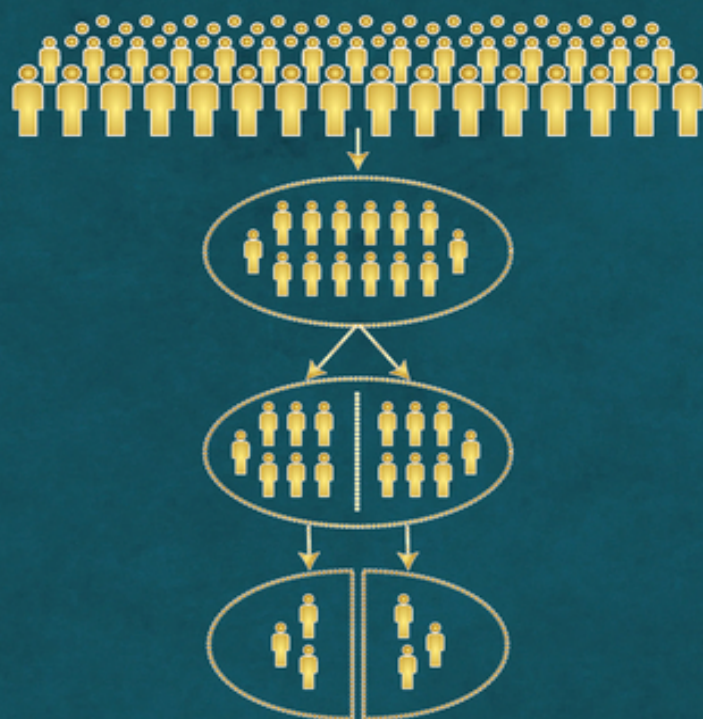


WILEY SERIES IN SURVEY METHODOLOGY

# Experimental Methods in Survey Research

Techniques that Combine Random Sampling  
with Random Assignment



Edited by

Paul J. Lavrakas ■ Michael W. Traugott ■ Courtney Kennedy  
Allyson L. Holbrook ■ Edith D. de Leeuw ■ Brady T. West



WILEY



**Experimental Methods in Survey Research**

## **WILEY SERIES IN SURVEY METHODOLOGY**

Established in Part by Walter A. Shewhart and Samuel S. Wilks

Editors: Mick P. Couper, Graham Kalton, Lars Lyberg, J. N. K. Rao, Norbert Schwarz,  
Christopher Skinner

Editor Emeritus: Robert M. Groves

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

# Experimental Methods in Survey Research

Techniques that Combine Random Sampling with Random Assignment

*Edited by*

*Paul J. Lavrakas*

NORC, IL, US

*Michael W. Traugott*

University of Michigan, MI, US

*Courtney Kennedy*

Pew Foundation, DC, US

*Allyson L. Holbrook*

University of Illinois-Chicago, IL, US

*Edith D. de Leeuw*

University of Utrecht, NL, TC

*Brady T. West*

University of Michigan, MI, US

**WILEY**

This edition first published 2019  
© 2019 John Wiley & Sons, Inc.

All rights reserved. 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 or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Paul J. Lavrakas, Michael W. Traugott, Courtney Kennedy, Allyson L. Holbrook, Edith D. de Leeuw, and Brady T. West to be identified as the authors of the editorial material in this work has been asserted in accordance with law.

*Registered Office*

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

*Editorial Office*

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at [www.wiley.com](http://www.wiley.com).

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

*Limit of Liability/Disclaimer of Warranty*

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

*Library of Congress Cataloging-in-Publication Data*

Names: Lavrakas, Paul J., editor.

Title: Experimental methods in survey research : techniques that combine random sampling with random assignment / edited by Paul J. Lavrakas, [and five others].

Description: New Jersey : Wiley, 2020. | Series: Wiley series in survey methodology | Includes bibliographical references and index. |

Identifiers: LCCN 2019015585 (print) | LCCN 2019021945 (ebook) | ISBN 9781119083757 (Adobe PDF) | ISBN 9781119083764 (ePub) | ISBN 9781119083740 (hardback)

Subjects: LCSH: Social surveys. | Social sciences--Research--Methodology. | Surveys--Evaluation. | BISAC: SOCIAL SCIENCE / Statistics.

Classification: LCC HM538 (ebook) | LCC HM538 .E97 2020 (print) | DDC 300.72/3--dc23

LC record available at <https://lccn.loc.gov/2019015585>

Cover design: Wiley

Cover image: © Angelina Babii/Shutterstock

Set in 10/12pt WarnockPro by SPi Global, Chennai, India

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

*We dedicate this book with gratitude to the contributions made to the use of experimental design by Donald T. Campbell, Thomas D. Cook, Robert Boruch, Marilyn B. Brewer, Jeanne Converse, Stephen Fienberg, Stanley Presser, Howard Schuman, and Eleanor Singer, and to their influence on our own usage of survey-based experiments.*

## Contents

**List of Contributors** *xix*

**Preface by Dr. Judith Tanur** *xxv*

**About the Companion Website** *xxix*

<b>1</b>	<b>Probability Survey-Based Experimentation and the Balancing of Internal and External Validity Concerns</b>	<b>1</b>
	<i>Paul J. Lavrakas, Courtney Kennedy, Edith D. de Leeuw, Brady T. West, Allyson L. Holbrook, and Michael W. Traugott</i>	
1.1	Validity Concerns in Survey Research	3
1.2	Survey Validity and Survey Error	5
1.3	Internal Validity	6
1.4	Threats to Internal Validity	8
1.4.1	Selection	8
1.4.2	History	9
1.4.3	Instrumentation	10
1.4.4	Mortality	10
1.5	External Validity	11
1.6	Pairing Experimental Designs with Probability Sampling	12
1.7	Some Thoughts on Conducting Experiments with Online Convenience Samples	12
1.8	The Contents of this Book	15
	References	15

### **Part I Introduction to Section on Within-Unit Coverage** 19

*Paul J. Lavrakas and Edith D. de Leeuw*

<b>2</b>	<b>Within-Household Selection Methods: A Critical Review and Experimental Examination</b>	<b>23</b>
	<i>Jolene D. Smyth, Kristen Olson, and Mathew Stange</i>	
2.1	Introduction	23
2.2	Within-Household Selection and Total Survey Error	24
2.3	Types of Within-Household Selection Techniques	24
2.4	Within-Household Selection in Telephone Surveys	25
2.5	Within-Household Selection in Self-Administered Surveys	26

