Measurement Errors in Surveys
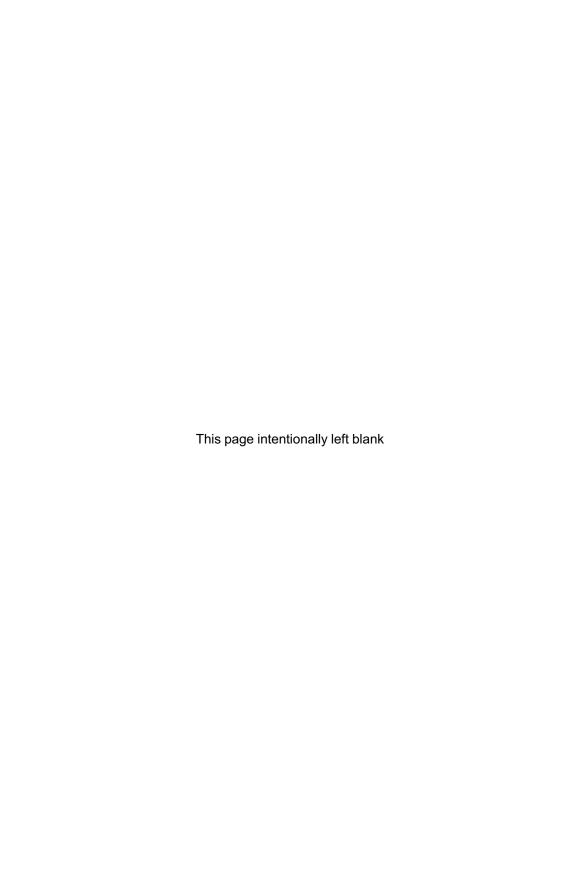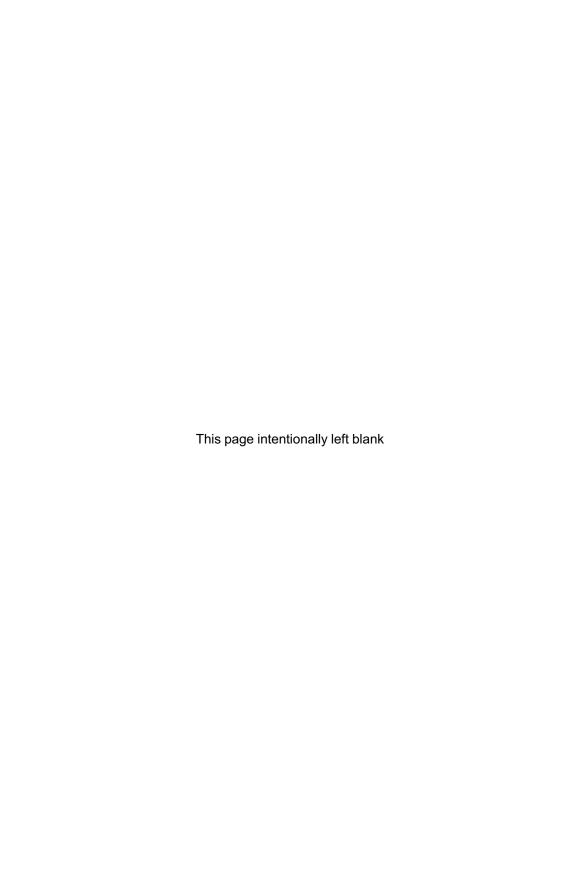
PAUL P. BIEMER

