PALGRAVE HANDBOOKS

THE PALGRAVE HANDBOOK OF SURVEY RESEARCH

Edited by
David L. Vannette & Jon A. Krosnick

The Palgrave Handbook of Survey Research

David L. Vannette • Jon A. Krosnick
Editors

The Palgrave Handbook of Survey Research

palgrave
macmillan

Editors
David L. Vannette
Stanford University
Stanford, California, USA

Jon A. Krosnick
Stanford University
Stanford, California, USA

Qualtrics, LLC
Provo, Utah, USA

ISBN 978-3-319-54394-9
ISBN 978-3-319-54395-6 (eBook)
https://doi.org/10.1007/978-3-319-54395-6

Library of Congress Control Number: 2017953116

© The Editor(s) (if applicable) and The Author(s) 2018
Open Access Chapter 23 is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). For further details see license information in the chapter.
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover illustration: Anna Berkut / Alamy Stock Photo

Printed on acid-free paper

This Palgrave Macmillan imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Section 1 Conventional Survey Research

1 Assessing the Accuracy of Survey Research 3
Jon A. Krosnick

2 The Importance of Probability-Based Sampling Methods for Drawing Valid Inferences 7
Gary Langer

3 Sampling for Single and Multi-Mode Surveys Using Address-Based Sampling 13
Colm O'Muircheartaigh

4 Evidence About the Accuracy of Surveys in the Face of Declining Response Rates 19
Scott Keeter

5 Sampling to Minimize Nonresponse Bias 23
J. Michael Brick

6 Cross-National Issues in Response Rates 29
Vasja Vehovar and Koen Beullens

7 Choosing a Mode of Survey Data Collection 43
Roger Tourangeau

8 **Mixed-Mode and Mixed-Device Surveys** 51
Edith Desiree de Leeuw and Vera Toepoel

9 **The Use and Effects of Incentives in Surveys** 63
Eleanor Singer

10 **Methods for Determining Who Lives in a Household** 71
Kathleen Targowski Ashenfelter

11 **Harmonization for Cross-National Comparative Social Survey Research: A Case Study Using the "Private Household" Variable** 79
Jürgen H. P. Hoffmeyer-Zlotnik and Uwe Warner

12 **Answering for Someone Else: Proxy Reports in Survey Research** 87
Curtiss Cobb

13 **Improving Question Design to Maximize Reliability and Validity** 95
Jon A. Krosnick

14 **Cognitive Interviewing in Survey Design: State of the Science and Future Directions** 103
Gordon Willis

15 **Survey Interviewing: Departures from the Script** 109
Nora Cate Schaeffer

16 **How to Improve Coding for Open-Ended Survey Data: Lessons from the ANES** 121
Arthur Lupia

17 **Applying Human Language Technology in Survey Research** 129
Mark Liberman

18 **Maintaining Respondent Trust and Protecting Their Data** 135
Roger Tourangeau

19 **Tackling Panel Attrition** 143
Peter Lynn

20 Respondent Attrition Versus Data Attrition and Their Reduction 155
Randall J. Olsen

21 Best Practices for Creating Survey Weights 159
Matthew DeBell

Section 2 Opportunities to Expand Data Collection

22 New Kinds of Survey Measurements 165
Carli Lessof and Patrick Sturgis

23 The Role of Surveys in the Era of "Big Data" 175
Mario Callegaro and Yongwei Yang

24 Getting the Most Out of Paradata 193
Frauke Kreuter

25 Open Probability-Based Panel Infrastructures 199
Marcel Das, Arie Kapteyn and Michael Bosnjak

26 Collecting Interviewer Observations to Augment Survey Data 211
Brady T. West

27 "Just One More Thing": Using Leave-Behind Measurement Supplements to Augment Survey Data Collection 217
Michael W. Link

28 Ecological Momentary Assessment in Survey Research 221
Arthur A. Stone

29 Biomarkers in Representative Population Surveys 227
David Weir

30 Measuring Past Events Using Calendar Aids 235
Robert F. Belli

31 Collecting Social Network Data 241
Tobias H. Stark

Section 3 Linking Survey Data with External Sources

32 Methods of Linking Survey Data to Official Records 257
Joseph W. Sakshaug

33 Linking Individual-Level Survey Data to Consumer File Records 263
Josh Pasek

34 Linking Survey Data to Administrative Records in a Comparative Survey Context 267
Annelies G. Blom and Julie Korbmacher

