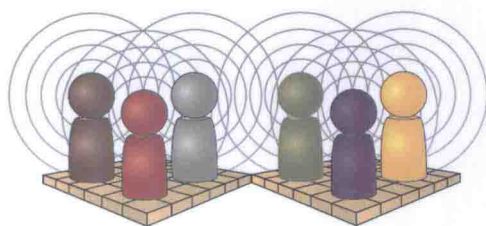
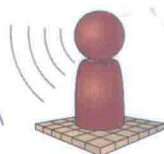
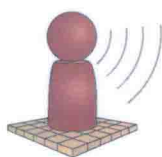
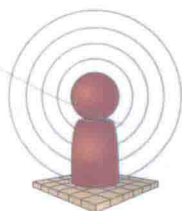


Social Media, Sociality, and Survey Research

EDITED BY

Craig A. Hill • Elizabeth Dean • Joe Murphy



WILEY

Social Media, Sociality, and Survey Research

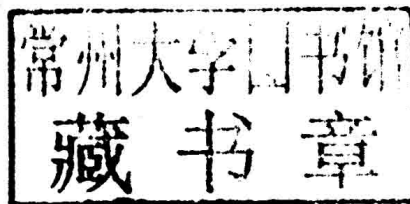
Edited by

Craig A. Hill

Elizabeth Dean

Joe Murphy

*RTI International
Research Triangle Park
North Carolina*



WILEY

Copyright © 2014 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) 750-4470, 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 United States at (800) 762-2974, outside the United States 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 formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data

Social media, sociality, and survey research / [edited by] Craig A. Hill (RTI International, Research Triangle Park, NC), Elizabeth Dean (RTI International, Research Triangle Park, NC), Joe Murphy (RTI International, Research Triangle Park, NC).

pages cm

Includes bibliographical references and index.

ISBN 978-1-118-37973-8 (cloth)

1. Online social networks—Research. 2. Social media—Research. 3. Social surveys—Methodology. I. Hill, Craig A., 1956- II. Dean, Elizabeth, 1974- III. Murphy, Joe, 1974- HM742.S62845 2013 302.23'1—dc23

2013018926

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1



Social Media, Sociality, and Survey Research

List of Figures

1.1	Media Types	5
1.2	Percentage of Online U.S. Population Engaging in Internet Activities	20
1.3	The Sociality Hierarchy	26
2.1	Binary Scale Example	43
2.2	Degree Scale Example	43
2.3	Phases of Sentiment Analysis Process	45
2.4	Spikes on Twitter about Health-Care Reform	49
3.1	ACA Timeline	63
3.2	Number of Health-Care Reform Tweets and Number of Twitter Users Tweeting about Health-Care Reform (March 2009–March 2011)	72
3.3	Adjusted Health-Care Reform Tweets and Health-Care Reform Authors as a Proportion of All Tweets and Twitter Authors (March 2009–March 2011)	73
3.4	Positive, Negative, or Neutral Sentiment of Tweets about Health-Care Reform by Month (March 2009–March 2011)	74
3.5	Perception that Family Would Be Better Off, Worse Off, or No Different as a Result of Health-Care Reform (KFF Survey, March 2009–March 2011)	75
3.6	Perception that Country Would Be Better Off, Worse Off, or No Different as a Result of Health-Care Reform (KFF Survey, April 2009–March 2011)	76
3.7	Comparison of Sentiment/Support for Health-Care Reform–KFF Survey (“Family Better Off”) and Twitter (March 2009–March 2011)	77
3.8	Comparison of Sentiment/Support for Health-Care Reform–KFF Survey (“Country Better Off”) and Twitter (March 2009–March 2011)	78

3.9	Sentiment of Health-Care Reform Tweets that Mention the Republican Party/GOP/Conservatives (March 2009–March 2011)	80
4.1	Gender Distribution Among Facebook Users in the United States Compared to 2010 Census	90
4.2	Age Distribution Among Facebook Users in the United States Compared to 2010 Census	90
4.3	Facebook’s API Structure Connects Datapoints, Such as People to People, Objects, and Events	92
4.4	A Personal Facebook Network Created Using Gephi SNA Software (Bastian et al., 2009)	96
4.5	Front Page of Reconnector Application	100
4.6	Example Question from Reconnector Application	101
4.7	Results Page of Reconnector Application	102
5.1	Tourangeau’s Four-Stage Cognitive Model	110
6.1	Nonsensitive Question Prompt Example	138
6.2	Innocuous Question Prompt Example	138
6.3	In-Person Interview Configuration	141
7.1	Our Smartphone Habits	155
7.2	Top U.S. Smartphone Operating Systems by Market Share	160
7.3	SurveyPulse Download Page in Google Play	163
7.4	SurveyPulse Download Page in iTunes	164
7.5	SurveyPulse Home Screen	167
7.6	Reward Tracking	168
7.7	Example of Result Sharing	169
7.8	SurveyPulse Users by Operating System	171
7.9	SurveyPulse Users by Age	171
7.10	SurveyPulse Users by Race	172
7.11	SurveyPulse Users by Gender	172
7.12	SurveyPulse Users by Highest Level of Education Completed	173
7.13	SurveyPulse Users by Income	173
7.14	SurveyPulse Contact Information Provided by Users	174
7.15	Notification Preferences of SurveyPulse Users	174
7.16	Notification Method Preferences of SurveyPulse Users	175
8.1	Screenshot of RTI Research Challenge Rules Page	184
8.2	Word Cloud of 2012 Research Challenge Entries	184