ROBERT M. GROVES

LARS E. LYBERG

NANCY A. MATHIOWETZ

SEYMOUR SUDMAN

This page intentionally left blank

Measurement Errors
in Surveys

This page intentionally left blank

Measurement Errors in Surveys

Edited by

PAUL P. BIEMER
Research Triangle Institute

ROBERT M. GROVES
U.S. Bureau of the Census
The University of Michigan

LARS E. LYBERG
Statistics Sweden

NANCY A. MATHIOWETZ
U.S. Bureau of the Census

SEYMOUR SUDMAN
University of Illinois at Urbana-Champaign

Morris H. Hansen

*This book is dedicated to the memory of Morris Hansen,
born in 1911 who died on October 9, 1990 in Washington,
D.C. In addition to his many important contributions to
survey sampling, Morris and his colleagues at the U.S.
Bureau of the Census pioneered the development of non-
sampling error research in the early 1950s and 1960s.
Those accomplishments, as well as his recent ones,
continue to influence the development of the field of
survey nonsampling error.*

This page intentionally left blank

CONTRIBUTORS

Frank M. Andrews, Survey Research Center, Institute of Social Research, The University of Michigan, Ann Arbor, Michigan

Barbara Bickart, Department of Marketing, University of Florida, Gainesville, Florida

Paul P. Biemer, Research Triangle Institute, Research Triangle Park, North Carolina

Johnny Blair, Survey Research Center, University of Maryland, College Park, Maryland

Norman M. Bradburn, University of Chicago and NORC, Chicago, Illinois

David Cantor, Westat, Inc., Rockville, Maryland

Rachel Caspar, Research Triangle Institute, Research Triangle Park, North Carolina

Edith D. de Leeuw, Vrije Universiteit, Amsterdam, The Netherlands

Wil Dijkstra, Vrije Universiteit, Amsterdam, The Netherlands

Don A. Dillman, U.S. Bureau of the Census, Washington, D.C. and Washington State University, Pullman, Washington

John E. Donmyer, A.C. Nielsen Co., Northbrook, Illinois

Judith Droitcour, U.S. General Accounting Office, Washington, D.C.

Solomon Dutka, Audits and Surveys, Inc., New York, New York

W. Sherman Edwards, Westat, Inc., Rockville, Maryland

Donna Eisenhower, Westat, Inc., Rockville, Maryland

James L. Esposito, National Center for Health Statistics, Hyattsville, Maryland

Trena M. Ezzati, National Center for Health Statistics, Hyattsville, Maryland

Ron Fecso, National Agricultural Statistical Service, U.S. Department of Agriculture, Washington, D.C.

Gösta Forsman, University of Linköping, Linköping, Sweden

Barbara H. Forsyth, *Research Triangle Institute, Research Triangle Park, North Carolina*

Floyd J. Fowler, Jr., Center for Survey Research, University of Massachusetts, Boston, Massachusetts

Lester R. Frankel, Audits and Surveys, Inc., New York, New York

Wayne A. Fuller, Department of Statistics, Iowa State University, Ames, Iowa

Robert M. Groves, U.S. Bureau of the Census, Washington, D.C. and Survey Research Center, The University of Michigan, Ann Arbor, Michigan

Daniel H. Hill, Opinion Research Institute, The University of Toledo, Toledo, Ohio

Hans-Jürgen Hippler, Zentrum für Umfragen, Methoden, und Analysen (ZUMA), Mannheim, Germany

David Holt, Department of Social Statistics, University of Southampton, Southampton, United Kingdom

Joop J. Hox, University of Amsterdam, Amsterdam, The Netherlands

Michael L. Hubbard, Research Triangle Institute, Research Triangle Park, North Carolina

Jared B. Jobe, National Center of Health Statistics, Hyattsville, Maryland

Daniel Kasprzyk, National Center for Education Statistics, Washington, D.C.

Ita G. G. Kreft, University of California at Los Angeles, Los Angeles, California

William Kruskal, Department of Statistics, University of Chicago, Chicago, Illinois

Judith T. Lessler, Research Triangle Institute, Research Triangle Park, North Carolina

Lars E. Lyberg, Statistics Sweden, Stockholm, Sweden

Nancy A. Mathiowetz, U.S. Bureau of the Census, Washington, D.C.

