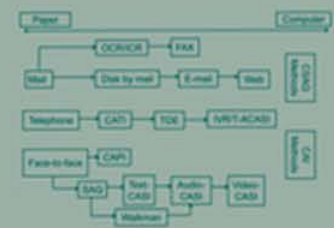


$$d^2 = \frac{v(\bar{y}_{ST})}{v_{SRS}(\bar{y})} = \frac{\sum_{h=1}^H W_h^2 \left(\frac{1-f_h}{n_h} \right) s_h^2}{\left(\frac{1-f}{n} \right) s^2}$$

Survey Methodology

SECOND EDITION

Robert M. Groves, Floyd J. Fowler, Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau

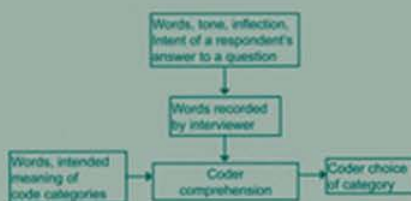


Frame Data	Interview Data
98012408A21	1214822079910867
98020604A54	026075412267825071231841231841
98020508B14	54579490121106074108821088210882
98040679A6	3646541886744610461046104610461
98060808A1	32105487446104610461046104610461
98080109B7	01784541081664010816640108166401
98071008B1	87784541081664010816640108166401
98080411C3	88165791025410264010264010264010264
98090133B4	88165791025410264010264010264010264
98100801D1	88165791025410264010264010264010264
98110002D6	88165791025410264010264010264010264
98120708A3	88165791025410264010264010264010264



interviewer design effect =

$$deff_{int} = 1 + \rho_{int}(m-1)$$



SURVEY METHODOLOGY

WILEY SERIES IN SURVEY METHODOLOGY

Established in Part by WALTER A. SHEWHART AND SAMUEL S. WILKS

Editors: *Robert M. Groves, Graham Kalton, J. N. K. Rao, Norbert Schwarz,
Christopher Skinner*

A complete list of the titles in this series appears at the end of this volume.

SURVEY METHODOLOGY

Second Edition

Robert M. Groves
Floyd J. Fowler, Jr.
Mick P. Couper
James M. Lepkowski
Eleanor Singer
Roger Tourangeau

 **WILEY**

A JOHN WILEY & SONS, INC., PUBLICATION

Copyright © 2009 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008 or online at <http://www.wiley.com/go/permission>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the U.S. at (800) 762-2974, outside the U.S. at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic format. For information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Survey methodology / Robert Groves ... [et al.]. — 2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-46546-2 (paper)

1. Surveys—Methodology. 2. Social surveys—Methodology. 3. Social sciences—Research—Statistical methods. I. Groves, Robert M.

HA31.2.S873 2009

001.4'33--dc22

2009004196

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE TO THE FIRST EDITION	xv
PREFACE TO THE SECOND EDITION	xix
ACKNOWLEDGMENTS	xxi
CHAPTER 1. AN INTRODUCTION TO SURVEY METHODOLOGY	
1.1 Introduction	2
1.2 A Brief History of Survey Research	3
1.2.1 The Purposes of Surveys	3
1.2.2 The Development of Standardized Questioning	5
1.2.3 The Development of Sampling Methods	6
1.2.4 The Development of Data Collection Methods	7
1.3 Some Examples of Ongoing Surveys	7
1.3.1 The National Crime Victimization Survey	8
1.3.2 The National Survey on Drug Use and Health	14
1.3.3 The Surveys of Consumers	17
1.3.4 The National Assessment of Educational Progress	20
1.3.5 The Behavioral Risk Factor Surveillance System	24
1.3.6 The Current Employment Statistics Program	27
1.3.7 What Can We Learn From the Six Example Surveys?	29
1.4 What is Survey Methodology?	30
1.5 The Challenge of Survey Methodology	32
1.6 About this Book	34
Keywords	35
For More In-Depth Reading	35
Exercises	36

CHAPTER 2. INFERENCE AND ERROR IN SURVEYS

2.1	Introduction	39
2.2	The Lifecycle of a Survey from a Design Perspective	41
2.2.1	Constructs	41
2.2.2	Measurement	43
2.2.3	Response	43
2.2.4	Edited Response	44
2.2.5	The Target Population	44
2.2.6	The Frame Population	45
2.2.7	The Sample	45
2.2.8	The Respondents	46
2.2.9	Postsurvey Adjustments	47
2.2.10	How Design Becomes Process	48
2.3	The Lifecycle of a Survey from a Quality Perspective	49
2.3.1	The Observational Gap between Constructs and Measures	50
2.3.2	Measurement Error: the Observational Gap between the Ideal Measurement and the Response Obtained	52
2.3.3	Processing Error: the Observational Gap between the Variable Used in Estimation and that Provided by the Respondent	53
2.3.4	Coverage Error: the Nonobservational Gap between the Target Population and the Sampling Frame	54
2.3.5	Sampling Error: The Nonobservational Gap between the Sampling Frame and the Sample	56
2.3.6	Nonresponse Error: The Nonobservational Gap between the Sample and the Respondent Pool	59
2.3.7	Adjustment Error	59
2.4	Putting It All Together	60
2.5	Error Notions in Different Kinds of Statistics	61
2.6	Nonstatistical Notions of Survey Quality	62
2.7	Summary	63
	Keywords	64
	For More In-Depth Reading	64
	Exercises	65