35 Linking Survey Data with the Catalist Commercial Database 275
Robert Blaemire

36 Challenges and Opportunities in Collecting Election Administration Data 289
Michael P. McDonald

37 Challenges with Validating Survey Data 293
Matthew K. Berent

38 The Promise of Collaborative Data Sharing Across Research Sectors 297
Robert M. Groves

Section 4 Improving Research Transparency and Data Dissemination

39 The Importance of Data Curation 303
Steven Ruggles

40 Improving the Usability of Survey Project Websites 309
David L. Vannette

41 The Role of Transparency in Maintaining the Legitimacy and Credibility of Survey Research 315
Arthur Lupia

42 Evidence-Based Survey Operations: Choosing and Mixing Modes 319
Michael Bosnjak

43 Best Practices for Survey Research 331
David L. Vannette

Section 5 Detailed Chapters

44 Reasons for Optimism About Survey Research 347
Jon A. Krosnick

45 Probability Versus Non-Probability Methods 351
Gary Langer

46 Address-Based and List-Based Sampling 363
Colm O'Muircheartaigh

47 The Impact of Survey Non-response on Survey Accuracy 373
Scott Keeter

48 Optimizing Response Rates 383
J. Michael Brick

49 Data Collection Mode 393
Roger Tourangeau

50 Survey Incentives 405
Eleanor Singer

51 Building Household Rosters Sensibly 417
Kathy Ashenfelter

52 Proxy Reporting 427
Curtiss Cobb

53 Questionnaire Design 439
Jon A. Krosnick

54 Cognitive Evaluation of Survey Instruments 457
Gordon Willis

55 Interviewer Deviations from Scripts 465
Nora Cate Schaeffer

56 Coding Open Responses 473
Arthur Lupia

57 What HLT Can Do for You (and Vice Versa) 489
Mark Liberman

58 Confidentiality, Privacy, and Anonymity 501
Roger Tourangeau

59 Panel Attrition 509
Randall Olsen

60 Computation of Survey Weights 519
Matthew DeBell

61 Paradata 529
Frauke Kreuter

62 Interviewer Observations 537
Brady T. West

63 Leave-Behind Measurement Supplements 549
Michael Link

64 Ecological Momentary Assessment and Experience Sampling 563
Arthur Stone

65 Biomarkers 573
David Weir

66 Specialized Tools for Measuring Past Events 587
Robert F. Belli

67 Linking Survey Data to Official Government Records 597
Joseph W. Sakshaug

68 Linking Knowledge Networks Web Panel Data with External Data 607
Josh Pasek

69 History and Promise and Blending Survey Data with Government Records on Turnout 621
Michael P. McDonald

70 Improving Information Quality and Availability Through Interactions Between Government and Academic, and Industry Survey Research Sectors 627
Robert M. Groves

71 Metadata and Preservation 635
Steven Ruggles

72 Usability of Survey Project Websites 645
David L. Vannette

73 Research Transparency and the Credibility of Survey-Based Social Science 655
Arthur Lupia

Index 667

37

Challenges with Validating Survey Data

Matthew K. Berent

Self-reports of voting behavior in many studies, including the American National Election Studies (ANES), are often about 20 percent higher than the official statistics on turnout. This has been a remarkably consistent finding since the 1960s, meaning that as actual turnout increases or decreases there are similar increases or decreases in self-reports of voting. Many people take this information to imply that the estimates generated by self-report survey data are not representative of the population because they don't match the "true" gold standard value provided by the government.

However, there are a number of different ways that this discrepancy could arise that warrants consideration. For example, a respondent might not answer the question because perhaps they believe voting behavior is private and sensitive and are unwilling to report their answer to the survey question, which leads to nonresponse and inaccuracy in the measure that is unrelated to the representativeness of the survey sample.

The second explanation is that there could be survey effects, meaning that participating in the survey could influence the voting behavior that the survey is attempting to measure. Many voting behavior studies involve very in-depth interviews before the election, and then an interview after the election. It is reasonable to expect that, after having answered a long battery of questions about politics before an election, respondents are made aware of

M.K. Berent (✉)
Matt Berent Consulting, Ohio, USA
e-mail: matt@mattberent.com

© The Author(s) 2018

D.L. Vannette, J.A. Krosnick (eds.), *The Palgrave Handbook of Survey Research*, https://doi.org/10.1007/978-3-319-54395-6_37

their political attitudes that were perhaps not as salient before the interview. This increased salience could then influence some respondents to go vote when they might not have otherwise. The result is that these surveys could inadvertently be biasing their own sample to over-represent voters.