Barbara Means, SRI International, Menlo Park, California

Geeta Menon, Department of Marketing, New York University, New York, New York

Nico Molenaar, Vrije Universiteit, Amsterdam, The Netherlands

David Morganstein, Westat, Inc., Rockville, Maryland

Ingrid M. E. Munck, Statistics Sweden, Stockholm, Sweden

Colm O'Muircheartaigh, London School of Economics and Political Science, London, United Kingdom

Jeroen Pannekoek, Central Bureau of Statistics, Voorburg, The Netherlands

Teresa L. Parsley, Research Triangle Institute, Research Triangle Park, North Carolina

Frank W. Piotrowski, A. C. Nielsen Co., Northbrook, Illinois

J. N. K. Rao, Department of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada

Willem E. Saris, University of Amsterdam, Amsterdam, The Netherlands

Nora Cate Schaeffer, Department of Sociology, University of Wisconsin, Madison, Wisconsin

Irwin Schreiner, U.S. Bureau of the Census, Washington, D.C.

Norbert Schwarz, Zentrum für Umfragen, Methoden, und Analysen (ZUMA), Mannheim, Germany

Stuart Scott, U.S. Bureau of Labor Statistics, Washington, D.C.

Adriana R. Silberstein, U.S. Bureau of Labor Statistics, Washington, D.C.

Christopher J. Skinner, University of Southampton, Southampton, United Kingdom

Johannes H. Smit, Vrije Universiteit, Amsterdam, The Netherlands

Tom W. Smith, NORC and University of Chicago, Chicago, Illinois

S. Lynne Stokes, Department of Management Sciences and Information Systems, University of Texas, Austin, Texas

Seymour Sudman, Survey Research Laboratory, University of Illinois, Urbana, Illinois

Gary E. Swan, SRI International, Menlo Park, California

John Tarnai, Washington State University, Pullman, Washington

D. Roland Thomas, Department of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada

Johannes van der Zouwen, Vrije Universiteit, Amsterdam, The Netherlands

Piet G. W. M. van Dosselaar, Central Bureau of Statistics, Voorburg, The Netherlands

Wendy Visscher, Research Triangle Institute, Research Triangle Park, North Carolina

Kirk M. Wolter, A. C. Nielsen Co., Northbrook, Illinois

CONTENTS

SECTION E. MODELING MEASUREMENT ERRORS
AND THEIR EFFECTS ON ESTIMATION
AND DATA ANALYSIS

This page intentionally left blank

PREFACE

Survey measurement error, as the term is used in this book, refers to error in survey responses arising from the method of data collection, the respondent, or the questionnaire (or other instrument). It includes the error in a survey response as a result of respondent confusion, ignorance, carelessness, or dishonesty; the error attributable to the interviewer, perhaps as a consequence of poor or inadequate training, prior expectations regarding respondents' responses, or deliberate errors; and error attributable to the wording of the questions in the questionnaire, the order or context in which the questions are presented, and the method used to obtain the responses. At the time survey responses are collected, all of these factors may intervene and interact in such a way as to degrade response accuracy. The consequences of measurement errors are survey results that, while ostensibly accurate and credible may in reality be quite inaccurate and misleading. Measurement error does not include the errors of nonobservation (such as nonresponse error), processing errors, or other errors occurring after data collection. Thus, these errors are intentionally not discussed in the present book.

Given the importance of the topic of measurement error in survey work, the Survey Research Methods Section (SRM) of the American Statistical Association (ASA) in 1986 determined that survey measurement error should be the topic of an SRM-sponsored conference and approached Paul Biemer to develop the idea. Two years later, while Biemer was visiting Statistics Sweden, the groundwork was laid for a measurement error conference and an edited monograph of the conference-invited papers. It was decided that the conference should seek the participation of researchers worldwide. Further, the scope of the conference was expanded from one which focused only on the statistical aspects of measurement error to one which also enveloped important research being conducted in sociology, psychometrics, psychology, market research, and other disciplines as well. A preliminary outline for the book, which would also serve as a general framework for the conference sessions, was developed.