2.6	Methodological Requirements of Experimentally Studying Within-Household Selection Methods	27
2.7	Empirical Example	30
2.8	Data and Methods	31
2.9	Analysis Plan	34
2.10	Results	35
2.10.1	Response Rates	35
2.10.2	Sample Composition	35
2.10.3	Accuracy	37
2.11	Discussion and Conclusions	40
	References	42
<b>3</b>	<b>Measuring Within-Household Contamination: The Challenge of Interviewing More Than One Member of a Household</b>	<b>47</b>
	<i>Colm O'Muircheartaigh, Stephen Smith, and Jaclyn S. Wong</i>	
3.1	Literature Review	47
3.2	Data and Methods	50
3.2.1	The Partner Study Experimental Design and Evaluation	52
	Investigators	53
	Field/Project Directors	53
3.2.1.1	Effects on Response Rates and Coverage	54
3.3	The Sequence of Analyses	55
3.4	Results	55
3.4.1	Results of Experimental Group Assignment	55
3.5	Effect on Standard Errors of the Estimates	57
3.6	Effect on Response Rates	58
3.6.1	Response Rates Among Primes	58
3.6.2	Response Rates Among Partners	60
3.7	Effect on Responses	61
3.7.1	Cases Considered in the Analyses	61
3.7.2	Sequence of Analyses	61
3.7.3	Dependent Variables of Interest	61
3.7.4	Effect of Experimental Treatment on Responses of Primes	61
3.7.4.1	Missingness	61
3.7.4.2	Substantive Answers	62
3.7.5	Changes in Response Patterns Over Time Among Primes	62
3.7.5.1	Missingness	62
3.7.5.2	Substantive Answers	62
3.7.5.3	Social Desirability	62
3.7.5.4	Negative Questions	62
3.7.5.5	Positive Questions	63
3.7.5.6	Effect on Responses of Partners	63
3.7.6	Conclusion	63
3.7.6.1	Design	63
3.7.6.2	Statistical Considerations	64
3.8	Substantive Results	64
	References	64

**Part II Survey Experiments with Techniques to Reduce  
Nonresponse 67**

*Edith D. de Leeuw and Paul J. Lavrakas*

- 4 Survey Experiments on Interactions and Nonresponse: A Case Study of  
Incentives and Modes 69**  
*A. Bianchi and S. Biffignandi*
- 4.1 Introduction 69
- 4.2 Literature Overview 70
- 4.3 Case Study: Examining the Interaction Between Incentives and Mode 73
- 4.3.1 The Survey 73
- 4.3.2 The Experiment 73
- 4.3.3 Data and Methods 75
- 4.3.4 Results of the Experiment: Incentives and Mode Effects 77
- 4.3.4.1 Effects on Participation 77
- 4.3.4.2 Effects on Subgroups 78
- 4.3.4.3 Effects on Data Quality 80
- 4.3.4.4 Effects on Costs 81
- 4.4 Concluding Remarks 83
- Acknowledgments 85
- References 86
- 5 Experiments on the Effects of Advance Letters in Surveys 89**  
*Susanne Vogl, Jennifer A. Parsons, Linda K. Owens, and Paul J. Lavrakas*
- 5.1 Introduction 89
- 5.1.1 Why Advance Letters? 89
- 5.1.2 What We Know About Effects of Advance Letters? 90
- 5.1.2.1 Outcome Rates 90
- 5.1.2.2 Nonresponse Bias 91
- 5.1.2.3 Cost Effectiveness 91
- 5.1.2.4 Properties of Advance Letters 93
- 5.2 State of the Art on Experimentation on the Effect of Advance Letters 93
- 5.3 Case Studies: Experimental Research on the Effect of Advance Letters 95
- 5.4 Case Study I: Violence Against Men in Intimate Relationships 96
- 5.4.1 Sample Design 96
- 5.4.2 Fieldwork 97
- 5.4.3 Sampling Bias 97
- 5.4.4 Analytical Strategies 97
- 5.4.5 Results 98
- 5.4.5.1 Effect of Advance Letters on Outcome Rates and Reasons for Refusals 98
- 5.4.5.2 Recruitment Effort 98
- 5.4.5.3 Effect of Advance Letters on Reporting on Sensitive Topics 98
- 5.4.6 Discussion 99
- 5.5 Case Study II: The Neighborhood Crime and Justice Study 100
- 5.5.1 Study Design 100
- 5.5.2 Experimental Design 100
- 5.5.3 Fieldwork 101
- 5.5.4 Analytical Strategies 102
- 5.5.5 Results 103

- 5.5.6 Discussion 103
- 5.6 Discussion 106
- 5.7 Research Agenda for the Future 107
- References 108

### **Part III Overview of the Section on the Questionnaire 111**

*Allyson Holbrook and Michael W. Traugott*

- 6 Experiments on the Design and Evaluation of Complex Survey Questions 113**  
*Paul Beatty, Carol Cosenza, and Floyd J. Fowler Jr.*
  - 6.1 Question Construction: Dangling Qualifiers 115
  - 6.2 Overall Meanings of Question Can Be Obscured by Detailed Words 117
  - 6.3 Are Two Questions Better than One? 119
  - 6.4 The Use of Multiple Questions to Simplify Response Judgments 121
  - 6.5 The Effect of Context or Framing on Answers 122
    - 6.5.1 Alcohol 123
    - 6.5.2 Guns 123
  - 6.6 Do Questionnaire Effects Vary Across Sub-groups of Respondents? 124
  - 6.7 Discussion 126
  - References 128
- 7 Impact of Response Scale Features on Survey Responses to Behavioral Questions 131**  
*Florian Keusch and Ting Yan*
  - 7.1 Introduction 131
  - 7.2 Previous Work on Scale Design Features 132
    - 7.2.1 Scale Direction 132
    - 7.2.2 Scale Alignment 133
    - 7.2.3 Verbal Scale Labels 133
  - 7.3 Methods 134
  - 7.4 Results 136
  - 7.5 Discussion 141
    - Acknowledgment 143
    - 7.A Question Wording 143
      - 7.A.1 Experimental Questions (One Question Per Screen) 143
      - 7.A.2 Validation Questions (One Per Screen) 144
      - 7.A.3 GfK Profile Questions (Not Part of the Questionnaire) 145
    - 7.B Test of Interaction Effects 145
    - References 146
- 8 Mode Effects Versus Question Format Effects: An Experimental Investigation of Measurement Error Implemented in a Probability-Based Online Panel 151**  
*Edith D. de Leeuw, Joop Hox, and Annette Scherpenzeel*
  - 8.1 Introduction 151
    - 8.1.1 Online Surveys Advantages and Challenges 151
    - 8.1.2 Probability-Based Online Panels 152
    - 8.1.3 Mode Effects and Measurement 152

