal of International Business Studies, 18, 35–49.
Park, A. , & Jowell, R. (1997). Consistencies and differences in a cross-national survey. The International Social Survey Programme (1995). Retrieved from: www.za.uni-koeln.de/-data/en/issp/codebooks/s2880app.pdf).
Park, D. C. (1999). Cognitive aging, processing resources, and self-report. In N. Schwarz , D. C. Park , B. Knäuper , & S. Sudman (Eds.), Aging, cognition, and self-reports. Philadelphia: Psychology Press.
Paxson, M. C. , Dillman, D. A. , & Tarnai, J. (1995). Improving response to business mail surveys. In B. G. Cox , D. A. Binder , B. N. Chinappa , A. Christianson , M. J. Colledge , & P. S. Kott (Eds.), Business survey methods. New York: Wiley.
Payne, S. L. (1951). The art of asking questions. Princeton: Princeton University Press.
Peeters, C. F. W. (2005). Measuring politically sensitive behavior. Amsterdam: Vrije Universiteit.
Pepper, S. C. (1981). Problems in the quantification of frequency expressions. In D. W. Fiske (Ed.), Problems with language imprecision: New directions for methodology of social and behavioral science, 9. San Francisco: Jossey-Bass.

Peschar, J. (1982). Quantitative aspects in cross-national comparative research: Problems and issues. In M. Niessen , & J. Peschar (Eds.), International Comparative Research. Oxford: Pergamon Press.
Pew Research Center for the People and the Press . (2004). Polls face growing resistance, but still representative survey experiment shows (Survey Reports). Retrieved November 14, 2005 , from: http://people-press.org/reports/print.php3?PageID=817.
Pfeffermann, D. (1996). The use of sampling weights for survey data analysis. Statistical Methods in Medical Research, 5, 239–261.
Phipps, P. , & Tupek, A. (1991). Assessing measurement errors in a touchtone recognition survey. Survey Methodology, 17, 15–26.
Ponce, N. A. , Lavarreda, S. A. , Yen, W. , Brown, E. R. , DiSogra C. , & Satter, D. (2004). The California health interview 2001. Translation of a major survey for California’s multiethnic population. Public Health Reports, 119, 1–18.
Porst, R. , & Jers, C. (2006). Die ALLBUS-“Gastarbeiter-Frage”. Zur Geschichte eines Standard-Instruments in der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS). Retrieved September 10, 2006 from: http://www.gesis.org/Publikationen/Berichte/ZUMA_Arbeitsberichte/2005.htm.
Porter, S. R. , & Whitcomb, M. E. (2003). The impact of contact type on web survey response rates. Public Opinion Quarterly, 67, 579–589.
Potaka, L. , & Cochrane, S. (2004). Developing bilingual questionnaires: Experiences from New Zealand in the development of the 2001 Maori Language Survey. Journal of Official Statistics, 20, 289–300.
Potter, F. (1990). A study of procedures to identify and trim extreme sampling weights. Proceedings of the Survey Research Methods Section, 13, 225–230.
Potter, F. (1988). Survey of procedures to control extreme sampling weights. Proceedings of the Survey Research Methods Section, 11, 453–458.
Presser, S. (1989). Pretesting: A neglected aspect of survey research. In F. J. Fowler Jr. (Ed.), Conference Proceeding of Health Survey Research Methods (pp. 35–38) (DHHS Pub. No. PHS 89–3447). Washington, DC: National Center for Health Services Research.
Presser, S. , & Blair, J. (1994). Survey pretesting: Do different methods produce different results? Sociological Methodology, 24, 73–104.
Raghunathan, T. E. , & Grizzle, J. E. (1995). A split questionnaire survey design. Journal of the American Statistical Association, 90, 54–63.
Raghunathan, T. E. , Lepkowski, J. M. , van Hoewyk, J. , & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology, 27, 85–95.
Ramos, M. , Sevedi, B. M. , & Sweet, E. M. (1998). Computerized self-administered questionnaires. In M. P. Couper , R. P. Baker , J. Bethlehem , C. Z. F. Clark , J. Martin , W. L. Nicholls II , et al. (Eds.), Computer assisted survey information collection (398–408). New York: Wiley.
Rao, J. N. K. , Wu, C. F. J. , & Yue, K. (1992). Some recent work on resampling methods for complex surveys. Survey Methodology, 18, 209–217.
Rasinski, K. A. , Rasinski, G. B. , Baldwin, A. K. , Yeh, W. , & Lee, L. (1999). Methods of data collection: Perception of risks and losses, and motivation to give truthful answers to sensitive survey questions. Applied Cognitive Psychology, 13, 465–484.
Rässler, S. (2002). Statistical matching: A frequentist theory, practical applications, and alternative Bayesian approaches. Lecture Notes in Statistics, 168. New York: Springer.
Rässler, S. , & Schnell, R. (2004). Multiple imputations for unit nonresponse versus weighting including a comparison with a nonresponse follow-up study. Diskussionspapier der Lehrstühle für Statistik 65/2004, Nürnberg, Germany.
Raudenbush, S. W. & Bryk, A. S. (2002). Hierarchical linear models. Newbury Park, CA: Sage.
Redline, C. , & Dillman, D. A. (2002). The influence of alternative visual designs on respondents’ performance with branching instructions in self-administered questionnaires. In R. Groves , D. Dillman , J. Eltinge , & R. Little (Eds.), Survey nonresponse. New York: Wiley.
Redline, C. , Dillman, D. A. , Dajani, A. , & Scaggs, M. (2003). Improving navigational performance in U.S. Census 2000 by altering the visually administered languages of

branching instructions. Journal of Official Statistics, 19, 403–420.