By the fall of 1988, an organizing/editing committee was formed consisting of: Paul P. Biemer (as Chair), Robert M. Groves, Lars Lyberg, Nancy A. Mathiowetz, and Seymour Sudman. Gösta Forsman was enlisted to assist in planning and conducting the conference. The

committee contacted numerous research organizations for monetary contributions. The committee also reviewed and finalized the monograph outline and began to identify and contact researchers throughout the world as potential authors. Abstracts were requested and 128 abstracts were received from researchers interested in writing for the monograph. From these, the committee selected 32 to become the chapters in this monograph and developed the conference program. Kate Roach at John Wiley & Sons was contacted and Wiley agreed to publish the book.

Four professional organizations were asked to sponsor the conference: the American Association of Public Opinion Research (AAPOR), the American Marketing Association (AMA), the American Statistical Association (ASA), and the International Association of Survey Statisticians (IASS). All four organizations enthusiastically agreed. Two of these (ASA and AAPOR) also contributed funds to support the project. In addition, the following research organizations contributed funds:

Australian Bureau of Statistics
Central Statistical Office of Finland
Istituto Centrale di Statistica, Italy
National Agricultural Statistics Service
National Center for Health Statistics
National Science Foundation
Nielsen Media Research
NORC
Office of Population Censuses and Surveys, United Kingdom
Research Triangle Institute
Statistics Sweden
U.S. Bureau of Labor Statistics
U.S. Bureau of the Census
U.S. General Accounting Office
Westat, Inc.

Without the financial support of these organizations, the conference and edited monograph would not have been possible.

The International Conference on Measurement Errors in Surveys was held on November 11–14, 1990 in Tucson, Arizona. It drew 421 attendees from 15 countries. The program consisted of 60 invited papers (including the 32 chosen for the present book) and 70 contributed papers. Additionally, two short courses dealing with measurement errors were presented. The number of presented papers and the number of attendees surprised and delighted the committee who, in the early planning stages, had anticipated a much smaller program and audience, because of the specialized topic of the conference.

During the conference, the committee was assisted by eight graduate

student fellows. Six of these were selected from applicants worldwide to attend the conference and the two others were sponsored by Iowa State University to attend the conference and to assist the committee. These eight fellows were: Tom Belin, Harvard University; Karen Bogen, University of Michigan; Joseph Croos, Iowa State University; Stasja Draisma, Vrije Universiteit, The Netherlands; Daniel Merkel, Northwestern University; Steven Pennell, University of Michigan; Todd Rockwood, Washington State University; and Todd Sanger, Iowa State University. The fellows took notes during discussions between section editors and authors and during floor discussions after the monograph papers were presented. At least one fellow was available at every session to ensure that the session would run smoothly.

In designing the book, the committee did not intend to merely publish a conference proceedings volume. Rather the aim was to collectively write a book dealing with the most important issues in the field of survey measurement error, attempting whenever possible to integrate diverse perspectives. Thus, each chapter has undergone extensive editing, review, and revision. The book is organized into five sections. The section titles and their editors are:

Section A: The Questionnaire (Seymour Sudman)
Section B: Respondents and Responses (Nancy A. Mathiowetz)
Section C: Interviewers and Other Means of Data Collection (Lars Lyberg)
Section D: Measurement Errors in the Interview Process (Robert M. Groves)
Section E: Modeling Measurement Errors and Their Effects on Estimation and Data Analysis (Paul P. Biemer)

The reader will notice that the sections are not mutually exclusive. The difficulty in knowing exactly where a paper should be placed merely reflects the complex nature of survey data collection.

In designing this volume, difficult decisions had to be made about what to exclude. Some topics were excluded because they were only marginally related to the main theme of survey measurement error. Other topics were omitted because we could identify no current research, despite the fact that new knowledge might be needed. Readers will note an emphasis on population and household surveys. We have, however, included three chapters on business and establishment surveys (Chapters 7, 12, and 18) and one on crop yield surveys (Chapter 17).