8.2	Experiments and Probability-Based Online Panels	153
8.2.1	Online Experiments in Probability-Based Panels: Advantages	153
8.2.2	Examples of Experiments in the LISS Panel	154
8.3	Mixed-Mode Question Format Experiments	154
8.3.1	Theoretical Background	154
8.3.2	Method	156
8.3.2.1	Respondents	156
8.3.2.2	Experimental Design	156
8.3.2.3	Questionnaire	157
8.3.3	Analyses and Results	157
8.3.3.1	Response and Weighting	157
8.3.3.2	Reliability	157
8.3.3.3	Response Styles	158
8.4	Summary and Discussion	161
	Acknowledgments	162
	References	162
<b>9</b>	<b>Conflicting Cues: Item Nonresponse and Experimental Mortality</b>	<b>167</b>
	<i>David J. Ciuk and Berwood A. Yost</i>	
9.1	Introduction	167
9.2	Survey Experiments and Item Nonresponse	167
9.3	Case Study: Conflicting Cues and Item Nonresponse	170
9.4	Methods	170
9.5	Issue Selection	171
9.6	Experimental Conditions and Measures	172
9.7	Results	173
9.8	Addressing Item Nonresponse in Survey Experiments	174
9.9	Summary	178
	References	179
<b>10</b>	<b>Application of a List Experiment at the Population Level: The Case of Opposition to Immigration in the Netherlands</b>	<b>181</b>
	<i>Mathew J. Creighton, Philip S. Brenner, Peter Schmidt, and Diana Zavala-Rojas</i>	
10.1	Fielding the Item Count Technique (ICT)	183
10.2	Analyzing the Item Count Technique (ICT)	185
10.3	An Application of ICT: Attitudes Toward Immigrants in the Netherlands	186
10.3.1	Data – The Longitudinal Internet Studies for the Social Sciences Panel	186
10.3.2	Implicit and Explicit Agreement	187
10.3.3	Social Desirability Bias (SDB)	188
10.4	Limitations of ICT	190
	References	192
	<b>Part IV Introduction to Section on Interviewers</b>	<b>195</b>
	<i>Brady T. West and Edith D. de Leeuw</i>	
<b>11</b>	<b>Race- and Ethnicity-of-Interviewer Effects</b>	<b>197</b>
	<i>Allyson L. Holbrook, Timothy P. Johnson, and Maria Krysan</i>	
11.1	Introduction	197
11.1.1	The Problem	197

11.1.2	Theoretical Explanations	198
11.1.2.1	Survey Participation	198
11.1.2.2	Survey Responses	199
11.1.3	Existing Evidence	199
11.1.3.1	Survey Participation	199
11.1.3.2	Survey Responses in Face-to-Face Interviews	200
11.1.3.3	Survey Responses in Telephone Interviews	202
11.1.3.4	Survey Responses in Surveys that Are Not Interviewer-Administered	204
11.1.3.5	Other Race/Ethnicity of Interviewer Research	204
11.2	The Current Research	205
11.3	Respondents and Procedures	207
11.4	Measures	207
11.4.1	Interviewer Race/Ethnicity	207
11.4.2	Sample Frame Variables	208
11.4.2.1	Disposition Codes	208
11.4.2.2	Neighborhood Level Variables	208
11.4.3	Respondent Variables	208
11.4.3.1	Race/Ethnicity	208
11.4.3.2	Perceived Race/Ethnicity	209
11.4.3.3	Confidence	209
11.4.3.4	Male	209
11.4.3.5	Age	209
11.4.3.6	Years of Education	209
11.4.3.7	Household Income	209
11.4.3.8	Political Ideology	209
11.4.3.9	Affirmative Action	209
11.4.3.10	Immigration Policy	210
11.4.3.11	Attitudes Toward Obama	210
11.5	Analysis	210
11.5.1	Sample Frame Analyses	210
11.5.2	Survey Sample Analyses	211
11.6	Results	211
11.6.1	Sample Frame	211
11.6.1.1	Survey Participation or Response	212
11.6.1.2	Refusals	212
11.6.1.3	Contact	213
11.6.2	Perceptions of Interviewer Race	214
11.6.3	Survey Responses	215
11.7	Discussion and Conclusion	219
11.7.1	Limitations	220
11.7.2	Future Work	221
	References	221
<b>12</b>	<b>Investigating Interviewer Effects and Confounds in Survey-Based Experimentation</b>	<b>225</b>
	<i>Paul J. Lavrakas, Jenny Kelly, and Colleen McClain</i>	
12.1	Studying Interviewer Effects Using a <i>Post hoc</i> Experimental Design	226
12.1.1	An Example of a <i>Post hoc</i> Investigation of Interviewer Effects	228
12.2	Studying Interviewer Effects Using <i>A priori</i> Experimental Designs	230
12.2.1	Assignment of Experimental Treatments to Interviewers	230

- 12.3 An Original Experiment on the Effects of Interviewers Administering Only One Treatment vs. Interviewers Administering Multiple Treatments 232
- 12.3.1 The Design of Our Survey-Based Experiment 232
- 12.3.2 Analytic Approach 233
- 12.3.3 Experimental Design Findings 234
- 12.4 Discussion 239
- References 242