Reips, U. D. (2002). Context effects in web-surveys. In B. Batinic , U. D. Reips , M. Bosnjak & A. Werner (Eds.), Online social sciences (95–104). Seattle: Hogrefe & Huber.
Rice, S. A. (1929). Contagious bias in the interview. American Journal of Sociology, 35, 420–423.
Richman, W. L. , Kiesler, S. , Weisband, S. , & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires and interviews. Journal of Applied Psychology, 84, 754–775.
Ritchie, J. & Lewis, J. (Eds.). (2003). Qualitative research practice. London: Sage.
Rizzo, L. J. , Brick, M. , & Park, I. (2004). A minimally intrusive method for sampling persons in random digit dial surveys. Public Opinion Quarterly, 68, 267–274.
Robert, C. P. , & Casella, G. (1999). Monte carlo statistical methods. New York: Springer.
Rock, D. A. , Werts, C. E. , Linn, R. L. , & Jöreskog, K. G. (1977). A maximum likelihood solution to the errors in variables and errors in equation models. Multivariate Behavioral Research, 12, 187–197.
Rogler, L. H. (1999). Methodological sources of cultural insensitivity in mental health research. American Psychologist, 54, 424–433.
Rokkan, S. (1962). The development of cross-national comparative research: A review of current problems and possibilities. Social Science Information, 1, 21–38.
Rokkan, S. (1969). Cross-national survey research: Historical, analytical and substantive contexts. In S. Rokkan , S. Verba , J. Viet , & E. Almasy (Eds.), Comparative Survey analysis. Paris: Mouton.
Rokkan, S. , & Szczerba-Likiernik , (1968). Introduction. In S. Rokkan (Ed.), Comparative research across cultures and nations, Paris: Mouton.
Rosén, B. , & Elvers, E. (1999). Quality concept for official statistics. Encyclopedia of Statistical Sciences, Update 3, 621–629. New York: Wiley.
Rosenthal, R. , & Rosnow, R. (1975). The volunteer subject. New York: Wiley.
Ross, M. (1989). The relation of implicit theories to the construction of personal histories. Psychological Review, 96, 341–357.
Rubin, D. B. (1976). Inference and missing data. Biometrika, 63, 581–590.
Rubin, D. B. (1977). Formalizing subjective notions about the effect of nonrespondents in sample surveys. Journal of the American Statistical Association, 72, 538–543.
Rubin, D. B. (1978). Multiple imputation in sample surveys—a phenomenological Bayesian approach to nonresponse. Proceedings of the Section on Survey Research Methods, 1, 20–40.
Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. Journal of Business and Economic Statistics, 4, 87–95.
Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. New York: Wiley.
Rubin, D. B. (1996). Multiple imputation after 18+ years. Journal of the American Statistical Association, 91, 473–489.
Rubin, D. B. (2004). The design of a general and flexible system for handling nonresponse in sample surveys. The American Statistician, 58, 298–302. (Original work published 1977)
Rubin, D. B. , & Schenker, N. (1987). Interval estimation from multiply imputed data: A case study using census agriculture industry codes. Journal of Official Statistics, 3, 375–387.
Rubin, D. B. , & Schenker, N. (1986). Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. Journal of the American Statistical Association, 81, 366–374.
Rubin, D. B. , Stern, H. S. , & Vehovar, V. (1995). Handling `Don't kno?' survey responses: The case of the Slovenian Plebiscite. Journal of the American Statistical Association, 90, 822–828.
Rubin, H. J. , & Rubin, I. S. (1995). Qualitative interviewing: The art of hearing data. Thousand Oaks, CA: Sage.
Rust, K. F. , & Rao, J. N. K. (1996). Variance estimation for complex surveys using replication techniques. Statistical Methods in Medical Research, 5, 283–310.
Ryan, T. P. (2000). Statistical methods for quality improvement. New York: Wiley.

Saeboe, H. V. , Byfuglien, J. , & Johannessen, R. (2003). Quality issues at statistics Norway. Journal of Official Statistics, 19, 287–303.
Salant, P. , & Dillman, D. A. (1994). How to conduct your own survey. New York: Wiley.
Saris, W. (1998). Ten years of interviewing without interviewers: The telepanel. In M. P. Couper , R. P. Baker , J. A. Harkness, Bethlehem , C. Z. F. Clark , J. R. van de Vijver , & P. MohlerMartin , W. L. Nicholls II , et al. (Eds.), Cross-Cultural Survey Methods. Hoboken, NJ: New York: Wiley.
Saris, W. (2003a). Multitrait-multimethod studies. In J. A. Harkness , F. J. R. van de Vijver , & P. Mohler (Eds.), Cross-Cultural Survey Methods. Hoboken, NJ: Wiley.
Saris, W. (2003b). Response function equality. In J. A. Harkness , F. J. R. van de Vijver , & P Mohler (Eds.), Cross-Cultural Survey Methods. Hoboken, NJ: Wiley.
Schaefer, D. R. , & Dillman, D. A. (1998). Development of a standard e-mail methodology: Results of an experiment. Public Opinion Quarterly, 62, 378–397.
Schaefer, N. (1991). Conversation with a purpose or conversation? Interaction in the standardized interview. In P. Biemer , R. Groves , L. Lyberger , N. Mathiowetz , & S. Sudman (Eds.), Measurement errors in surveys (367–393). New York: Wiley
Schafer, J. L. (1997). Analysis of incomplete multivariate data. New York: Chapman and Hall.
Schafer, J. L. , & Schenker, N. (2000). Inference with imputed conditional means. Journal of the American Statistical Association, 95, 144–154.
Scheers, N. J. , & Dayton, M. C. (1988) Covariate randomized response models. Journal of the American Statistical Association, 83, 969–974.