Each section editor had responsibilities as a secondary editor for at least one other section as well. The authors of the chapters, in addition to their extensive writing and revising activities, were also involved in the review of the other monograph chapters. They were encouraged to seek

outside reviews for their chapters on their own. Thus, the monograph reflects the efforts and contributions of scores of writers, editors, and reviewers.

The diversity of orientations of the authors for the monograph made it impossible to impose a unified terminology and set of notation across all chapters. Two chapters (Chapter 1 and Chapter 24) have been provided as an attempt to link the terminological and notational conventions of the various disciplines represented in the book. Further, except for one section (Section E), the statistical level of the monograph is quite accessible by graduate students in sociology, psychology, or marketing research. Section E, however, which deals with the more theoretical statistical side of measurement error research, requires a fairly thorough grounding in survey sampling and mathematical statistics, at the level of *Introduction to Mathematical Statistics*, 3rd ed., by R. V. Hogg and A. Craig, 1970.

Although the present book can serve as a course text, the primary audience is researchers having some prior training in survey research and/or survey methods. Since it contains a number of review articles on measurement error research in several disciplines, it will be useful to researchers actively engaged in measurement error research who want an introduction to the techniques and research issues in the field from a different theoretical perspective. The book will also be useful to the survey methodologist or survey practitioner who wants to learn more about the causes, consequences, and cures of survey error in order to improve the quality of surveys through better design, data collection, and analytical techniques. This book is far from the final word on measurement error. It does, however, reflect current knowledge in 1990, to the best of our editorial judgment. As a group, we hope that its publication will stimulate future research in this exiting field.

Gösta Forsman deserves great appreciation for all the activities he performed so ably for the conference and book project. We are also truly grateful to Lee Decker of ASA, who tirelessly and efficiently handled an enormous number of logistical details associated with the conference.

Sincere thanks go to Cheryl Crawford and Mary Kay Martinez, at New Mexico State University, and to Linda Miller, at Research Triangle Institute, who performed many clerical and secretarial functions associated with the project. Thanks are due to Graham Kalton and Wayne Fuller for their efforts, while chairing the Survey Research Methods Section of the ASA, in promoting the idea of a conference on nonsampling error. We are appreciative of the efforts of Pat Dean, Dan Kasprzyk, and Lynne Stokes, who assisted in the review of a number of manuscripts, and Pat Ellis, who prepared the final copy of the reference list. Our employing organizations also deserve great appreciation for supporting our activi-

ties in assembling the book: New Mexico State University and Research
Triangle Institute (Biemer); Statistics Sweden (Lyberg); Survey Research
Center at the University of Michigan (Groves); National Center for
Health Services Research (Mathiowetz); U.S. Bureau of the Census
(Groves and Mathiowetz); and University of Illinois at Urbana–Champaign (Sudman).

<div align="right">

PAUL BIEMER

ROBERT GROVES

LARS LYBERG

NANCY MATHIOWETZ

SEYMOUR SUDMAN

</div>

Research Triangle Park, North Carolina

Washington, D.C.

Stockholm, Sweden

Washington, D.C.

Urbana, Illinois

November 1991

This page intentionally left blank

INTRODUCTION

William Kruskal
University of Chicago

This volume summarizes the International Conference on Measurement Errors in Surveys, held in November 1990 in Tucson, Arizona. I agreed to be the opening speaker as an opportunity to present some thoughts to the conference participants: primarily to illustrate how pervasive in society are concepts of measurement error, how important they are to understanding our world, and thus to motivate this volume.

Dare we ask what is truth? Some readers of this volume are surely members of the AARP, the American Association of Retired Persons, which has a close relationship with the NRTA, the National Retired Teachers Association. Members receive a journal called *Modern Maturity*, a somewhat bland, inspirational magazine that elaborates on Robert Browning's "Grow old along with me." In last March's *Modern Maturity* I was delighted to find an excellent article on philosophy by Steven Ross (1990), City University of New York. Ross treats the big, big questions of epistemology and ethics in a down to earth, up-to-date, and lucid way that is relevant to a big question for this conference: is there a true value for what is measured by an observation or a survey?