**Part V Introduction to Section on Adaptive Design 245**

*Courtney Kennedy and Brady T. West*

- 13 Using Experiments to Assess Interactive Feedback That Improves Response Quality in Web Surveys 247**  
*Tanja Kunz and Marek Fuchs*
- 13.1 Introduction 247
- 13.1.1 Response Quality in Web Surveys 247
- 13.1.2 Dynamic Interventions to Enhance Response Quality in Web Surveys 249
- 13.2 Case Studies – Interactive Feedback in Web Surveys 251
- 13.3 Methodological Issues in Experimental Visual Design Studies 258
- 13.3.1 Between-Subjects Design 258
- 13.3.2 Field Experiments vs. Lab Experiments 259
- 13.3.3 Replication of Experiments 259
- 13.3.4 Randomization in the Net Sample 260
- 13.3.5 Independent Randomization 262
- 13.3.6 Effects of Randomization on the Composition of the Experimental Conditions 262
- 13.3.7 Differential Unit Nonresponse 263
- 13.3.8 Nonresponse Bias and Generalizability 264
- 13.3.9 Breakoff as an Indicator of Motivational Confounding 264
- 13.3.10 Manipulation Check 265
- 13.3.11 A Direction for Future Research: Causal Mechanisms 268
- References 269
- 14 Randomized Experiments for Web-Mail Surveys Conducted Using Address-Based Samples of the General Population 275**  
*Z. Tuba Suzer-Gurtekin, Mahmoud Elkasabi, James M. Lepkowski, Mingnan Liu, and Richard Curtin*
- 14.1 Introduction 275
- 14.1.1 Potential Advantages of Web-Mail Designs 275
- 14.1.2 Design Considerations with Web-Mail Surveys 276
- 14.1.3 Empirical Findings from General Population Web-Mail Studies 277
- 14.2 Study Design and Methods 278
- 14.2.1 Experimental Design 278
- 14.2.2 Questionnaire Design 279
- 14.2.3 Weights 280
- 14.2.4 Statistical Methods 280
- 14.3 Results 281
- 14.4 Discussion 285
- References 287

**Part VI Introduction to Section on Special Surveys 291***Michael W. Traugott and Edith D. de Leeuw*

- 15 Mounting Multiple Experiments on Longitudinal Social Surveys: Design and Implementation Considerations 293**  
*Peter Lynn and Annette Jäckle*
- 15.1 Introduction and Overview 293
- 15.2 Types of Experiments that Can Be Mounted in a Longitudinal Survey 294
- 15.3 Longitudinal Experiments and Experiments in Longitudinal Surveys 295
- 15.4 Longitudinal Surveys that Serve as Platforms for Experimentation 296
- 15.5 The *Understanding Society* Innovation Panel 298
- 15.6 Avoiding Confounding of Experiments 299
- 15.7 Allocation Procedures 301
- 15.7.1 Assignment Within or Between Households and Interviewers? 302
- 15.7.2 Switching Treatments Between Waves 303
- 15.8 Refreshment Samples 304
- 15.9 Discussion 305
- 15.A Appendix: Stata Syntax to Produce Table 15.3 Treatment Allocations 306
- References 306
- 16 Obstacles and Opportunities for Experiments in Establishment Surveys Supporting Official Statistics 309**  
*Diane K. Willimack and Jaki S. McCarthy*
- 16.1 Introduction 309
- 16.2 Some Key Differences Between Household and Establishment Surveys 310
- 16.2.1 Unit of Analysis 310
- 16.2.2 Statistical Products 310
- 16.2.3 Highly Skewed or Small Specialized Target Populations 310
- 16.2.4 Sample Designs and Response Burden 311
- 16.2.5 Question Complexity and Availability of Data in Records 311
- 16.2.6 Labor-intensive Response Process 311
- 16.2.7 Mandatory Reporting 312
- 16.2.8 Availability of Administrative and Auxiliary Data 312
- 16.3 Existing Literature Featuring Establishment Survey Experiments 312
- 16.3.1 Target Populations 312
- 16.3.2 Research Purpose 313
- 16.3.3 Complexity and Context 313
- 16.4 Key Considerations for Experimentation in Establishment Surveys 314
- 16.4.1 Embed the Experiment Within Production 314
- 16.4.2 Conduct Standalone Experiments 315
- 16.4.3 Use Nonsample Cases 315
- 16.4.4 Exclude Large Establishments 315
- 16.4.5 Employ Risk Mitigation Strategies and Alternatives 316
- 16.5 Examples of Experimentation in Establishment Surveys 318
- 16.5.1 Example 1: Adaptive Design in the Agricultural Resource Management Survey 318
- 16.5.1.1 Experimental Strategy: Embed the Experiment Within the Production Survey 318
- 16.5.1.2 Experimental Strategy: Exclude Large Units from Experimental Design 319
- 16.5.2 Example 2: Questionnaire Testing for the Census of Agriculture 319

- 16.5.2.1 Experimental Strategy: Conduct a Standalone Experiment 319
- 16.5.2.2 Experimental Strategy: Exclude Large Establishments 320
- 16.5.3 Example 3: Testing Contact Strategies for the Economic Census 320
- 16.5.3.1 Experimental Strategy: Embed Experiments in Production Surveys 321
- 16.5.4 Example 4: Testing Alternative Question Styles for the Economic Census 321
- 16.5.4.1 Experimental Strategy: Use Nonproduction Cases Alongside a Production Survey 321
- 16.5.5 Example 5: Testing Adaptive Design Alternatives in the Annual Survey of Manufactures 322
- 16.5.5.1 Experimental Strategy: Exclude Large Units 322
- 16.5.5.2 Experimental Strategy: Embed the Experiment Within the Production Survey 322
- 16.6 Discussion and Concluding Remarks 323
- Acknowledgments 324
- References 324

## **Part VII Introduction to Section on Trend Data 327**

*Michael W. Traugott and Paul J. Lavrakas*

- 17 Tracking Question-Wording Experiments Across Time in the General Social Survey, 1984–2014 329**  
*Tom W. Smith and Jaesok Son*
- 17.1 Introduction 329
- 17.2 GSS Question-Wording Experiment on Spending Priorities 330
- 17.3 Experimental Analysis 330
- 17.4 Summary and Conclusion 338
- 17.A National Spending Priority Items 339
- References 340
- 18 Survey Experiments and Changes in Question Wording in Repeated Cross-Sectional Surveys 343**  
*Allyson L. Holbrook, David Sterrett, Andrew W. Crosby, Marina Stavrakantonaki, Xiaoheng Wang, Tianshu Zhao, and Timothy P. Johnson*
- 18.1 Introduction 343
- 18.2 Background 344
- 18.2.1 Repeated Cross-Sectional Surveys 344
- 18.2.2 Reasons to Change Question Wording in Repeated Cross-Sectional Surveys 345
- 18.2.2.1 Methodological Advances 345
- 18.2.2.2 Changes in Language and Meaning 346
- 18.2.3 Current Practices and Expert Insight 346
- 18.3 Two Case Studies 347
- 18.3.1 ANES 347
- 18.3.1.1 Description of Question Wording Experiment 347
- 18.3.1.2 Data Collection 348
- 18.3.1.3 Measures 349
- 18.3.1.4 Analysis 352
- 18.3.1.5 Results 352
- 18.3.1.6 Implications for Trend Analysis 355
- 18.3.2 General Social Survey (GSS) 356