Scheers, N. J. , & Mitchell, C.D. (1988). Covariate randomized response models. Journal of the American Statistical Association 83, 969–974.
Schenker, N. & Taylor, J. M. G. (1996). Partially parametric techniques for multiple imputation. Computational Statistics & Data Analysis, 22, 425–446.
Scherpenzeel, A. C. (1995). A question of quality: Evaluating survey questions by multitrait-multimethod studies. Leidschendam: KPN research.
Scheuch, E. K. (1990). The development of comparative research: Towards causal explanations. In E. Øyen (Ed.), Comparative methodology: Theory and practice in international social research. Wiltshire, U.K.: Sage.
Scheuren, F. (2004). What is a survey? Alexandria, VA: American Statistical Association. Retrieved July 24, 2006 , from: http://www.whatisasurvey.info.
Scheuren, F. (2001). How important is accuracy? Proceedings of Statistics Canada Symposium, Statistics Canada, Ottawa, Canada.
Schillewaert, N. , Langerak, F. , & Duhamel, T. (1998). Non probability sampling for WWW surveys: A comparison of methods. Journal of the Market Research Society, 40, 307–313.
Schober, M. F. , & Conrad, F. G. (2002). A collaborative view of standardized survey interviews. In D. Maynard , H. Houtkoop-Steenstra , N. Schaeffer , & J. van der Zouwen (Eds.), Standardization and tacit knowledge (67–94). New York: Wiley.
Schober, M. F. , & Conrad, F. G. (1998). Does conversational interviewing improve survey data quality beyond the laboratory? ASA Proceedings of the Annual Meeting of the American Statistical Association. Arlington, VA: ASA
Schober, M. F. , & Conrad, F. G. (1997). Does conversational interviewing reduce survey measurement error? Public Opinion Quarterly, 61, 576–602.
Schonlau, M. , Van Soest, A. , Kapteyn, A. , Couper, M. P. , & Winter, J. (2004). Attempting to adjust for selection bias in web surveys with propensity scores: The case of the health and retirement survey (HRS). Proceedings of the section on survey research methods, 24.
Schuman, H. , & Converse, J. (1971). The effect of black and white interviewers on black responses in 1968. Public Opinion Quarterly, 35, 44–68.
Schuman, H. , & Ludwig, J. (1983). The norm of even-handedness in surveys as in life. American Sociological Review, 48, 112–120.
Schuman, H. , & Presser, S. (1981). Questions and answers in attitude surveys: Experiments in question form, wording, and context. New York: Academic Press.
Schulz, W. , & Sibberns, H. (Eds.). (2004). IEA Civic Education Study technical report. Amsterdam: IEA.

Schwarz, N. (1996). Cognition and communication: Judgmental biases, research methods, and the logic of conversation. Hillsdale: Erlbaum.
Schwarz, N. (1997). Questionnaire design: The rocky road from concepts to answers. In: L. Lyberg , P. Biemer , M. Collins , E. de Leeuw , C. Dippo , N. Schwarz , et al. (Eds.), Survey measurement and process quality (29–45). New York: Wiley.
Schwarz, N. (1999a). Self-reports: How the questions shape the answers. American Psychologist, 54, 93–105.
Schwarz, N. (1999b). Frequency reports of physical symptoms and health behaviors: How the questionnaire determines the results. In D. C. Park , R. W. Morrell & K. Shifren (Eds.), Processing of medical information in aging patients: Cognitive and human factors perspectives. Mahaw: Erlbaum.
Schwarz, N. (2003). Culture-sensitive context effects: A challenge for cross-cultural surveys. In J. A. Harkness , F. J. R. van de Vijver , & P. Mohler (Eds.), Cross-cultural survey methods. Hoboken, NJ: Wiley.
Schwarz, N. , & Bless, H. (1992a). Constructing reality and its alternatives: Assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), The construction of social judgments. Hillsdale: Erlbaum.
Schwarz, N. , & Bless, H. (1992b). Scandals and the public's trust in politicians: Assimilation and contrast effects. Personality and Social Psychology Bulletin, 18, 574–579.
Schwarz, N. , & Bless, H. (2007). Mental construal processes: The inclusion/exclusion model. In D. A. Stapel and J. Suls (Eds.), Assimilation and contrast in social psychology. Philadelphia, PA: Psychology Press.
Schwarz, N. , & Hippler, H. J. (1995). Subsequent questions may influence answers to preceding questions in mail surveys. Public Opinion Quarterly, 59, 93–97.
Schwarz, N. , & Hippler, H. J. (1991). Response alternatives: The impact of their choice and ordering. In P. Biemer , R. Groves , N. Mathiowetz & S. Sudman (Eds.), Measurement error in surveys. New York: Wiley.
Schwarz, N. , Hippler, H. J. , Deutsch, B. , & Strack, F. (1985). Response scales: Effects of category range on reported behavior and subsequent judgments. Public Opinion Quarterly, 49, 388–395.
Schwarz, N. , Knäuper, B. , Hippler, H. J. , Noelle-Neumann, E. , & Clark, F. (1991). Rating scales: Numeric values may change the meaning of scale labels. Public Opinion Quarterly, 55, 570–582.
Schwarz, N. , & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. American Journal of Evaluation, 22, 127–160.
Schwarz, N. , Park, D. C. , Knäuper, B. , Davidson, N. , & Smith, P. (1998, April). Aging, cognition, and self-reports: Age-dependent context effects and misleading conclusions about age differences in attitudes and behavior. Paper presented at the Cognitive Aging Conference, Atlanta, GA.
Schwarz, N. , & Scheuring, B. (1992). Selbstberichtete Verhaltens- und Symptomhäufigkeiten: Was Befragte aus Anwortvorgaben des Fragebogens lernen. [Frequency-reports of psychosomatic symptoms: What respondents learn from response alternatives] Zeitschrift für Klinische Psychologie, 22, 197–208.