Ross sketches the traditional positivist position and the attacks on it by Wittgenstein and Quine. We are left with no sharp lines — perhaps even no fuzzy lines — between theory and observation. Many of us feel uncomfortable with that position, for if there is no true value lurking in the platonic underbrush how can we talk sensibly about error? And if we cannot talk sensibly about error, how can we design surveys and do inferences?

There are other philosophical issues for surveys: for example, should we regard interviewers as introducing fixed but unknown systematic effects each, or should we think of their individual effects as randomly chosen? Or should we do both . . . perhaps in separate analyses? Or again

should we be more realistic (aha, note my recourse to realism as a goal) and introduce models with stochastic dependence among interviewers. A similar distinction arises in so-called dual systems analyses . . . of which I will say more later. One of the two samples may be taken idealistically as perfect in terms of measurement error; at another extreme, the two samples might be regarded as replicates, i.e., independent and identically distributed.

How we handle these questions can make major differences in our analyses of error structures and hence yield major differences in conclusions. Our choice may also be relevant to debates among survey organizations, for example, about whether it is better to maintain a relatively fixed, highly trained interviewing staff or to arrange frequent turnover, short-term training, and consequently possible lower esprit de corps?

There is a widespread desire to cling to faith in an objective reality out there; I confess that I feel it strongly on Mondays, Wednesdays, and Fridays. Lionel Trilling (1951, p. 4) paraphrased the historian Vernon L. Parrington (before criticizing him):

> There exists . . . a thing called *reality*; it is one and immutable, it is wholly external, it is irreducible. Men's minds may waver, but reality is always reliable, always the same, always easily to be known.

The last part of that statement — "always easily to be known" — is not part of my thrice a week mind-frame, and I expect that few statisticians would agree that reality is easily known. Indeed we earn our livelihoods because of difficulties in knowing reality, or whatever each of us accepts as a substitute for, or reflection of, reality.

Another American historian, Oscar Handlin, said more recently (1979, p. 405):

> Truth is absolute, it is as absolute as the world is real . . . the historian's vocation depends on this minimal operational article of faith.

The opposing viewpoint, that there are no true values, just different ways of measurement, has had statistical proponents for many years, perhaps most notably W. E. Deming (e.g., Deming, 1950, Ch. 1). Another example is a 1939 (p. 135) quotation by Walter A. Shewhart:

> . . . Consider in contrast the concept of the true value X' of a quality characteristic, for example the length of a line AB, or the velocity of light. I am not able even to conceive of a physical operation of observing or experiencing a true length X'. You may argue that there are ways of measuring the length of a line, by any one of which you may obtain a sequence of observations; you may even argue that the limiting average $\bar{X}'$ is equal to X'. But the physical operation is a

method of obtaining $\bar{X}'$, not X'. Whether $\bar{X}' = X'$ we shall never know. The true length X' is the given, unknowable, unapproachable, ineffable. It is removed from the pale of observation as securely as $\sqrt{-1}$ is removed from the realm of real numbers; there is not even the question of approximating $\sqrt{-1}$ with the rational and irrational numbers.

Attacks on some sort of absolute truth have a long literary history as well. I have a hard time trying to understand continental literary figures like Jacques Derrida, but their frequent scorn for simple truth comes through clearly. One continental literary figure, now turned national leader, Vaclav Havel, is relatively understandable. He was recently quoted in Senator Daniel Patrick Moynihan's newsletter (April 20, 1990). Havel, speaking to a joint meeting of the Congress about events in eastern Europe, said that he had "one great certainty: Consciousness precedes Being, and not the other way around, as the Marxists claim." Moynihan adds that "this is a real issue to intellectuals such as Havel; an issue men and women have died for; it holds that the beliefs create the 'real' world and not vice versa."