CHAPTER 3. TARGET POPULATIONS, SAMPLING FRAMES, AND COVERAGE ERROR

3.1	Introduction	69
3.2	Populations and Frames	69

3.3	Coverage Properties of Sampling Frames	72
3.3.1	Undercoverage	72
3.3.2	Ineligible Units	76
3.3.3	Clustering of Target Population Elements Within Frame Elements	77
3.3.4	Duplication of Target Population Elements in Sampling Frames	79
3.3.5	Complicated Mappings between Frame and Target Population Elements	80
3.4	Alternative Frames for the Target Population of Households or Persons	81
3.4.1	Area Frames	81
3.4.2	Telephone Number Frames for Households or Persons	81
3.4.3	Frames for Web Surveys of General Populations	83
3.5	Frame Issues for Other Common Target Populations	84
3.5.1	Customers, Employees, or Members of an Organization	84
3.5.2	Organizations	85
3.5.3	Events	86
3.6.4	Rare Populations	87
3.6	Coverage Error	87
3.7	Reducing Undercoverage	88
3.7.1	The Half-Open Interval	88
3.7.2	Multiplicity Sampling	90
3.7.3	Multiple Frame Designs	91
3.7.4	Increasing Coverage While Including More Ineligible Elements	93
3.8	Summary	94
	Keywords	95
	For More In-Depth Reading	95
	Exercises	95

CHAPTER 4. SAMPLE DESIGN AND SAMPLING ERROR

4.1	Introduction	97
4.2	Samples and Estimates	99
4.3	Simple Random Sampling	103
4.4	Cluster Sampling	106
4.4.1	The Design Effect and Within-Cluster Homogeneity	110
4.4.2	Subsampling within Selected Clusters	113

4.5	Stratification and Stratified Sampling	113
4.5.1	Proportionate Allocation to Strata	116
4.5.2	Disproportionate Allocation to Strata	122
4.6	Systematic Selection	123
4.7	Complications in Practice	125
4.7.1	Two-Stage Cluster Designs with Probabilities Proportionate to Size (PPS)	127
4.7.2	Multistage and Other Complex Designs	129
4.7.3	How Complex Sample Designs are Described: The Sample Design for the NCVS	130
4.8	Sampling US Telephone Households	133
4.9	Selecting Persons Within Households	136
4.10	Summary	138
	Keywords	139
	For More In-Depth Reading	139
	Exercises	139

CHAPTER 5. METHODS OF DATA COLLECTION

5.1	Alternative Methods of Data Collection	150
5.1.1	Degree of Interviewer Involvement	153
5.1.2	Degree of Interaction with the Respondent	154
5.1.3	Degree of Privacy	155
5.1.4	Channels of Communication	156
5.1.5	Technology Use	157
5.1.6	Implications of these Dimensions	158
5.2	Choosing the Appropriate Method	159
5.3	Effects of Different Data Collection Methods on Survey Errors	160
5.3.1	Measuring the Marginal Effect of Mode	160
5.3.2	Sampling Frame and Sample Design Implications of Mode Selection	162
5.3.3	Coverage Implications of Mode Selection	163
5.3.4	Nonresponse Implications of Mode Selection	166
5.3.5	Measurement Quality Implications of Mode Selection	168
5.3.6	Cost Implications	173
5.3.7	Summary on the Choice of Method	174
5.4	Using Multiple Modes of Data Collection	175
5.5	Summary	177
	Keywords	178
	For More In-Depth Reading	179
	Exercises	179

CHAPTER 6. NONRESPONSE IN SAMPLE SURVEYS

6.1	Introduction	183
6.2	Response Rates	183
6.2.1	Computing Response Rates	184
6.2.2	Trends in Response Rates Over Time	186
6.3	Impact of Nonresponse on the Quality of Survey Estimates	189
6.4	Thinking Causally About Survey Nonresponse Error	191
6.5	Dissecting the Nonresponse Phenomenon	192
6.5.1	Unit Nonresponse Due to Failure to Deliver the Survey Request	193
6.5.2	Unit Nonresponse Due to Refusals	197
6.5.3	Unit Nonresponse Due to the Inability to Provide the Requested Data	201
6.6	Design Features to Reduce Unit Nonresponse	201
6.7	Item Nonresponse	208
6.8	Are Nonresponse Propensities Related to Other Error Sources?	210
6.9	Summary	210
	Keywords	211
	For More In-Depth Reading	211
	Exercises	211

CHAPTER 7. QUESTIONS AND ANSWERS IN SURVEYS

7.1	Alternatives Methods of Survey Measurement	217
7.2	Cognitive Processes in Answering Questions	218
7.2.1	Comprehension	220
7.2.2	Retrieval	220
7.2.3	Estimation and Judgment	222
7.2.4	Reporting	223
7.2.5	Other Models of the Response Process	223
7.3	Problems in Answering Survey Questions	225
7.3.1	Encoding Problems	225
7.3.2	Misinterpreting the Questions	226
7.3.3	Forgetting and Other Memory Problems	229
7.3.4	Estimation Processes for Behavioral Questions	234
7.3.5	Judgment Processes for Attitude Questions	236
7.3.6	Formatting the Answer	237
7.3.7	Motivated Misreporting	240
7.3.8	Navigational Errors	241
7.4	Guidelines for Writing Good Questions	242