- 18.3.2.1 Purpose of Experiment 356
- 18.3.2.2 Data Collection 356
- 18.3.2.3 Measures 356
- 18.3.2.4 Analysis 357
- 18.3.2.5 Results 360
- 18.4 Implications and Conclusions 362
- 18.4.1 Limitations and Suggestions for Future Research 363
- Acknowledgments 364
- References 364

## **Part VIII Vignette Experiments in Surveys 369**

*Allyson Holbrook and Paul J. Lavrakas*

- 19 Are Factorial Survey Experiments Prone to Survey Mode Effects? 371**  
*Katrin Auspurg, Thomas Hinz, and Sandra Walzenbach*
- 19.1 Introduction 371
- 19.2 Idea and Scope of Factorial Survey Experiments 372
- 19.3 Mode Effects 373
- 19.3.1 Typical Respondent Samples and Modes in Factorial Survey Experiments 373
- 19.3.2 Design Features and Mode Effects 374
- 19.3.2.1 Selection of Dimensions and Levels 374
- 19.3.2.2 Generation of Questionnaire Versions and Allocation to Respondents 375
- 19.3.2.3 Number of Vignettes per Respondent and Response Scales 376
- 19.3.2.4 Interviewer Effects and Social Desirability Bias 377
- 19.3.3 Summing Up: Mode Effects 378
- 19.4 Case Study 378
- 19.4.1 Data and Methods 378
- 19.4.1.1 Survey Details 378
- 19.4.1.2 Analysis Techniques 380
- 19.4.2 Mode Effects Regarding Nonresponse and Measurement Errors 380
- 19.4.2.1 Nonresponse 380
- 19.4.2.2 Response Quality 382
- 19.4.3 Do Data Collection Mode Effects Impact Substantive Results? 385
- 19.4.3.1 Point Estimates and Significance Levels 385
- 19.4.3.2 Measurement Efficiency and Sources of Unexplained Variance 387
- 19.5 Conclusion 388
- References 390
- 20 Validity Aspects of Vignette Experiments: Expected “What-If” Differences Between Reports of Behavioral Intentions and Actual Behavior 393**  
*Stefanie Eifler and Knut Petzold*
- 20.1 Outline of the Problem 393
- 20.1.1 Problem 393
- 20.1.2 Vignettes, the Total Survey Error Framework, and Cognitive Response Process Perspectives 394
- 20.1.3 State of the Art 395
- 20.1.4 Research Strategy for Evaluating the External Validity of Vignette Experiments 397

- 20.2 Research Findings from Our Experimental Work 399
- 20.2.1 Experimental Design for Both Parts of the Study 399
- 20.2.2 Data Collection for Both Parts of the Study 400
- 20.2.2.1 Field Experiment 400
- 20.2.2.2 Vignette Experiment 402
- 20.2.3 Results for Each Part of the Study 405
- 20.2.3.1 Field Experiment 405
- 20.2.3.2 Vignette Experiment 405
- 20.2.4 Systematic Comparison of the Two Experiments 406
- 20.2.4.1 Assumptions and Precautions About the Comparability of Settings, Treatments, and Outcomes 406
- 20.2.4.2 Assumptions and Precautions About the Comparability of Units 407
- 20.2.4.3 Results 408
- 20.3 Discussion 411
- References 413

**Part IX Introduction to Section on Analysis 417**  
*Brady T. West and Courtney Kennedy*

- 21 Identities and Intersectionality: A Case for Purposive Sampling in Survey-Experimental Research 419**  
*Samara Klar and Thomas J. Leeper*
- 21.1 Introduction 419
- 21.2 Common Techniques for Survey Experiments on Identity 420
- 21.2.1 Purposive Samples 422
- 21.2.2 The Question of Representativeness: An Answer in Replication 423
- 21.2.3 Response Biases and Determination of Eligibility 425
- 21.3 How Limited Are Representative Samples for Intersectionality Research? 426
- 21.3.1 Example from TESS 426
- 21.3.2 Example from a Community Association-Based LGBTQ Sample 427
- 21.3.3 Example from Hispanic Women Sample in Tucson, AZ 430
- 21.4 Conclusions and Discussion 430
- Author Biographies 431
- References 431
  
- 22 Designing Probability Samples to Study Treatment Effect Heterogeneity 435**  
*Elizabeth Tipton, David S. Yeager, Ronaldo Iachan, and Barbara Schneider*
- 22.1 Introduction 435
- 22.1.1 Probability Samples Facilitate Estimation of Average Treatment Effects 437
- 22.1.2 Treatment Effect Heterogeneity in Experiments 438
- 22.1.3 Estimation of Subgroup Treatment Effects 440
- 22.1.3.1 Problems with Sample Selection Bias 440
- 22.1.3.2 Problems with Small Sample Sizes 440
- 22.1.4 Moderator Analyses for Understanding Causal Mechanisms 441
- 22.1.4.1 Problems of Confounder Bias in Probability Samples 442
- 22.1.4.2 Additional Confounder Bias Problems with Nonprobability Samples 442
- 22.1.4.3 Problems with Statistical Power 443

22.1.5	Stratification for Studying Heterogeneity	444
22.1.5.1	Overview	444
22.1.5.2	Operationalizing Moderators	445
22.1.5.3	Orthogonalizing Moderators	445
22.1.5.4	Stratum Creation and Allocation	445
22.2	Nesting a Randomized Treatment in a National Probability Sample: The NSLM	446
22.2.1	Design of the NSLM	446
22.2.1.1	Population Frame	446
22.2.1.2	Two-Stage Design	447
22.2.1.3	Nonresponse	447
22.2.1.4	Experimental Design	447
22.2.1.5	Treatment	447
22.2.1.6	Outcomes	447
22.2.2	Developing a Theory of Treatment Effect Heterogeneity	447
22.2.2.1	Goals	447
22.2.2.2	Potential Moderators	447
22.2.2.3	Operationalizing School Achievement Level	448
22.2.2.4	Orthogonalizing for Minority Composition	449
22.2.3	Developing the Final Design	449
22.2.3.1	Stratum Creation	449
22.2.3.2	Stratum Allocation	449
22.2.3.3	Implementing Sample Selection	451
22.2.3.4	Final Sample	451
22.3	Discussion and Conclusions	451
22.3.1	Estimands	452
22.3.2	Contextual Effects Matter	452
22.3.3	Plan for Moderators	452
22.3.4	Designing Strata	452
22.3.5	Power Concerns	453
	Acknowledgments	453
	References	453
<b>23</b>	<b>Design-Based Analysis of Experiments Embedded in Probability Samples</b>	<b>457</b>
	<i>Jan A. van den Brakel</i>	
23.1	Introduction	457
23.2	Design of Embedded Experiments	458
23.3	Design-Based Inference for Embedded Experiments with One Treatment Factor	460
23.3.1	Measurement Error Model and Hypothesis Testing	460
23.3.2	Point and Variance Estimation	462
23.3.3	Wald Test	464
23.3.4	Special Cases	464
23.3.5	Hypotheses About Ratios and Totals	465
23.4	Analysis of Experiments with Clusters of Sampling Units as Experimental Units	466
23.5	Factorial Designs	468
23.5.1	Designing Embedded $K \times L$ Factorial Designs	468