Schwarz, N. , Strack, F. , & Mai, H. P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. Public Opinion Quarterly, 55, 3–23.
Schwarz, N. , Strack, F. , Müller, G. , & Chassein, B. (1988). The range of response alternatives may determine the meaning of the question: Further evidence on informative functions of response alternatives. Social Cognition, 6, 107–117.
Schwarz, N. , & Sudman, S. (1996). Answering questions: Methodology for determining cognitive and communicative processes in survey research. San Francisco: Jossey-Bass.
Schwarz, N. , & Sudman, S. (1994). Autobiographical memory and the validity of retrospective reports. New York: Springer Verlag.
Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. Journal of Personality and Social Psychology, 51, 515–530.

Seber, G. A. F. (1996). A review of estimating animal abundance II. International Statistical Review, 60, 129–166.
Serpell, R. (1990). Audience, culture and psychological explanation: A reformulation of the emicetic problem in cross-cultural psychology. The quarterly newsletter of the Laboratory of Comparative Human Cognition, 12, 99–132.
Shao, J. (2002). Replication methods for variance estimation in complex sample surveys with imputed data. In R. M. Groves , D. A. Dillman , J. L. Eltinge , & R. J. A. Little (Eds.), Survey nonresponse. New York: Wiley.
Sharma, D. C. (2005). Net telephony set to surge, IDC says. Retrieved November 14, 2005 , from: http://news.com.com/Net+telephony+set+to+surge2C+IDC+says/2100-7352_3-5653437.html.
Sheatsley, P. (1983). Questionnaire construction and item writing. In P. Rossi , J. D. Wright , & A. B. Anderson (Eds.), Handbook of survey research. New York: Academic Press.
Shuttles, C. D. , Welch, J. S. , Hoover, J. G. , & Lavrakas, P. J. (2002, May). The Development and Experimental Testing of an Innovative Approach to Training Telephone Interviewers to Avoid Refusals. Paper presented at the 2002 Annual Meeting of the American Association for Public Opinion Research, St. Petersburg, FL.
SIBIS (2003). Towards the information society in Europe and the US. Bonn, Germany: SIBIS. Retrieved November 14, 2005 , from: http://www.empirica.biz/sibis/statistics/data/1-10.htm.
Sieber, J. E. , & Stanley, B. (1988). Ethical and professional dimensions of socially sensitive research. American Psychologist, 43, 49–55.
Siemiatycki, J. (1979). A comparison of mail, telephone, and home interview strategies for household health surveys. American Journal of Public Health, 69, 238–245.
Singer, E. (1978). Informed consent: Consequences for response rate and response quality in social surveys. American Sociological Review, 43, 144–162.
Singer, E. (2003). Exploring the meaning of consent: Participation in research and beliefs about risks and benefits. Journal of Official Statistics, 19, 273–285.
Sirken, M. , Hermann, D. , Schechter, S. , Schwarz, N. , Tanur, J. , & Tourangeau, R. (Eds.) (1999). Cognition and survey research. New York: Wiley.
Singer, E. , Mathiowetz, N. , & Couper, M. P. (1993). The impact of privacy and confidentiality concerns on census participation. Public Opinion Quarterly, 57, 465–482.
Singer, E. , Van Hoewyk, J. , Gebler, N. , Raghunatan, T. , & McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. Journal of Official Statistics, 15, 217–230.
Singer, E. , Van Hoewijck, J. , & Maher, M. P. (2000). Experiments with incentives in telephone surveys. Public Opinion Quarterly, 64, 171–188.
Singer, E. , Van Hoewyk, J. , & Neugebauer, R. (2003). Attitudes and behavior: The impact of privacy and confidentiality concerns on participation in the 2000 census. Public Opinion Quarterly, 65, 368–384.
Skevington, S. M. (2002). Advancing cross-cultural research on quality of life: Observations drawn from the WHOQOL development. Quality of Life Research, 11, 135–144.
Skevington, S. M. , Sartorious, N. , Amir, M. , & The WHOQOL-Group (2004). Developing methods for assessing quality of life in different cultural settings. The history of the WHOQOL instruments. Social Psychiatry and Psychiatry Epidemiology, 39, 1–8.
Skevington, S. M. , & Tucker, C. (1999). Designing response scales for cross-cultural use: Data from the development of the UK WHOQOL. British Journal of Medical Psychology, 72, 51–61.
Skinner, C. J. , Holt, D. , & Smith, T. M. F. (1989). Analysis of complex surveys. Chichester: Wiley.
Skinner, H. A. , & Allen, B. A. (1983). Does the computer make a difference? Computer versus face-to-face, versus self-report assessment of alcohol, drugs, and tobacco use. Journal of Consulting and Clinical Psychology, 51, 267–275.
Skjåk, K. K. , & Harkness, J. (2003). Data collection methods. In J. A. Harkness , F. J. R. van de Vijver , & P. Mohler (Eds.), Cross-cultural survey methods. New York: Wiley.
Smid, M. , & Hess, D. (2003). Harmonizing sampling frames and indicators in international market research: A German perspective. In J. H. P. Hoffmeyer-Zlotnik , & C. Wolf (Eds.), Advances in cross-national comparison. A European working book for demographic and

socio-economic variables. New York: Kluwer Academic/Plenum Publishers.
Smith, A. , Christopher, S. , & McCormick, A. K. (2004). Development and implementation of a culturally sensitive cervical health survey: A community-based participatory approach. Women and Health, 40, 67–86.