I might happily continue along these lines. For example, I might repeat to you Higgins's Law from a John McPhee *New Yorker* article on maritime trade (April 2, 1990, p. 74):

Use one electronic [location] system and you always know where you are. Use more than one and you're always in doubt.

A great potential text for a statistical sermon. Or I might present a book review of Miles Orvell's recent *The Real Thing* (1989), in part named after Henry James's story about genuine aristocrats who were far less able to pose for an artist illustrating a book about aristocrats than were lower-class models. That summary rides brutally over Jamesian subtleties, but I have no choice. We must go on to other subtopics, keeping in mind the breadth of our interests, and awaiting Bruce Spencer's conference paper on error in true values (Spencer, 1990).

Let me first, however, note a topic that would merit treatment at length if we had the length. There may be true values that lie behind each individual measurement: your true age, your true belief about celestial music, the true income of family i. Then there may be true values for an entire population. For additive qualities like money that is not such a big step. For beliefs and opinions, true value for the entirety can be troublesome.

This is a good time to present a brief outline of topics. I have dealt briefly with the problem of true values. Next I turn to scope and ask how widely or narrowly the concept of survey might be construed. I then develop a fundamental approach to measurement error, with consideration of the apparent human need for a single figure of merit. Also

discussed are the bottom-up and top-down versions of total survey error. We end with a few remarks about the study of error structure.

Scope; what are surveys? Let me place aside various claims or restrictions. It is artificial to limit our measurement error discussion to social surveys, to person-to-person question and answers, or to finite populations. We want to be able to cover not only standard questions about age, sex, ethnicity, prices, incomes, etc., not only attitude questions about public figures and policies, but also questions about luminosity of stars, pollen counts in Chicago, how fast Peter Pan flies (for studies of myth), the economic backgrounds of members of Parliament in 1806, medical diagnoses, and limitlessly on. Robert Groves (Chapter 1, this book) writes about many of these vistas. At present I pause mainly to express a caution about overdefensiveness. Social surveys are sometimes given rough handling by so-called hard scientists, cynical newspaper columnists, and congressmen at budget hearings. Surely some of those attacks are justified, but a great many are not, and the attacks often come from glass houses. All of us, no doubt, have inspected glass houses of biology and medicine. As to physical sciences, many of you have doubtless seen the famous graph of thermal conductivity of copper that is poles away from one traditional view of physics as a pure science. It has been popularized by J. S. Hunter and appears in Turner and Martin (1984, p. 15). I do not at all wish to denigrate physical science, but I wonder how much of its literature is open to sharp statistical criticism. Problems arise in connection with estimating fundamental constants of nature, for example, see W. L. Wiese et al. (1966) on atomic transition probabilities. Other problems arise in contexts with political and economic facets: examples are cold fusion, anomalous water, and space probe reliability. There are other attacks on the crystal house of physics, for example, a 1983 book by Nancy Cartwright with the vigorous title *How the Laws of Physics Lie*. A related lively discussion is given by Huber (1990).

In a spirit of continued professional iconoclasm, let me mention some other traits often said to characterize the field of sample surveys, although I am not persuaded about any. Thus I see no crisp lines between surveys and other parts of statistics.

Emphasis on finite populations, to repeat a point, is frequently mentioned, but it does not strike me as of the essence. Further, *all* statistical populations are finite because of rounding resolution of measuring instruments; our ruler only measures to 1/16 inch or whatever. Traditional continuous models are useful approximations, not gospel.

A related trait is the usual exceedingly wide family of probability models common in sample survey theory . . . as opposed to much narrower parametric models in traditional statistics. Yet that boundary strikes me

as hazy because, on the one hand, there is a movement toward more nearly parameterized models for sampling studies; on the other hand, nonparametrics uses extremely wide models, similar to those in sampling theory, that depend on broad structures only, like symmetry, independence, and equidistribution.

Also related is the feeling by some that survey sampling deals primarily with totals and averages, leaving more complex quantities to other hands. I think that simply misleading.