7.4.1 Nonsensitive Questions About Behavior	243
7.4.2 Sensitive Questions About Behavior	246
7.4.3 Attitude Questions	248
7.4.4 Self-Administered Questions	251
7.5 Summary	252
Keywords	254
For More In-Depth Reading	254
Exercises	255

CHAPTER 8. EVALUATING SURVEY QUESTIONS

8.1 Introduction	259
8.2 Expert Reviews	260
8.3 Focus Groups	261
8.4 Cognitive Interviews	263
8.5 Field Pretests and Behavior Coding	265
8.6 Randomized or Split-Ballot Experiments	267
8.7 Applying Question Standards	268
8.8 Summary of Question Evaluation Tools	269
8.9 Linking Concepts of Measurement Quality to Statistical Estimates	274
8.9.1 Validity	274
8.9.2 Response Bias	279
8.9.3 Reliability and Simple Response Variance	281
8.10 Summary	286
Keywords	287
For More In-Depth Reading	287
Exercises	288

CHAPTER 9. SURVEY INTERVIEWING

9.1 The Role of the Interviewer	291
9.2 Interviewer Bias	292
9.2.1 Systematic Interviewer Effects on Reporting of Socially Undesirable Attributes	292
9.2.2 Systematic Interviewer Effects on Topics Related to Observable Interviewer Traits	292
9.2.3 Systematic Interviewer Effects Associated with Interviewer Experience	294
9.3 Interviewer Variance	295
9.3.1 Randomization Requirements for Estimating	

Interviewer Variance	296
9.3.2 Estimation of Interviewer Variance	297
9.4 Strategies for Reducing Interviewer Bias	300
9.4.1 The Role of the Interviewer in Motivating Respondent Behavior	300
9.4.2 Changing Interviewer Behavior	301
9.5 Strategies for Reducing Interviewer-Related Variance	302
9.5.1 Minimizing Questions that Require Nonstandard Interviewer Behavior	303
9.5.2 Professional, Task-Oriented Interviewer Behavior	304
9.5.3 Interviewers Reading Questions as They Are Worded	305
9.5.4 Interviewers Explaining the Survey Process to the Respondent	306
9.5.5 Interviewers Probing Nondirectively	308
9.5.6 Interviewers Recording Answers Exactly as Given	311
9.5.7 Summary on Strategies to Reduce Interviewer Variance	312
9.6 The Controversy About Standardized Interviewing	312
9.7 Interviewer Management	315
9.7.1 Interviewer Selection	315
9.7.2 Interviewer Training	316
9.7.3 Interviewer Supervision and Monitoring	317
9.7.4 The Size of Interviewer Workloads	318
9.7.5 Interviewers and Computer Use	318
9.8 Validating the Work of Interviewers	319
9.9 The Use of Recorded Voices (and Faces) in Data Collection	322
9.10 Summary	323
Keywords	324
For More In-Depth Reading	324
Exercises	325

CHAPTER 10. POSTCOLLECTION PROCESSING OF SURVEY DATA

10.1 Introduction	329
10.2 Coding	331
10.2.1 Practical Issues of Coding	332
10.2.2 Theoretical Issues in Coding Activities	334
10.2.3 “Field Coding” – An Intermediate Design	334
10.2.4 Standard Classification Systems	337

10.2.5 Other Common Coding Systems	341
10.2.6 Quality Indicators in Coding	342
10.2.7 Summary of Coding	344
10.3 Entering Numeric Data into Files	344
10.4 Editing	345
10.5 Weighting	347
10.5.1 Weighting with a First-Stage Ratio Adjustment	348
10.5.2 Weighting for Differential Selection Probabilities	349
10.5.3 Weighting to Adjust for Unit Nonresponse	350
10.5.4 Poststratification Weighting	352
10.5.5 Putting All the Weights Together	352
10.6 Imputation for Item-Missing data	354
10.7 Sampling Variance Estimation for Complex Samples	359
10.8 Survey Data Documentation and Metadata	363
10.9 Summary	365
Keywords	366
For More In-Depth Reading	367
Exercises	367

CHAPTER 11. PRINCIPLES AND PRACTICES RELATED TO ETHICAL RESEARCH

11.1 Introduction	371
11.2 Standards for the Conduct of Research	371
11.3 Standards for Dealing with Clients	374
11.4 Standards for Dealing with the Public	375
11.5 Standards for Dealing with Respondents	376
11.5.1 Legal Obligations to Survey Respondents	376
11.5.2 Ethical Obligations to Respondents	377
11.5.3 Informed Consent: Respect for Persons	379
11.5.4 Beneficence: Protecting Respondents from Harm	381
11.5.5 Efforts at Persuasion	383
11.6 Emerging Ethical Issues	384
11.7 Research About Ethical Issues in Surveys	384
11.7.1 Research on Informed Consent Protocols	385
11.7.2 Research on Confidentiality Assurances and Survey Participation	390
11.8 Administrative and Technical Procedures for Safe-Guarding Confidentiality	392
11.8.1 Administrative Procedures	392
11.8.2 Technical Procedures	393
11.9 Summary and Conclusions	398

Keywords	400
For More In-Depth Reading	400
Exercises	400

CHAPTER 12. FAQs ABOUT SURVEY METHODOLOGY

12.1 Introduction	405
12.2 The Questions and Their Answers	405

REFERENCES	421
------------	-----

INDEX	451
-------	-----

PREFACE TO THE FIRST EDITION

We wrote this book with a specific purpose in mind. We are all survey methodologists—students of the theories and practices of the various data collection and analysis activities that are called “survey research.” Surveys (in a form that would be recognizable today) are approximately 60–80 years old. Over the past two decades, a set of theories and principles has evolved that offer a unified perspective on the design, conduct, and evaluation of surveys. This perspective is most commonly labeled the “total survey error” paradigm. The framework guides modern research about survey quality and shapes how practicing survey professionals approach their work. The field arising out of this research domain can appropriately be called “survey methodology.”