Smith, P. J. , Hoaglin, D. C. , & Battaglia, M. P. (2005). Statistical methodology of the national immunization survey, 1994–2002. Vital Health Statistics 2, 138. National Center for Health Statistics. Retrieved November 14, 2005 , from:
http://www.cdc.gov/nchs/data/series/sr_02/sr02_138.pdf.
Smith, T. W. (1995, May). Little Things Matter: How differences in questionnaire format can affect survey responses. Paper presented at Annual Conference of the American Association for Public Opinion Research, Fort Lauderdale, FL.
Smith, T. W. (2003). Developing comparable questions in cross-national surveys. In J. A. Harkness , F. J. R. van de Vijver , & P. Mohler (Eds.), Cross-cultural survey methods. Hoboken, NJ: Wiley.
Smith, T. W. (2004). Developing and evaluating cross-national survey instruments. In S. Presser , J. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , & E. Singer (Eds.), Methods for testing and evaluating survey questionnaires. Hoboken, NJ: Wiley.
Smyth, J. D. , Dillman, D. A. , & Christian, L.M. (2006, January). Does “yes or no” on the telephone mean the same as “check-all-that-apply” on the web? Paper presented at Telephone Survey Methodology II, Miami, FL.
Smyth, J. D. , Dillman, D. A. , Christian, L. M. , & Stern, M. F. (2006). Comparing check-all and forced-choice question formats in web surveys. Public Opinion Quarterly 70, 66–77.
Snell-Hornby, M. (1988). Translation studies: An integrated approach. Amsterdam: John Benjamins.
Snijkers, G. (2002). Cognitive laboratory experiences on pre-testing computerised questionnaires and data quality. Heerlen: Statistics Netherlands.
Snijkers, G. , Hox, J. , & De Leeuw, E. (1999). Interviewer?’ Tactics for Fighting Survey Nonresponse. Journal of Official Statistics, 15, 185–198. (Also on www.jos.nu)
Social and Economic Sciences Research Center . (2005). Instructing web and telephone respondents to report date answers in format desired by the surveyor (Technical Report 05–067). Washington, DC: Christian, L. M., Dillman, D. A., & Smyth, J. D.
Somers, H. (2005, December). Round-trip translation: What is it good for? Proceedings of the Australasian Language Technology Workshop 2005, Sydney, Australia. Retrieved September 6, 2006 from:
http://www.alta.asn.au/events/altw2005/cdrom/pdf/ALTA200519.pdf#search=%22Round-Trip%20Translation%3A%20What%20Is%20It%20Good%20For%3F%22
Sperber, A. D. , Devellis, R. F. , & Boehlecke, B. (1994). Cross-cultural translation: Methodology and validation. Journal of Cross-cultural Psychology, 25, 501–524.
Srinivasan, R. & Hanway, S. (1999). A new kind of survey mode difference: Experimental results from a test of inbound voice recognition and mail surveys. Paper presented at the meeting of the American Association of Public Opinion Research, St. Pete Beach, FL.
Statistics Canada . (2003). Survey methods and practices. Ottawa: Ministry of Industry, Statistics Canada, Social Survey Division.
Statistics Sweden (2001). The future development of the Swedish register system. (R&D Report 2001:1). Stockholm: Statistics Sweden.
Steeh, C. (2004, May). A new era for telephone interviewing. Paper presented at the Fifty-ninth Annual Conference of the American Association for Public Opinion Research, Phoenix, AZ.
Steeh, C. (1981). Trends in nonresponse rates, 1952–1979. Public Opinion Quarterly, 45, 40–57.
Steeh, C. , Kirgis, N. , Cannon, B. , & De Witt, J. (2001). Are they really as bad as they seem? Nonresponse rates at the end of the twentieth century. Journal of Official Statistics, 17, 227–247.
Steeh, C. , & Piekarski, L. (2006, January). Accommodating new technologies: The rejuvenation of telephone surveys. Paper presented at the Second International Conference on Telephone Survey Methodology, Miami, FL.

Stephens, R. T. (2003). Utilizing Metadata as a knowledge communication tool. Proceedings of the International Professional Communication Conference 2004, Minneapolis, MN. Institute of Electrical and Electronics Engineers, Inc.
Stern, M. F. , & Dillman, D. A. (2006). Community participation, social ties, and use of the internet. City and Community. 5, 409–424.
Stewart, D. W. , & Shamdasani, P. N. (1990). Focus groups: Theory and practice. Thousand Oaks, CA: Sage.
Stouffer, S. A. (1963). Communism, conformity, and civil liberties. Gloucester, Mass: Peter Smith.
Strack, F. (1994). Response processes in social judgment. In R. S. Wyer , & T. K. Srull (Eds.), Handbook of social cognition, volume 1 (2nd ed.). Hillsdale: Erlbaum.
Strack, F. , & Martin, L. (1987). Thinking, judging, and communicating: A process account of context effects in attitude surveys. In H. J. Hippler , N. Schwarz , & S. Sudman (Eds.), Social information processing and survey methodology. New York: Springer Verlag.
Strack, F. , Schwarz, N. , & Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, mood, and mode of thinking. Journal of Personality and Social Psychology, 49, 1460–1469.
Strack, F. , Schwarz, N. , & Wänke, M. (1991). Semantic and pragmatic aspects of context effects in social and psychological research. Social Cognition, 9, 111–125.
Strube, G. (1987). Answering survey questions: The role of memory. In H. J. Hippler , N. Schwarz , & S. Sudman (Eds.), Social information processing and survey methodology. New York: Springer Verlag.
Sturgeon, K. , & Winter, S. (1999). International marketing on the world wide web. New opportunities for research: What works, what does not and what is next. Proceedings of the ESOMAR Worldwide Internet Conference Net Effects, 191–200.
Subcommittee on Disclosure Limitation Methodology (1994). Federal committee on statistical methodology (Statistical policy working paper No. 22). Washington DC: Statistical Policy Office, Office of Information and Regulatory Affairs, OMB.