This raises a third potential mechanism for higher levels of voting behavior in surveys than in the general population, it is possible that survey response propensity is positively associated with the propensity to vote, meaning that the same types of people who are willing to respond to a long interview about politics may be the same types of people who go out and vote; this would be another way that the sample of survey respondents could be biased from the general population, not demographically but behaviorally.

Lastly, it is entirely possible that respondents misreport their voting behavior, meaning that their answers are factually incorrect. This could be caused by respondents misinterpreting the question, misremembering the behavior, reporting based on their typical voting behavior rather than their specific behavior in the focal election, or intentional lying which might be due to social desirability bias if the respondent believes that it would reflect poorly on them to report not voting in the election to the interviewer.

Misreporting is the most commonly cited cause of the discrepancy between survey data and the official records, leading to the typical conclusion that official records should be used instead of self-reports. However, the approach of using official records is not completely straightforward. To employ this method, researchers must obtain the official records of turnout histories, match each respondent to his or her official record, and then determine the "correct" turnout status for each respondent.

There are two primary problems with the validation task. The first is that over 200 million people are currently eligible to vote in U.S. elections, thus the matching task is not trivial simply on the basis of the size of the databases involved. Second, the federal government does not aggregate individual voting records, meaning that researchers must collect the records from each individual state and there is substantial variability between states in the availability, accuracy, and content of these records.

Researchers are increasingly turning to commercial vendors of political and voter file databases to conduct the matching between survey data and the official voter records. However, these researchers face another set of challenges when working with these databases. For example, vendors vary in their level of transparency, which means that researchers need to identify and specify the level of uncertainty that they are willing to accept with regard to the provenance and quality of their data. Vendors are also typically unwilling

to provide complete details about their matching algorithms, which makes it impossible to estimate the reliability or validity of the matches.

Some researchers opt to attempt in-house matching by obtaining the government registration and turnout records from a sample of states. Publicly available computer applications, such as LinkPlus from the Centers for Disease Control, can then be used to match survey respondents to their official government records. This approach requires more work on the part of the researcher but the benefits are complete control over the data cleaning and matching processes.

Recent evidence from an ANES in-house matching project indicates that two main factors are contributing to the discrepancy between self-reports and official records of voting behavior (Berent et al. 2016). The first is a downward bias in government records that occurs when the records incorrectly identify some respondents as having not turned out when in fact they did. These are all cases where a record match cannot be found, not cases where the self-report and government record data disagree explicitly. The second factor is an upward bias in the self-reports that occurs because people who participate in surveys are more likely to vote. These biases are additive, not offsetting, and seem to account for nearly all of the discrepancies between the self-report data and government records.

Areas for future research:

- Developing a better set of tools for researchers to conduct their own transparent validation studies without needing to use commercial vendors
- Identifying and understanding the correlates of survey participation and turnout to better understand the potential mechanism that drives the higher levels of turnout among survey respondents

Reference and Further Reading

Berent, M. K., Krosnick, J. A., & Lupia, A. (2016). Measuring Voter Registration and Turnout in Surveys Do Official Government Records Yield More Accurate Assessments? *Public Opinion Quarterly*, 80 (3), 597–621.

Matthew K. Berent received a PhD in Social Psychology from The Ohio State University in 1995. Early in his career, Matthew held faculty positions at Colgate University, Idaho State University, University of California – Santa Cruz, and

Florida Atlantic University. He has published research on survey question design, attitude theory, personality theory, and even audiology. More than 1,000 scholarly publications have cited his work.

Matthew is currently the president of Matt Berent consulting. His consulting clients include academics, high-tech companies, start-ups companies, manufacturing companies, and legal firms. He recently worked with researchers from Stanford University and the University of Michigan to develop and evaluate best practices for survey question design and coding open-ended data, to investigate problems with field interviewer behaviors, and to identify factors that cause distortions in time series data. You can find some of his work on these topics at his consulting website mattberent.com.