We increasingly noticed a mismatch, however, between the texts related to surveys and how the science of surveys was evolving. Many survey research texts focused on the application of tools and deemphasized the theories and science underlying those tools. Many texts told students to do things that were no longer or never supported by the methodological research in the area. In short, there were books that emphasized “how to do” surveys but neglected the science underlying the practices that were espoused.

Most harmful we thought was the impression conveyed to those who read the texts that surveys were merely a recipe-like task; if step-by-step instructions were followed, high quality would be guaranteed. In contrast, we saw surveys as requiring the implementation of principles in unique ways to fit a particular substantive purpose for a particular target population.

These issues became particularly important to us when the demand for a one semester graduate level (and senior undergraduate level) course became obvious at the Joint Program in Survey Methodology (JPSM), a consortium graduate program funded by the U.S. Federal statistical agencies in which the authors teach. The students would often have advanced education in another field (e.g., economics, statistics, or psychology) but no formal exposure to the field of survey methodology. We planned a 14-week lecture course with exercises and examinations that began in the Fall of 1998, and we immediately suffered from the absence of a text that could accompany the lectures and motivate the exercises.

We began to envision a text describing the basic principles of survey design discovered in methodological research over the past years and the guidance they offered for decisions that are made in the execution of good quality surveys. We wanted to include exercises that would help integrate an understanding of the field. We wanted to convey that the field is based on experimental and other research findings and that practical survey design was not a mere matter of judgment and opinion but rather the result of a body of research findings.

We drafted this book over several years. After we wrote the first couple of chapters, we hit a dry spell, which was ended when our colleague Nancy

Mathiowetz kicked us back in gear. We appreciated her energy in getting us going again.

The manuscript profited greatly from the critique of our student colleagues. The text had a dry run in the Summer of 2003 in a class at the University of Michigan Survey Research Center Summer Institute in Survey Research Techniques, entitled “Introduction to Survey Research Techniques,” taught by Maria Krysan and Sue Ellen Hansen. We thank these instructors for helping improve the manuscript. We learned much from the criticisms and ideas of both Krysan and Hansen and the students in the class: Nike Adebisi, Jennifer Bowers, Scott Compton, Sanjay Kumar, Dumile Mkhwanazi, Hanne Muller, Vuyelwa Nkambule, Laurel Park, Aaron Russell, Daniel Spiess, Kathleen Stack, Kimiko Tanaka, Dang Viet Phuong, and Christopher Webb.

It is fair to say that this book strongly reflects the lessons taught by many of our own mentors. One deserves special mention. All of the authors were friends and students of Charlie Cannell (some formally; all informally). Charles F. Cannell began his survey career with Rensis Likert at the U.S. Department of Agriculture, Division of Program Surveys. Cannell later joined Likert and others in founding the University of Michigan Survey Research Center in 1946. He was the first director of field operations at the Center and had a long and distinguished career in survey methodology. In memory of Charlie and his work, the Institute for Social Research (the larger institute of which the SRC is part) established the Charles F. Cannell Fund in Survey Methodology. All royalties that result from the sales of this text will be contributed to this fund. The endowment from the fund is designated for support of young scholars developing their research careers in survey methodology. We can think of no better use.

We designed the text to be used in a class where the participants had taken one or more courses in statistics. The key relevant skill is the reading of statistical notation, including summation signs, notation for expected values, and simple algebraic manipulation of summed quantities. Some chapters present quantitative analyses using regression and logistic regression models, and students unfamiliar with linear modeling need some help in understanding these results.

This book has 12 chapters, in the order in which they are presented in the semester-length course on survey methodology called “Fundamentals of Survey Methodology” at the JPSM. We envision that instructors will want to assign additional readings, often from one or more review articles referenced in the chapters.

The first two chapters (“Introduction to Survey Methodology” and “Inference And Error In Surveys”) are conceptual in nature. Chapter 1 presents six example surveys that are used throughout the book to illustrate various principles and practices. The instructor can supplement the text by displaying the Web pages of these surveys in the class and leading class discussions about the key design features and products of the surveys.

The second chapter presents the key components of the total survey error paradigm. Again, at this early stage in the class, we have found that providing students with examples of key error components by referencing the example surveys aids in student understanding. A defining characteristic of surveys as we see them is that they are designed to produce statistical descriptions of populations. Although there are computer programs that will calculate statistics, we think it is critical that a survey methodologist understand the calculations that underlie those statistics. Hence, the book routinely presents statistical notation along with a conceptual discussion of what is being calculated.

The treatment of Chapter 2 would be a good time to devote a class to statistical notation, which, once it is learned, will help the students be more comfortable throughout the rest of the book.