Sudman, S. (1983). Applied sampling. In P. Rossi , J. D. Wright , & A. B. Anderson (Eds.), Handbook of survey research. New York: Academic Press.
Sudman, S. , & Bradburn, N. M. (1974). Response effects in surveys. Chicago: Aldine.
Sudman, S. , & Bradburn, N. M. (1982). Asking questions: a practical guide to questionnaire design. San Francisco: Jossey Bass.
Sudman, S. , Bradburn, N. M. , & Schwarz, N. (1996). Thinking about answers: The application of cognitive processes to survey methodology. San Francisco: Jossey-Bass.
Sudman, S. , & Schwarz, N. (1989). Contributions of cognitive psychology to advertising research. Journal of Advertising Research, 29, 43–53.
Sundgren, B. (1993). Guidelines on the design and implementation of statistical metainformation systems. R&D Report 4. Stockholm: Statistics Sweden.
Sundgren, B. (1973). An infological approach to data bases. Unpublished doctoral thesis, University of Stockholm, Sweden.
Supple, A. J. , Aquilino W. S. , & Wright, D. L. (1999). Collecting sensitive self-report data with laptop computers: Impact on the response tendencies of adolescents in a home interview. Journal of Research on Adolescence, 9, 467–488.
Sykes, W. , & Collins, M. (1992). Anatomy of the survey interview. Journal of Official Statistics, 8, 277–291.
Tafforeau, J. , López Cobo, M. , Tolonen, H. , Scheidt-Nave, C. , & Tinto, A. (2005). Guidelines for the development and criteria for the adoption of health survey instruments. European Commission. Retrieved June 23, 2006 , from: http://forum.europa.eu.int/Public/irc/dsis/-health/library?l=/reports/healthsinterviewssurvey/guidelines_instruments/_EN_1.0_&a=d
Tanner, M. A. , & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. Journal of the American Statistical Association, 82, 528–540.
Tanzer, N. K. (2005). Developing tests for use in multiple languages and cultures: A plea for simultaneous development. In R. K. Hambleton , P. F. Merenda , & C. D. Spielberger (Eds.), Adapting educational and psychological tests for cross-cultural assessment. Mahwah, NJ: Erlbaum.

Teune, H. (1990). Comparing countries: Lessons learned. In E. Øyen (Ed.), Comparative methodology: Theory and practice in international social research. London: Sage.
Thomas, E. L. , & Robinson, H. A. (1972). Improving reading in everyday class: A sourcebook for teachers. Boston: Allyn and Bacon.
Thomas, R. (2002). Material prepared by R. Thomas for Center for Applied Social Surveys (CASS) course 2002 on "Pretesting Survey Questionnaires" conducted by Campanelli, Collins, & Thomas. Southampton, UK: University of Southampton, Southampton Statistical Science Research Institute, Center for Applied Social Surveys.
Thompson, S. K. (2002). Sampling (2nd ed.). New York: Wiley Interscience.
Thomson, K. , Nicolaas, G. , Bromley, C. , & Park, A. (2001). Welsh assembly election study 1999: Technical report. London: National Center for Social Research.
Thornberry, J. , Bhaskar, B. , Wesley, B. , Krulewitch, C.J. , Wesley, B. , Hubbard, M.L. , Das, A. , Foudin, L. , Adamson, M. 2002. Audio computerized self-report interview use in prenatal clinics: Audio computer assisted self interview with touch screen to detect alcohol consumption in pregnant women application of a new technology to an old problem. Computer Information Nursing 20: 46–52.
Tjøstheim, I. (2005, April). Mobile self-interviewing an opportunity for location-based market research. Are privacy concerns a showstopper? Paper presented at the Association of Survey Computing Conference on Mobile Computing, London, UK.
Tjøstheim, I. , & Thalberg, S. (2005). Are the mobile phone users ready for MCASI - mobile computer assisted self-interviewing? In Excellence in international research 2005. ESOMAR.
Tomas, R.K. (2004).
Tortora, R. (2004). Response trends in a national random digit dial survey. Advances in Methodology and Statistics, 1, 21–32.
Tourangeau, R. (1984). Cognitive science and survey methods: A cognitive perspective. In T. Jabine , M. Straf , J. Tanur , & R. Tourangeau (Eds.), Cognitive aspects of survey methodology: Building a bridge between disciplines. Washington, DC: National Academy Press.
Tourangeau, R. , Couper, M.P. , & Steiger, D.M. (2001). Social presence in web surveys. Federal Committee on Statistical Methodology conference, 2001. Retrieved August 2007 from http://www.websm.org.
Tourangeau, R. , Rasinski, K. , Jobe, J. , Smith, T. , & Pratt, W. (1997). Sources of error in a survey of sexual behaviour, Journal of Official Statistics, 13, 341–365.
Tourangeau, R. , Rips, L. J. , & Rasinski, K. (2000). The psychology of survey response. Cambridge: Cambridge University Press.
Tourangeau, R. , & Smith, T. W. (1998). Collecting sensitive data with different modes of data collection. In M. P. Couper , R. P. Baker , J. Betlehem , C. Z. F. Clark , J. Martin , W. L. Nichols II , et al. (Eds.), Computer assisted survey information collection (431–453). New York: Wiley.
Tourangeau, R. , & Smith, T. W. (1996). Asking sensitive questions. The impact of data collection, mode, question format, and question context. Public Opinion Quarterly, 60, 275–304.
Tourangeau, R. , Steiger, D. M. & Wilson, D. (2002). Self-administered questions by telephone: Evaluating Interactive Voice Response. Public Opinion Quarterly, 66, 265–278.