23.5.2	Testing Hypotheses About Main Effects and Interactions in $K \times L$ Embedded Factorial Designs	469
23.5.3	Special Cases	471
23.5.4	Generalizations	471
23.6	A Mixed-Mode Experiment in the Dutch Crime Victimization Survey	472
23.6.1	Introduction	472
23.6.2	Survey Design and Experimental Design	472
23.6.3	Software	474
23.6.4	Results	474
23.7	Discussion	477
	Acknowledgments	478
	References	478
<b>24</b>	<b>Extending the Within-Persons Experimental Design: The Multitrait-Multierror (MTME) Approach</b>	<b>481</b>
	<i>Alexandru Cernat and Daniel L. Oberski</i>	
24.1	Introduction	481
24.2	The Multitrait-Multierror (MTME) Framework	482
24.2.1	A Simple but Defective Design: The Interview–Reinterview from the MTME Perspective	483
24.2.2	Designs Estimating Stochastic Survey Errors	485
24.3	Designing the MTME Experiment	487
24.3.1	What are the main types of measurement errors that should be estimated?	487
24.3.2	How can the questions be manipulated in order to estimate these types of error?	487
24.3.3	Is it possible to manipulate the form pair order?	488
24.3.4	Is there enough power to estimate the model?	488
24.3.5	How can data collection minimize memory effects?	489
24.4	Statistical Estimation for the MTME Approach	489
24.5	Measurement Error in Attitudes Toward Migrants in the UK	491
24.5.1	Estimating Four Stochastic Error Variances Using MTME	491
24.5.1.1	What are the main types of measurement errors that should be estimated?	491
24.5.1.2	How can the questions be manipulated in order to estimate these types of errors?	492
24.5.1.3	Is it possible to manipulate the form pair order?	492
24.5.1.4	Is there enough power to estimate the model?	492
24.5.1.5	How can data collection minimize memory effects?	493
24.5.2	Estimating MTME with four stochastic errors	493
24.6	Results	494
24.7	Conclusions and Future Research Directions	497
	Acknowledgments	498
	References	498
	<b>Index</b>	<b>501</b>

## List of Contributors

### ***Katrin Auspurg***

Ludwig Maximilian University of Munich  
Munich, Germany

### ***Paul Beatty***

Center for Behavioral Science Methods  
U.S. Census Bureau  
Washington, DC, USA

### ***A. Bianchi***

Department of Management, Economics and  
Quantitative Methods  
University of Bergamo  
via dei Caniana 2, 24127 Bergamo, Italy

### ***S. Biffignandi***

Department of Management, Economics and  
Quantitative Methods  
University of Bergamo  
via dei Caniana 2, 24127 Bergamo, Italy

### ***Philip S. Brenner***

Department of Sociology and Center for  
Survey Research  
University of Massachusetts  
100 Morrissey Blvd., Boston, MA 02125, USA

### ***Alexandru Cernat***

Social Statistics Department  
University of Manchester  
Manchester, M13 9PL, UK

### ***David J. Ciuk***

Department of Government  
Franklin & Marshall College  
Lancaster, PA, USA

### ***Carol Cosenza***

Center for Survey Research  
University of Massachusetts  
Boston, MA, USA

### ***Mathew J. Creighton***

School of Sociology, Geary Institute for  
Public Policy  
University College Dublin  
Stillorgan Road, Dublin 4, Ireland

### ***Andrew W. Crosby***

Department of Public Administration  
Pace University  
New York, NY 10038, USA

### ***Richard Curtin***

Institute for Social Research  
University of Michigan  
Ann Arbor, MI 48106, USA

### ***Edith D. de Leeuw***

Department of Methodology & Statistics  
Utrecht University  
Utrecht, the Netherlands

### ***Stefanie Eifler***

Department of Sociology  
Catholic University of Eichstätt-Ingolstadt  
85072 Eichstätt, Germany

### ***Mahmoud Elkasabi***

ICF, The Demographic and Health Surveys  
Program  
Rockville, MD 20850, USA

**Floyd J. Fowler Jr**

Center for Survey Research  
University of Massachusetts  
Boston, MA, USA

**Marek Fuchs**

Darmstadt University of Technology  
Institute of Sociology  
Karolinenplatz 5, 64283 Darmstadt, Germany

**Thomas Hinz**

University of Konstanz  
Konstanz, Germany

**Allyson L. Holbrook**

Departments of Public Administration and  
Psychology and the Survey Research  
Laboratory  
University of Illinois at Chicago  
Chicago, IL 60607, USA

**Joop Hox**

Department of Methodology and Statistics  
University of Utrecht  
Utrecht, the Netherlands

**Ronaldo Iachan**

ICF International  
Fairfax, VA 22031, USA

**Annette Jäckle**

University of Essex  
Institute for Social and Economic Research  
Colchester CO4 3SQ, UK

**Timothy P. Johnson**

Department of Public Administration and  
the Survey Research Laboratory  
University of Illinois at Chicago  
Chicago, IL 60607, USA

**Jenny Kelly**

NORC, University of Chicago  
55 East Monroe Street, Chicago, IL  
60603, USA

**Courtney Kennedy**

Pew Research Center  
Washington, DC, USA

**Florian Keusch**

Department of Sociology, School of Social  
Sciences  
University of Mannheim  
Mannheim, Germany

**Samara Klar**

School of Government & Public Policy  
University of Arizona  
Tucson, AZ 85721, USA

**Maria Krysan**

Department of Sociology and the Institute of  
Government & Public Affairs  
University of Illinois at Chicago  
Chicago, IL, USA

**Tanja Kunz**

Leibniz-Institute for the Social Sciences  
P.O.Box 12 21 55, 68072, Mannheim,  
Germany