Starting with Chapter 3 (“Target Populations, Sampling Frames, and Coverage Error”), each chapter deals with a different component of total survey error and the methodological research discoveries that guide best practices. The focus of these chapters is deliberately the research on which best practices in survey research are based. We have often found that students beginning the study of survey methodology have the perspective that their opinions on a specific design feature are diagnostic of the best practices. The material that is presented in Chapters 3–11 attempts to show that there are scientific studies of survey methods that inform best practice; opinions are of little value unless they are research-based. Some of these studies do not have intuitively obvious findings. Hence, a student of the field must review the past methodological literature and at times do novel research to determine good design. There are two devices in the text that can help convey this perspective of the field. One is the set of embedded references to research in the discussions. The other is the presentation of illustrative boxes that give short descriptions of classic research in the domain covered in the chapter. These are summaries that describe the design, findings, limitations, and impact of the research. The full articles on this research can be used as supplementary readings, which could be discussed in class. There are also suggested supplementary readings at the end of each chapter.

Chapter 4 (“Sample Design and Sampling Error”) uses more statistical notation than most of the other chapters. When many participants in the course need remedial instruction in reading and understanding statistical notation, we have referred them to the small monograph by Kalton, *An Introduction to Survey Sampling* (Sage, 1983). In some editions of the course, we have spent three weeks on the coverage and sampling chapters.

Each of Chapters 5–10 is normally covered in one week of the course. We have found it useful to emphasize the parallels between equations expressing coverage error and nonresponse error. We have also emphasized how the basic principles of intraclass correlations apply both to sample clustering effects and interviewer variance.

Chapter 11 (“Principles and Practices Related to Scientific Integrity”) is included not just as sensitivity training but because it includes both conceptual frameworks underlying ethical treatment of human subjects and also recent theory and practice regarding disclosure analysis of survey data. Again, we describe how research, as well as judgement, can affect decisions related to ethical issues.

We wrote Chapter 12 (“FAQs About Survey Methodology”) in a very different style. It is a tradition in the course, in a review session prior to the final examination, to have an open question section. At this time, we found students asking the kind of questions that come from attempts to integrate their learning of specific lessons with their larger worldview. Hence, we constructed a “frequently asked questions” format including those global questions and offering our answers to them.

The manuscript was greatly improved by the editorial wisdom of Sarah Dipko and Sonja Ziniel. Adam Kelley assisted in computer-based processing of figures and tables. Lisa Van Horn at Wiley is a production editor with a wonderful sense of when intervention is needed and when it isn't. We thank them all.

It was great fun writing this book, assembling our views on key research

areas, and debating how to convey the excitement of survey methodology as an area of knowledge. We hope you have as much fun as we did.

Ann Arbor, Michigan
Boston, Massachusetts
Ann Arbor, Michigan
Ann Arbor, Michigan
Ann Arbor, Michigan
College Park, Maryland
March 2004

ROBERT M. GROVES
FLOYD J. FOWLER, JR.
MICK P. COUPER
JAMES M. LEPKOWSKI
ELEANOR SINGER
ROGER TOURANGEAU

PREFACE TO THE SECOND EDITION

We have been pleased by the acceptance of the first edition of *Survey Methodology*. It has now been used by instructors around the world and has been translated into several other languages. Some of these instructors and their students have graciously pointed out weaknesses and errors in some sections of the text. Some of them gave us great ideas to improve the text.

In addition, as survey methodologists actively conducting research in the field, we became increasingly aware that some of the lessons in the text were becoming out of date. This was most true of sections of the book that concern the role of survey nonresponse in the quality of survey estimates and the rapidly evolving new modes of data collection.

For those reasons, we assembled the group of coauthors and agreed to update parts of chapters that could most profit from changes. As the reader will see, there is increased discussion of sampling frame issues for mobile telephone and web surveys in Chapter 3. There is an integration of some of the example surveys into the presentation of sample designs in Chapter 4, along with a new section on selection of persons within households. The changes in Chapter 5 update the findings on mobile phone and web surveys. Chapter 6, on survey nonresponse, is radically changed, reflecting new insights into how nonresponse rates and nonresponse errors relate to one another. Chapter 8, on evaluating survey questions, highlights new research findings on effective questionnaire development techniques. Chapter 11, on ethical issues in survey research, is reorganized to emphasize the growing research results on privacy, informed consent, and confidentiality issues. The remaining chapters provide the reader with more recent methodological research findings, especially when they expand our understanding of survey errors. The chapters have about 50% more exercises, following feedback from instructors that such additions would benefit their use of the text.

Two assistants labored over this edition's manuscript: Michael Guterbock and Kelly Smid. Some Ph.D. students at Michigan read draft chapters (Ashley Bowers, Matthew Jans, Courtney Kennedy, Joe Sakshaug, and Brady West). When we signed the contract with Wiley, we demanded that Lisa Van Horn continue as our production editor. All of the above went beyond the call to make this edition a success. We thank them.

As with the last edition, we want to use the royalties of the text to help persons newly entering the field of survey methodology. They will be given to the Rensis Likert Fund for Research on Survey Methodology, which directly benefits graduate students in survey methodology.

Ann Arbor, Michigan
Boston, Massachusetts
Ann Arbor, Michigan
Ann Arbor, Michigan
Ann Arbor, Michigan
College Park, Maryland
March, 2009

ROBERT M. GROVES
FLOYD J. FOWLER, JR.
MICK P. COUPER
JAMES M. LEPKOWSKI
ELEANOR SINGER
ROGER TOURANGEAU

ACKNOWLEDGMENTS

Reprinted tables and figures in *Survey Methodology* and their copyright holders are listed below. The authors appreciate permission to adapt or reprint them.