Traub, R. E. (1994). Reliability for the Social Sciences. Thousand Oaks, CA: Sage.
Triandis, H. C. (1972). The analysis of subjective culture. Oxford, UK: Wiley-Interscience.
Trice, A. (1987). Informed consent: Biasing of sensitive self-report data by both consent and information. Journal of Social Behavior and Personality, 2, 369–374.
Trussell, N. , & Lavrakas, P. J. (2004). The influence of incremental increases in token cash incentives on mail survey response: Is there an optimal amount? Public Opinion Quarterly, 68, 349–367.
Tsuchiya, T. , Hiray, Y. , & Ono, S. (2007). A study of the properties of the item count technique. Public Opinion Quarterly, 71, 253–272.
Tuckel, P. , & O'Neill, H. (2002). The vanishing respondent in telephone surveys. Journal of Advertising Research, 42, 26–48.
Tucker, C. J. , Brick, M. , Meekins, B. , & Esposito, J. (2005). Household telephone service and usage patterns in the U.S. in 2004: Implications for telephone samples. Bureau of Labor

Statistics: Unpublished report.
Tucker, C. , Casady, R. , & Lepkowski, J. (1992). Sample allocation for stratified telephone sample designs. Proceedings of the section on survey research methods, 15, 291–296.
Tucker, C. , Lepkowski, J. M. , & Piekarski, L. (2002). The current efficiency of list-assisted telephone sampling designs. Public Opinion Quarterly, 66, 321–328.
Turner, C. F. , Forsyth, B. H. , O'Reilly, J. M. , Cooley, P. C. , Smith, T. K. , Rogers, S. M. , & Miller, H. G. (1998). Automated self-interviewing and the survey measurement of sensitive behaviors. In M. P. Couper , R. P. Baker , J. Betlehem , C. Z. F. Clark , J. Martin , W. L. Nichols II , et al. (Eds.), Computer assisted survey information collection. New York: Wiley.
Turner, C. F. , Miller, H. G. , Smith, T. K. , Cooley, P. C. , & Rogers, S. M. (1996). Telephone Audio Computer-Assisted Self-Interviewing (T-ACASI) and survey measurements of sensitive behaviors: Preliminary results. In R. Banks , J. Fairgrieve , L. Gerrard , T. Orchard , C. Payne , & A. Westlake (Eds.), Survey and statistical computing 1996: Proceedings of the seconds ASC International Conference. Chesham, Bucks, U.K.: Association for Survey Computing.
Tuskegee Syphilis Study Ad Hoc Advisory Panel . (1973). Final report. Washington, DC: US Department of Health, Education, and Welfare.
Tuten, T. L. , Bosnjak, M. , & Bandilla, W. (1999/2000). Banner-advertised web-surveys. Marketing Research, 11, 16–21.
U.S. Energy Information Administration . (1996). Residential energy consumption survey quality profile. Washington, DC: U.S. Department of Energy.
U.S. Federal Committee on Statistical Methodology . (2001). Measuring and reporting the quality of survey data (Statistical Policy Working Paper 31). Washington, D.C.: U.S. Office of Management and Budget.
Van Buuren, S. , Brand, J. P. L. , Oudshoorn, C. G. M. , & Rubin, D. B. (in press). Fully conditional specification in multivariate imputation. Journal of Statistical Computation and Simulation, 76, 1049–1064.
Van Buuren, S. , & Oudshoorn, C. G. M. (2000). Multivariate imputation by chained equations: MICE V1.0 User's Manual (Report PG/VGZ/00.038). Leiden, The Netherlands: TNO Preventie en Gezondheid.
Van de Vijver, F. J. R. (2003). Bias and equivalence: Cross cultural perspectives. In J. A. Harkness , F. J. R. van de Vijver , & P. Mohler (Eds.), Cross-cultural survey methods, Hoboken, NJ: Wiley.
Van der Heijden, P. G. M. , van Gils, G. , Bouts, J. , & Hox, J. J. (2000). A comparison of randomized response, computer-assisted self-interview and face-to-face direct questioning. Sociological Methods and Research, 28, 505–537.
Van der Stadt, H. , Ten Cate, A. , Hundepool, A. J. & Keller, W. J. (1986). Koopkracht in kaart gebracht: Een statistiek van de inkomensdynamiek. Statistische Onderzoekingen, M28. Voorburg: Statistics Netherlands.
Van der Zouwen, J. , & Smit, J. H. (2004). Evaluating survey questions by analysing patterns of behaviour codes and question-answer sequences: A diagnostic approach. In S. Presser , J. M. Rothgeb , M. P. Couper , J. T. Lessler , E. Martin , J. Martin , et al. (Eds.), Methods for testing and evaluating survey questionnaires. New York: Wiley.
Van Deth, J. W. (1998). Equivalence in comparative political research. In J. W. van Deth (Ed.), Comparative politic: The problem of equivalences. London: Routledge.
Van Hattum, M. J. C. & de Leeuw, E. D. (1999). A disk-by-mail survey of pupils in primary schools: data quality and logistics. Journal of Official Statistics, 15, 413–429. (also at www.jos.nu)
Vartivarian S. , & Little, R. (2002). On the formation of weighting adjustment cells for unit nonresponse. Proceedings of the Survey Research Methods Section, 15.
Vehovar, V. , Batagelj, Z. , Lozar Mnafreda, K. , & Zaletel, M. (2002). Nonresponse in web surveys. In R. M. Groves , D. A. Dillman , J. L. Eltinge , R. J. A. Little (Eds), Survey Nonresponse. New York: Wiley.
Vehovar, V. , Belak, E. , Batagelj, Z. , & Čikić, S. (2003). Mobile phone surveys in Slovenia. Unpublished manuscript.