**Paul J. Lavrakas**

NORC, University of Chicago  
55 East Monroe Street, Chicago, IL  
60603, USA

**Thomas J. Leeper**

Department of Methodology  
London School of Economics and Political  
Science  
London WC2A 2AE, UK

**James M. Lepkowski**

Institute for Social Research  
University of Michigan  
Ann Arbor, MI 48106, USA

**Mingnan Liu**

Facebook, Inc.  
Menlo Park, CA, USA

**Peter Lynn**

University of Essex  
Institute for Social and Economic Research  
Colchester CO4 3SQ, UK

**Jaki S. McCarthy**

United States Department of Agriculture  
National Agricultural Statistics Service  
(USDA/NASS)  
1400 Independence Avenue, SW,  
Washington, DC 20250-2054, USA

**Colleen McClain**

Program in Survey Methodology at the  
University of Michigan, Institute for Social  
Research  
426 Thompson Street, Ann Arbor, MI  
48104, USA

**Daniel L. Oberski**

Department of Methodology & Statistics  
Utrecht University  
Utrecht, 3584 CH, the Netherlands

**Kristen Olson**

Department of Sociology  
University of Nebraska-Lincoln  
Lincoln, NE, USA

**Colm O'Muircheartaigh**

Harris School of Public Policy  
University of Chicago and NORC at the  
University of Chicago  
Chicago, IL, United States

**Linda K. Owens**

Stephens Family Clinical Research Institute  
Carle Foundation Hospital, 611 W. Park Street  
Urbana, IL 61801, USA

**Jennifer A. Parsons**

Survey Research Laboratory  
University of Illinois at Chicago  
Chicago, IL, USA

**Knut Petzold**

Sociology Section  
Ruhr-Universität Bochum  
44801 Bochum, Germany

**Annette Scherpenzeel**

Chair for Economics of Aging, School of  
Management  
Technical University of Munich  
Munich, Germany

**Peter Schmidt**

Department of Political Science and Centre  
for Environment and Development (ZEU)  
University of Giessen  
Karl-Glöcknerstrasse 21 E, 35394 Giessen  
Germany

**Barbara Schneider**

College of Education and Sociology  
Department  
Michigan State University  
East Lansing, MI 48824, USA

**Stephen Smith**

NORC at the University of Chicago  
Chicago, IL, United States

**Tom W. Smith**

Center for the Study of Politics and Society  
NORC at the University of Chicago  
1155 East 60th Street Chicago, IL 60637, USA

**Jolene D. Smyth**

Department of Sociology  
University of Nebraska-Lincoln  
Lincoln, NE, USA

**Jaesok Son**

Center for the Study of Politics and Society  
NORC at the University of Chicago  
1155 East 60th Street Chicago, IL 60637, USA

**Mathew Stange**

Mathematica Policy Research  
Ann Arbor, MI, USA

**Marina Stavrakantonaki**

Department of Public Administration  
University of Illinois at Chicago  
Chicago, IL 60607, USA

**David Sterrett**

NORC at the University of Chicago  
Chicago, IL 60603, USA

**Z. Tuba Suzer-Gurtekin**

Institute for Social Research  
University of Michigan  
Ann Arbor, MI 48106, USA

**Elizabeth Tipton**

Human Development Department, Teachers  
College  
Columbia University  
New York, NY 10027, USA

and

Statistics Department  
Northwestern University  
Evanston, IL 60201, USA

**Michael W. Traugott**

Center for Political Studies  
Institute for Social Research, University of  
Michigan  
Ann Arbor, MI, USA

**Jan A. van den Brakel**

Department of Statistical Methods  
Statistics Netherlands  
Heerlen, the Netherlands

and

Department of Quantitative Economics  
Maastricht University School of Business and  
Economics  
Maastricht, the Netherlands

**Susanne Vogl**

Department of Education  
Department of Sociology, University of  
Vienna  
Vienna, Austria

**Sandra Walzenbach**

University of Konstanz,  
Konstanz, Germany  
and

ISER/University of Essex  
Colchester, UK

**Xiaoheng Wang**

Department of Public Administration  
University of Illinois at Chicago  
Chicago, IL 60607, USA

**Brady T. West**

Survey Research Center  
Institute for Social Research, University of  
Michigan  
Ann Arbor, MI, USA

**Diane K. Willimack**

United States Department of Commerce  
U.S. Census Bureau  
4600 Silver Hill Road, Washington, DC  
20233-0001, USA

**Jaclyn S. Wong**

Department of Sociology  
University of South Carolina  
Columbia, SC, United States

**Ting Yan**

Westat  
Rockville, MD, USA

**David S. Yeager**

Psychology Department  
University of Texas at Austin  
Austin, TX 78712, USA

**Berwood A. Yost**

Franklin & Marshall College  
Floyd Institute for Public Policy and Center  
for Opinion Research  
Lancaster, PA, USA

***Diana Zavala-Rojas***

Research and Expertise Centre for Survey  
Methodology and European Social Survey  
ERIC, Department of Political and Social  
Sciences

Universitat Pompeu Fabra

C/de Ramon Trias Fargas, 25-27, 08005

Barcelona, Spain

***Tianshu Zhao***

Department of Public Administration

University of Illinois at Chicago

Chicago, IL 60607, USA



## Preface by Dr. Judith Tanur

I am enormously flattered to be asked to supply a preface for this path-breaking volume of essays about experiments embedded in surveys. I had hoped to contribute a chapter to the volume with my friend and long-term collaborator, Stephen Fienberg, but Steve, active to the end, lost his long battle with cancer before he was able to make the time to work on the chapter we had planned to write. So, I would like to use this opportunity to write something about the work we had done, and had hoped to continue, on the parallels between experimental and survey methodology, on experiments embedded in surveys (and vice-versa), and some of the considerations for analysis occasioned by such embedding.

We had long noted (e.g. Fienberg and Tanur 1987, 1988, 1989, 1996) that there are a great many parallels between elements of survey design and experimental design. Although in fact surveys and experiments had developed very long and independent traditions by the start of the twentieth century, it was only with the rise of ideas associated with mathematical statistics in the 1920s that the tools for major progress in these areas became available. The key intellectual idea was the role of randomization or random selection, both in experimentation and in sampling, and both R.A. Fisher and Jerzy Neyman utilized that idea, although in different ways. The richness of the two separate literatures continues to offer new opportunities for cross-fertilization of theory and tools for survey practice in particular.