Figure 1.5c, from Mokdad, Ford, Bowman, Dietz, Vinicor, Bales, and Marks (2003) with permission of the American Medical Association. Copyright © 2003.

Table 5.1 from Groves (1989) with permission of John Wiley and Sons. Copyright © 1989.

Figure 6.5 from Groves (2006) and text surrounding from Groves and Peytcheva (2008) with permission of the American Association for Public Opinion Research. Copyright © 2006, 2008.

Table in box on page 185, from Merkle and Edelman in Groves, Dillman, Eltinge, and Little (2002) with permission of John Wiley and Sons. Copyright © 2002.

Figure 7.2, from Tourangeau, Rips, and Rasinski (2000) reprinted with permission of Cambridge University Press. Copyright © 2000.

Box on page 236, from Schwarz, Hippler, Deutsch, and Strack (1985) reprinted with permission of the American Association for Public Opinion Research. Copyright © 1985.

Figure 7.3, from Jenkins and Dillman in Lyberg, Biemer, Collins, de Leeuw, Dippo, Schwarz, and Trewin (1997) reprinted with permission of John Wiley and Sons. Copyright © 1997.

Box on p. 268, from Oksenberg, Cannell, and Kalton (1991) reprinted with permission of Statistics Sweden. Copyright © 1991.

Box on p. 293, from Schuman and Converse (1971) reprinted with permission of the American Association for Public Opinion Research. Copyright © 1971.

Box on p. 299, from Kish (1962) reprinted with permission of the American Statistical Association. Copyright © 1962.

Table 9.2, from Fowler and Mangione (1990) reprinted with permission of Sage Publications. Copyright © 1990.

Table 10.5, from Campanelli, Thomson, Moon, and Staples in Lyberg, Biemer, Collins, de Leeuw, Dippo, Schwarz, and Trewin (1997) reprinted with permission of John Wiley and Sons. Copyright © 1997.

CHAPTER ONE

AN INTRODUCTION TO SURVEY METHODOLOGY

A Note to the Reader

You are about to be exposed to a system of principles called “survey methodology” for collecting information about the social and economic world. We have written this book in an attempt to describe the excitement of designing, conducting, analyzing, and evaluating sample surveys. To appreciate this fully, use the devices we have placed in each chapter to enrich your memory of the material. Throughout the book, you will see boxes with illustrations and examples of key principles, terminology notes, and highlights of classic research studies in the field. In the outside margin of each page you will find key terms, at the point where they are defined. At the end of each chapter is a set of exercises that you can use to test your understanding of that chapter’s material. The best strategy is to read the text through, then, at the end of each chapter, go back, read the boxes, and review the key terms.

At 8:30 AM on the day before the first Friday of each month, a group of economists and statisticians enter a soundproof and windowless room in a building at 2 Massachusetts Avenue, NE, in Washington, DC, USA. Once those authorized are present, the room is sealed.

Those in the room are professional staff of the U.S. Bureau of Labor Statistics (BLS), and their task is to review and approve a statistical analysis of key economic data. Indeed, they have spent the week poring over sets of numbers, comparing them, examining indicators of their qualities, looking for anomalies, and writing drafts of a press release describing the numbers. They write the press release in simple language, understandable by those who have no technical knowledge about how the numbers were produced.

At 8:00 AM the next day, a group of journalists assemble in a monitored room in the nearby main Department of Labor building, removed from any contact with the outside world. The BLS staff enter the room and then reveal the results to the journalists. The journalists immediately prepare news stories based on the briefing. At exactly 8:30 AM, they simultaneously electronically transmit their stories to their news organizations and sometimes telephone editors and producers.

The statistics revealed are the unemployment rate of the prior month and the number of jobs created in the prior month. The elaborate protections and security used prior to their release stem from the enormous impact the numbers can have

on society. Indeed, in months when the numbers signal important changes in the health of the U.S. economy, thousands of stock market investors around the world make immediate buy and sell decisions. Within 45 minutes of the announcement, trillions of dollars can move in and out of markets around the world based on the two numbers revealed at 8:30 AM.

Both the unemployment rate and the jobs count result from statistical surveys. A household survey produces the unemployment rate; an employer survey, the jobs count. The households and employers surveyed have been carefully selected so that their answers, when summarized, reflect the answers that would be obtained if the entire population were questioned. In the surveys, thousands of individual people answer carefully phrased questions about their own or their company's attributes. In the household survey, professional interviewers ask the questions and enter the answers onto laptop computers. In the employer survey, the respondents complete a standardized questionnaire either on paper or electronically. Complex data processing steps follow the collection of the data, to assure internal integrity of the numbers.

These two numbers have such an impact because they address an important component of the health of the nation's economy, and they are credible. Macroeconomic theory and decades of empirical results demonstrate their importance. However, only when decision makers believe the numbers do they gain value. This is a book about the process of generating such numbers through statistical surveys and how survey design can affect the quality of survey statistics. In a real sense, it addresses the question of when numbers from surveys are credible and when they are not.