Some say that survey sampling deals only with *descriptive* statistics, as opposed to *inferential* statistics. Yet inferential questions arise widely in surveys. In particular, surveys have had central roles in properly experimental studies.

Finally, some colleagues believe that the essence of survey sampling is the fundamental presence of *labels*, identifiers of the sampled units by name, number, etc.. That's a fruitful assertion but I think labels neither necessary nor sufficient for sample surveydom.

That leaves me without sharp boundaries for this subtaxonomy of statistics, a condition of mild discomfort yet honesty.

Approach to measurement error. It is high time to turn to our main theme: an approach to measurement error that seems to me basic and not always acknowledged. To start with, let's suppose that we have a population of people (or cars or stars or trout) and that we have a measurement system for *age*. It might be self-report, lookup in administrative files, counting scales or teeth, whatever. Let's say age in years for simplicity and let's, at least pro tempore, suppose that there is a true age for each individual of interest, i.e., in a specified population. We can think in principle of attaching to each individual a distribution for observed age. For Jim Brown, true age 31, that distribution might be

observed age	21	...	29	30	31	32	33	...	41
probability	0.1		0.075	0.1	0.5	0.075	0.05		0.1

That is, the observation error in measuring Jim Brown's age to the closest year is zero half the time, usually two years or less, and with a nonnegligible chance of a ten-year error exactly.

Our motivations for attacking error of age and other measurements are mainly to gain improved understanding of the process, thus making better decisions based on the measurements and deciding what components should get priority for improvement.

Now one might consider such separate hypothetical distributions, one for each individual, but I suspect that many of us would like to simplify by supposing the observed age distribution to depend only on the true age. In any case, this approach pushes one to consider explicitly what simplifying assumptions are being made.

Starting with such distributions — usually hypothetical — of observed age, the next logical step is introduction of sampling. (I know that this volume is on measurement, not sampling, yet the two broad topics are intertwined.) For any suggested sampling procedures and any set of little distributions like the example, a putting-together provides one with a table having columns (say) for true age in years, rows for observed age, and cell entries for proportions. I think of all entries as summing to one. A proportion 0.017 in column 16 and row 18 would mean that with probability 17 out of 1,000 a probabilistically chosen population member will in fact be 16 and measured as 18 years old.

When true value is a questionable concept, and even when it makes sense, the joint distribution for two or more modes of measurement might naturally be considered. For example, age might be measured (1) by asking a person directly for age, (2) by asking for year of birth, (3) by asking another household member, (4) by looking at a birth certificate, (5) by interviewer guess; there might well be others and the empirical study of their interconnections presents a constant task for survey organizations. Instead of starting with a univariate distribution for Jim Brown's age, then, we might start with a bivariate distribution in which rows and columns correspond to observed age by mode 1 and mode 2 respectively; one cell, for example, would give the probability that measurement mode 1 provides 21 years and mode 2, 29 years. Then as before one might introduce a sampling scheme and obtain a three-dimensional table in which one dimension corresponds to true age and the other two dimensions correspond to the two measurement modes. Or one might omit the true age dimension as meaningless. Decisions such as the latter one might be appropriate for attitude studies or for surveys of beliefs about heaven.

Age is, of course, a metric variable and a relatively objective one . . . unless one gets down to fine detail, as in some evaluations of astrology. One can also go through an analogous discussion for qualitative measurements: ethnicity, state of birth, breed of horse, chromosome on which a gene sits, and so on. The central point is that an approach of the kind sketched pushes one into fundamental thought about measurement error.

Simplification becomes necessary in practice, and in the case of quantitative variables like income, height, length of life, boiling point, etc., it may be natural to look at diagonal differences in the basic table, that is, to look at the observed value *minus* the true value . . . or, when true value is questionable, to look at the difference between observed values for two modes.

Caution: we should by no means automatically jump to differences. For example, if we are interested in age-heaping, round numbers play a