Vehovar, V. , Lozar Manfreda, K. , & Batagelj, Z. (2000, May). Design issues in WWW surveys. Paper presented at the 55th Annual Conference of American Association for Public Opinion Research, Portland, OR.

Verba, S. (1971). Cross-national survey research: The problem of credibility. In I. Vallier (Ed.), Comparative methodology on sociology: Essays on trends and applications. Berkeley: University of California Press.
Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. Cognitive Psychology, 18, 225–252.
Wänke, M. , & Schwarz, N. (1997). Reducing question order effects: The operation of buffer items. In L. Lyberg , P. Biemer , M. Collins , E. de Leeuw , C. Dippo & N. Schwarz (Eds.), Survey measurement and process quality. Chichester, UK: Wiley.
Ware, J. (1987). Standards for validating health measures: Definition and content. Journal of Chronic Diseases, 40, 473–480.
Warner, S.L. (1965). Randomized response: A technique for eliminating evasive response bias. Journal of the American Statistical Association, 60, 63–69.
Warren, R. D. , White, J. K. , & Fuller, W. A. (1974). An errors in variables analysis of managerial role performance. Journal of the American Statistical Association, 69, 886–893.
Weeks, M. F. (1992). Computer-Assisted Survey Information Collection: A review of CASIC methods and their implications for survey operations. Journal of Official Statistics, 8, 445–465.
Weir, P. , Laurence, M. , & Blessing, C. (2000). A comparison of the use of telephone interview to telephone audio CASI in a customer satisfaction survey. Proceedings of the Survey Research Methods Section, 23, 828–833.
Weisband, S. , & Kiesler, S. (1996). Self-disclosure on computer forms: Meta-analysis and implications. Tucson AR: University of Arizona. Retrieved from: www.al.arizona.ude/~weisbrand/chi/chi96.html.
Werking G. , & Clayton, R. (1993). Automated telephone collection techniques for improving timeliness and data quality. Washington DC: Bureau of labour Statistics.
Whitcomb, M. E. , & Porter, S. R. (2004). E-mail contacts: A test of complex graphical designs in survey research. Social Science Computer Review, 22, 370–376.
White, G. K. (1996). An on-line survey of food and beverage consumers on the Internet: An evaluation of the survey methodology. Marketing Research On-line, 1, 39–59. Retrieved November 11, 2001 , from: http://mro.massey.ac.nz/online3.pdf
Widman, L. , & Vogelius, L. (2002, June). Daily reach using SMS: Measuring and reporting reach in real time. Paper presented at the Week of Audience Measurement (WAM), Cannes, France.
Wilkins, J. R. , Hueston, W. D. , Crawford, J. M. , Steele, L. L. , & Gerken, D. F. (1997). Mixed-mode survey of female veterinarians yields high response rate. Occupational Medicine, 47, 458–462.
Wilkinson, R. K. , & Hines, C. J. (1991). Data collection methods in the STPDS Surveys: Improving response rates in mixed mode surveys. Retrieved April 15, 2005 , from: http://srsstats.sbe.nsf.gov/research/3-91.pdf.
Williams, M. D. , & Hollan, J. D. (1981). The process of retrieval from very long term memory. Cognitive Science, 5, 87–119.
Willis, G. (2004). Overview of methods for developing equivalent measures across multiple cultural groups. In S. B. Cohen , & J. M. Lepkowski (Eds.), Eighth Conference on Health Survey Research Methods, Hyattsville, US. Department of Health and Human Services.
Willis, G. (1994). Cognitive interviewing and questionnaire design: A training manual. Washington, DC: Office of Research Methodology, National Center for Health Statistics.
Willis, G. , & Lessler, J. (1999). Questionnaire appraisal system-1999. Research Triangle Park, NC: Research Triangle Institute.
Willke, J. , Adams, C. O. , & Girnius, Z. (1999). Internet testing: A landmark study of the differences between mall intercept and on-line interviewing in the United States. Proceedings of the ESOMAR Worldwide Internet Conference Net Effects, 145–157.
Winkel, F. W. , & Vrij, A. (1998). Who is in need of victim support? The issue of accountable, empirically validated selection and victim referral. Expert Evidence: The International Digest of Human Behaviour, Science and Law, 6, 23–41.
Winkielman, P. , Knäuper, B. , & Schwarz, N. (1998). Looking back at anger: Reference periods change the interpretation of (emotion) frequency questions. Journal of Personality and Social Psychology, 75, 719–728.

Wolter, K. M. (1985). Introduction to variance estimation. New York: Springer-Verlag.
Woodall, G. (1998). Market research on the Internet. Great Falls, VA: Rockbridge Associates, Inc. Retrieved November 11, 2001 , from: http://www.rockresearch.com/Articles/nmr01/nmr01.html
Zaller, J. R. (1992). The nature and origins of mass opinion. Cambridge: Cambridge University Press.
Zarkovich, S. (1966). Quality of statistical data. Rome: Food and Agricultural Organization of the United Nations.
Zdep, S. M. , Rhodes, J. N. , Schwartz, R. M. , & Kilkenny, M. J. (1979). The validity of the Randomized Response Technique. Public Opinion Quarterly, 43, 544–549.
Zelditch, M. Jr. (1971). Intelligible comparisons. In I. Vallier (Ed.), Comparative methods in sociology: Essays on trends and applications. Berkeley: University of California Press.
Zukerberg, A. , Nichols, E. , & Tedesco, H. (1999, May). Designing surveys for the next millennium: Internet questionnaire design issues. Paper presented at the 1999 Association for Public Opinion Research Conference, St. Petersburg, FL.
Zull, C. , Weber, R. , & Mohler, P. (1989). Computer aided text classification for the social ssciences: The general enquirer III. Mannheim, BRD: ZUMA.