1.1 INTRODUCTION

This chapter is an introduction to survey methodology as a field of knowledge, as a profession, and as a science. The initial sections of the chapter define the field so that the reader can place it among others. At the end of the chapter, readers will have a sense of what survey methodology is and what survey methodologists do.

survey

A "survey" is a systematic method for gathering information from (a sample of) entities for the purposes of constructing quantitative descriptors of the attributes of the larger population of which the entities are members. The word "systematic" is deliberate and meaningfully distinguishes surveys from other ways of gathering information. The phrase "(a sample of)" appears in the definition because sometimes surveys attempt to measure everyone in a population and sometimes just a sample.

statistic

descriptive
statistic

analytic
statistic

The quantitative descriptors are called "statistics." Statistics are quantitative summaries of observations on a set of elements. Some are "descriptive statistics," describing the size and distributions of various attributes in a population (e.g., the mean years of education of persons, the total number of persons in the hospital, the percentage of persons supporting the president). Others are "analytic statistics," measuring how two or more variables are related (e.g., a regression coefficient describing how much increases in income are associated with increases in years of education; a correlation between education and number of books read in the last year). That goal sets surveys apart from other efforts to describe people or events. The statistics attempt to describe basic characteristics or experiences of large and small populations in our world.

Almost every country in the world uses surveys to estimate their rate of unemployment, basic prevalence of immunization against disease, opinions about the central government, intentions to vote in an upcoming election, and people's satisfaction with services and products that they buy. Surveys are a key tool in tracking global economic trends, the rate of inflation in prices, and investments in new economic enterprises. Surveys are one of the most commonly used methods in the social sciences to understand the way societies work and to test theories of behavior. In a very real way, surveys are a crucial building block in a modern information-based society.

Although a variety of activities are called surveys, this book focuses on surveys that have the following characteristics:

- 1) Information is gathered primarily by asking people questions.
- 2) Information is collected either by having interviewers ask questions and record answers or by having people read or hear questions and record their own answers.
- 3) Information is collected from only a subset of the population to be described—a sample—rather than from all members.

Since “ology” is Greek for “the study of,” survey methodology is the study of survey methods. It is the study of sources of error in surveys and how to make the numbers produced by surveys as accurate as possible. Here the word “error” refers to deviations or departures from the desired outcome. In the case of surveys, “error” is used to describe deviations from the true values applicable to the population studied. Sometimes, the phrase “statistical error” is used to differentiate this meaning from a reference to simple mistakes.

survey
methodology

error
statistical error

The way each of the above steps is carried out—which questions are asked, how answers are collected, and which people answer the questions—can affect the quality (or error properties) of survey results. This book will describe how to conduct surveys in the real world and how to evaluate the quality of survey results. It will describe what is known, and not known, about how to minimize error in survey statistics. Most of all, this book will attempt to distill the results of 100 years of scientific studies that have defined the theories and principles, as well as practices, of high-quality survey research.

1.2 A BRIEF HISTORY OF SURVEY RESEARCH

Converse (1987) has produced an important account of the history of survey research in the United States, and we recount some of the highlights here. There are four perspectives on surveys that are worth describing: the purposes to which surveys were put, the development of question design, the development of sampling methods, and the development of data collection methods.

1.2.1 The Purposes of Surveys

Perhaps the earliest type of survey is the census, generally conducted by governments. Censuses are systematic efforts to count an entire population, often for

census

Schuman (1997) on “Poll” Versus “Survey”

What is the difference between a poll and a survey? The word “poll” is most often used for private-sector opinion studies, which use many of the same design features as studies that would be called “surveys.” “Poll” is rarely used to describe studies conducted in government or scientific domains. There are, however, no clear distinctions between the meanings of the two terms. Schuman notes that the two terms have different roots: “Poll” is a four letter word, generally thought to be from an ancient Germanic term referring to ‘head,’ as in counting heads. The two-syllable word ‘survey,’ on the other hand, comes from the French *survee*, which in turn derives from Latin *super* (over) and *videre* (to look). The first is therefore an expression with appeal to a wider public, the intended consumers of results from Gallup, Harris, and other polls. The second fits the needs of academicians in university institutes who wish to emphasize the scientific or scholarly character of their work.” (page 7)

purposes of taxation or political representation. In the United States, the Constitution stipulates that a census must be conducted every ten years, to reapportion the House of Representatives reflecting current population residence patterns. This gives the statistics from a census great political import. Because of this, they are often politically contentious (Anderson, 1990).

A prominent early reason for surveys was to gain understanding of a social problem. Some people trace the origins of modern survey research to Charles Booth, who produced a landmark study titled *Life and Labour of the People of London* (1889–1903) (<http://booth.lse.ac.uk/>). As Converse recounts it, Booth spent his own money to collect voluminous data on the poor in London and the reasons why they were poor. He wrote at least 17 volumes based on the data he collected. He did not use methods like the ones we use today—no well-defined sampling techniques, no standardized questions. Indeed, interviewer observation and inference produced much of the information. However, the Booth study used quantitative summaries from systematic measurements to understand a fundamental societal problem.

In contrast to studies of social problems, journalism and market research grew to use surveys to gain a systematic view of “the man on the street.” A particular interest was reactions to political leaders and preferences in upcoming

elections. That interest led to the development of modern public opinion polling.

In a related way, market research sought knowledge about reactions of “real” people to existing and planned products or services. As early as the 1930s, there was serious research on what programs and messages delivered via the radio would be most popular. The researchers began to use surveys of broader samples to produce information more useful to commercial decision makers.