Although the ideas are parallel, often the purpose these ideas serve is different in the two domains. For example, experimental randomization is used in experimental design in order to justify the assumption that the experimental and control groups are equivalent a priori, but it finds its analog in probability sampling in surveys in order to assure the “representativeness” of the sample *vis a vis* the population to which generalizations are to be made. On the other hand, similar processes to create homogeneous groups occur in both experimental and sampling designs and serve similar purposes. In experimentation, blocking creates homogeneous groups, exerting control of experimental error by segregating the effects of extraneous sources of variation. In sampling, stratification serves the same purpose by drawing sample members from each of the homogeneous groups into which the population can be divided. The list could be lengthened (as indeed, Steve and I did in the papers cited above) to include such parallels in design as those between Latin and Graeco-Latin squares on the one hand and lattice sampling or “deep stratification” on the other and between split plot designs and cluster sampling. In the analysis stage, we pointed to the parallel between covariance adjustment in experiments and poststratification in surveys. We embarked on a project to find and describe more modern parallels and to encourage researchers in each of these fields to look in the literature of the other for ideas about design and analysis.

We wrote a series of papers and envisioned a book, *Reaching Conclusions: The Role of Randomized Experiments and Sample Surveys*, that will now, unfortunately, remain a draft, that explored the ramifications of these parallels, pointed out newer and less obvious parallels

between experiments and surveys, and noted the frequent embedding of experiments in surveys and vice versa. In particular, we urged – sometimes by example – that when such embedding took place, the analyst should take advantage of the embedded structure in planning and carrying out the analysis.

In our 1989 *Science* paper, we addressed these issues of analysis of embedded experiments, reiterating the point that although there are formal parallels between the structures of surveys and experiments there are fundamental inferential differences as described above. We pointed out three inferential stances that could be used. In the context of this volume, it is perhaps worth quoting that discussion at some length (p. 243).

(1) One can use the *standard experiment paradigm*, which relies largely on internal validity based on randomization and local control (for example, the device of blocking) and on the assumption that the unique effects of experimental units and the treatment effects can be expressed in a simple additive form, without interaction (Fisher 1935, Ch. 4; Kempthorne 1952, Ch. 9). Then inference focuses on within-experiment treatment differences.

(2) One can use the *standard sampling paradigm*, which, for a two-treatment experiment embedded in a survey relies largely on external validity and generalizes the observations for each of the treatments to separate but paired populations of values. Each unit or individual in the original population from which the sample was drawn is conceived to have a pair of values, one for each treatment. But only one of these is observable, depending on which treatment is given. Then inferences focus on the mean difference or the difference in the means of the two populations.

(3) One can conceptualize a *population of experiments*, of which the present embedded experiment is a unit or sample of units, and thus capitalize on the internal validity created by the design of the present embedded experiment as well as the external validity created by the generalization from the present experiment to the conceptual population of experiments. Then inferences focus on treatment differences in a broader context than simply the present embedded experiment.

Then we took advantage of the generosity of Roger Tourangeau and Kenneth Rasinski to reanalyze data they had collected in a probability sample survey on context effects in attitude surveys (1989). They had crossed four issues at differing levels of familiarity with four orders of presentation (balanced in the form of a Latin Square), two versions of the context questions, and two methods of structuring the context question. Thus, there were 16 versions of the questionnaire plus two extra versions with neutral context questions. These 18 versions of the questionnaire were crossed with four interviewers. We considered the interviewers as blocks, and so within each block, we had five replications of an 18-treatment experiment, where 16 of the treatments represent a  $4 \times 2 \times 2$  factorial design. Focusing separately on two of the issues (abortion and welfare), we had four treatment combinations (positive vs. negative contexts by scattered vs. massed structure of the context-setting questions). The context-setting questions for the abortion question dealt with women's rights or traditional values, while those for the welfare question concerned fraud and waste in government programs or government responsibility to provide services. Using logit models to predict a favorable response on the target question and contrasting the four basic treatment combinations with the neutral context, we carried out detailed analyses and found very complicated results. I shall try to sketch only a fraction of results of those analyses – I urge the reader with a serious interest in these issues to refer to the original paper (Fienberg and Tanur, 1987). In short, we found that context matters – but its effect depends of how often the respondent agrees with the context-setting questions. And we found that all three of the inferential stances detailed above gave similar results for the abortion question, while the first and third gave similar results for the welfare question.

This volume concentrates on the embedding of experiments within surveys, often to answer questions about survey methodology as did the Tourangeau/Rasinski experiment discussed above. I would guess that such procedures are the most common uses of embedding. But perhaps it is worth bearing in mind that there are many good examples of the reverse – the embedding of surveys in experiments and such embedded surveys often serve a more substantive purpose. A set of particularly good examples were the massive social experiments of the mid-twentieth century such as the Negative Income Tax experiments and the Health Insurance Study. Many of the measurements of the outcome variables in those randomized experiments were necessarily carried out via surveys of the participants.

I find the contents of this volume fascinating, both for the broad sweep of the topics examined and for the variety of disciplines represented by the contributing authors and what I know of their expertise. I look forward to reading many of the articles.

I wish Steve's work could have been included in a more specific and current way than I have been able to achieve in this preface – he had plans for updating that I am not able to carry out. In that context, I am especially pleased to note the inclusion of a chapter by Jan Van den Brakel, whose work I know Steve admired and whose chapter I expect will contain much of the new material that would have informed our chapter had Steve lived to complete it.

Prolific as Steve was, he left so much good work unfinished and so much more not even yet contemplated. I miss him.

Montauk, New York  
December 2017

Judith Tanur

## References

- Fienberg, S.E. and Tanur, J.M. (1987). Experimental and sampling structures: parallels diverging and meeting. *International Statistical Review* 55 (1): 75–96.
- Fienberg, S.E. and Tanur, J.M. (1988). From the inside out and the outside in: combining experimental and sampling structures. *Canadian Journal of Statistics* 16: 135–151.
- Fienberg, S.E. and Tanur, J.M. (1989). Experimental combining cognitive and statistical approaches to survey design. *Science* 243: 1017–1022.
- Fienberg, S.E. and Tanur, J.M. (1996). Reconsidering the fundamental contributions of Fisher and Neyman on experimentation and sampling. *International Statistical Review* 64: 237–253.
- Fisher, R.A. (1935). *The Design of Experiments*. Edinburgh: Oliver and Boyd.
- Kempthorne, O. (1952). *The Design and Analysis of Experiments*. New York: Wiley.
- Tourangeau, R. and Rasinski, K.A. (1989). Carryover effects in attitude surveys. *Public Opinion Quarterly* 53: 495–524.