Over the early 20th century, public opinion polling and market research, sometimes done by the same companies, evolved to use mail surveys and telephone surveys. They often sampled from available lists, such as telephone, driver’s license, registered voter, or magazine subscriber listings. They collected their data primarily by asking a fixed set of questions; observations by interviewers and proxy reporting of other people’s situations were not part of what they needed. These features were directly tied to the most important difference between what they were doing and what those who had gone before had done; rather than collecting data about facts and objective characteristics of people, the polling and market research surveyors were interested in what people knew, felt, and thought.

The measurement of attitudes and opinions is a key foundation of the modern management philosophies that place much weight on customer satisfaction.

Customer satisfaction surveys measure expectations of purchasers about the quality of a product or service and how well their expectations were met in specific transactions. Such surveys are ubiquitous tools of management to improve the performance of their organizations.

Politicians and political strategists now believe that opinion polls are critical to good decisions on campaign strategy and messages to the public about important issues. Indeed, a common criticism of modern politicians is that they rely too heavily on polling data to shape their personal opinions, choosing to reflect the public's views rather than provide leadership to the public about an issue.

1.2.2 The Development of Standardized Questioning

The interest in measuring subjective states (i.e., characteristics that cannot be observed, internalized within a person) also had the effect of focusing attention on question wording and data collection methods. When collecting factual information, researchers had not thought it important to carefully word questions. Often, interviewers were sent out with lists of objectives, such as age, occupation, and education, and the interviewers would decide on how the questions would be worded. Experienced researchers often did the interviewing, with great confidence that they knew how to phrase questions to obtain good answers.

However, the market research and polling organizations were doing large numbers of interviews, using newly hired people with no special background in the social sciences. Of necessity, researchers needed to specify more carefully the information sought by the survey. Further, researchers found that small changes in wording of an attitude question sometimes had unusually large effects on the answers.

Thus, early in the development of opinion surveys, attention began to be paid to giving interviewers carefully worded questions that they were to ask exactly the same way for each interview. Also, as interviewers were used more to ask questions, it was found that how they asked questions and recorded answers could affect the results. This led eventually to researchers training and supervising interviewers more formally than earlier.

Question wording also was influenced as the academics started to pay some attention to what the commercial researchers were doing. Psychometricians, psychologists who quantify psychological states, had been interested in how to put meaningful numbers on subjective states. Measuring intelligence was the first effort in this direction. However, people such as Thurstone also worked on how to assign numbers to attitudes, feelings, and ratings (e.g., Thurstone and Chave, 1929).

For the most part, their approaches were extremely cumbersome and were used primarily when they could get captive college student volunteers to fill out lengthy, highly redundant questionnaires. Such instruments were not going to be useful for most survey interviews with representative samples; they took too long to measure one or a few attitudes. Rensis Likert in his PhD dissertation (Likert, 1932), however, demonstrated that a single, streamlined question, with a scaled set of answers, could accomplish much the same thing as a lengthy series of paired comparisons. Likert applied the work to surveys (and later founded the University of Michigan Survey Research Center in 1946).

1.2.3 The Development of Sampling Methods

Early researchers, such as Booth, essentially tried to collect data on every element of a defined population. Such censuses avoided problems of errors arising from measuring just a subset of the population, but were clearly impractical for large populations. Indeed, the difficulty of analyzing complete census data led to early efforts to summarize a census by taking a sample of returns. Early efforts to sample would study a “typical” town, or they would purposively try to collect individuals to make the samples look like the population—for example, by interviewing about half men and half women, and trying to have them distributed geographically in a way that is similar to the population.

probability
sample

Although the theory of probability was established in the 18th century, its application to practical sample survey work was largely delayed until the 20th century. The first applications were the taking of a “1 in N ” systematic selection from census returns. These were “probability samples”; that is, every record had a known nonzero chance of selection into the sample.

area
probability
sample

A big breakthrough in sampling came from people who did research on agriculture. In order to predict crop yields, statisticians had worked out a strategy they called “area probability sampling.” This is just what it sounds like: they would sample areas or plots of land and find out what farmers were doing with those plots in the spring (for example, if they were planting something on them and, if so, what) in order to project what the fall crops would look like. The same technique was developed to sample households. By drawing samples of geographic blocks in cities or tracts of land in rural areas, listing the housing units on the blocks or rural tracts, then sampling the housing units that were listed, samplers found a way to give all households and, by extension, the people living in them, a chance to be sampled. The attraction of this technique included the elimination of the need for a list of all persons or all households in the population prior to drawing the sample.

The Depression and World War II were major stimuli for survey research. One of the earliest modern probability samples was drawn for the Monthly Survey of Unemployment, starting in December, 1939, led by a 29-year-old statistician, Morris Hansen, who later became a major figure in the field (Hansen, Hurwitz, and Madow, 1953). During the war, the federal government became interested in conducting surveys to measure people’s attitudes and opinions, such as interest in buying war bonds, as well as factual information. Considerable resources were devoted to surveys during the war, and researchers who were recruited to work with the government during the war later came to play critical roles in the development of survey methods. When the war was over, methodologists understood that in order to produce good population-based statistics it was necessary to attend to three aspects of survey methodology: how questions were designed; how the data were collected, including the training of interviewers; and how samples were drawn.

Probability samples are the standard by which other samples are judged. They are routinely used by almost all government statistical agencies when data are used to provide important information for policy makers. They are used for surveys used in litigation. They are used for measurement of media audience sizes, which in turn determine advertising rates. In short, whenever large stakes ride on the value of a sample, probability sampling is generally used.