8.3	Screenshot of Amazon Mechanical Turk Hit Template for Snus Data Collection	192
8.4	Map of Tobacco Retailers that Sell Snus in Various Chicago Zip Codes	193
8.5	Screenshot of Waze Traffic Updates	197
8.6	Screenshot of U.S. Census Return Rate Challenge on Kaggle	198
9.1	Introductory Survey Screenshot	209
9.2	Participation Process	210
9.3	Mean Response Time (in Hours) by Time of Question Tweet	214
9.4	Mean Response Time (in Hours) by Question	215
9.5	Example Tweets Diary Entries	225
9.6	Example DMs	226
10.1	Second Life Classifieds Search Results	238
10.2	Recruitment Coordinator's Twitter Profile	239
10.3	Study Notification on Facebook	240
10.4	Recruitment Coordinator Giving Presentation at HealthInfo Island	241
10.5	Notecard Giver at the Chilbo Community in Second Life	242
10.6	Notecard Giver/Sign at Virtual RTI Interviewing Facility	243
11.1	Interest Over Time in Gamification	254
11.2	Example of Typical Survey Questionnaire	255
11.3	Average Index: Fun vs. Answer Time vs. Straightlining	256
11.4	Experiment: Critical Feedback to Advertisement	264
11.5	Application of Gaming Rules to Survey Questions	266
11.6	Time Limit Encourages Increased Responses	270
11.7	Predicting the Future of a Brand Name	272
11.8	Respondent Average Consideration Time (in Seconds)	272
11.9	Example of Complex Media Planning Task Exercise	274
11.10	Images Versus No Images in a Survey Question	276
11.11	Image of Faucet Added to Survey Question	277
11.12	(a) Examples of Visually Creative Layouts (b) Examples of Visually Creative Layouts	278
11.13	(a) Results of Addition of Icons to Sliders (b) Enjoyment Rating (c) Consideration Time (d) Standard Deviation	279
11.14	Cross-Country Consistency in Responses	281
11.15	Star Stamp Effect Used with Survey Questions	282

11.16	Gamified Versus Standard Responses to Questions	283
11.17	Vacation Word Clouds: Personal Versus Editors' Perspectives	285
11.18	Word Clouds of Favorite Stores	286
11.19	Media Planning Task from Two Perspectives	287
11.20	Standard Versus Game Word Association Responses	288
11.21	Average Improvement in Enjoyment Ratings per Question (Net Promoter Scores) (Standard Versus Games)	290
11.22	Average Improvement in Consideration Time per Question (Seconds) (Basic Versus Games)	290
11.23	Average Standard Deviation in Responses per Respondent (Basic Versus Game)	291
11.24	Level of Fun Versus Respondents' Level of Effort	292
12.1	Sampling Frame Coverage Errors	299
12.2	Sampling Frame Coverage Errors and Sample Estimates	301
12.3	Current and Future Uses of Social Media by Sociality Hierarchy Level	309

List of Tables

2.1	Example Components of Search Criteria	47
3.1	Top 10 Categories of Topics Discussed in Health-Care Reform Tweets	79
5.1	Hypothetical Examples of the Three Cognitive Interviewing Techniques	111
5.2	Participant Demographics	119
5.3	Audio and Video Quality of Second Life and Skype Interviews	121
5.4	Quality Indicators of Cognitive Interview Video Recordings	123
5.5	Participant Disengagement, by Mode	125
5.6	Nonverbal Cues, by Mode	125
5.7	Total Number of Problems, by Mode	126
5.8	Number and Type of Problem, by Mode (Mean, Standard Deviation)	126
6.1	Respondent Demographics by Interview Mode	140
6.2	Belief that Coin Toss Outcome Was Private	143
8.1	Employment Background of Participants	185
8.2	Location of Mechanical Turk Workers	188
8.3	Gender Breakdown of U.S. Mechanical Turk Workers	188
9.1	Study Design	207
9.2	Average Number of Question Tweets Replied to and Response Times	212
9.3	Average Response Time in Hours, by Age	215
9.4	Response Rates by Question Format	217
9.5	Participants Who Completed Debriefing Survey by Diary	222
9.6	Response Mode by Diary Topic	225
10.1	Recruitment Type and Setting	236
10.2	Recruitment Methods	245
10.3	Recruitment by Chronic Condition	246

Contributors

- Sarah Cook**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
- Elizabeth Dean**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
- Jill Dever**, RTI International, 701 13th Street, N.W., Suite 750, Washington, DC 20005-3967
- Carol Haney**, 75 West Street, Dummerston, VT 05301
- Heather Hansen**, RTI International, 701 13th Street, N.W., Suite 750, Washington, DC 20005-3967
- Saira Haque**, RTI International, 230 W Monroe Street, Suite 2100, Chicago, IL 60606
- Brian Head**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
- Craig Hill**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
- Michael Keating**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
- Annice Kim**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
- Joe Murphy**, RTI International, 230 W Monroe Street, Suite 2100, Chicago, IL 60606
- Rebecca Powell**, Survey Research and Methodology, 201 North 13th Street, Lincoln, NE 68588-0241
- Jon Puleston**, GMI, 1 Bedford Avenue, London WC1B 3AU, United Kingdom
- Bryan Rhodes**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
- Ashley Richards**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709
- David Roe**, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709

Adam Sage, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709

Jodi Swicegood, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709

Yuying Zhang, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709

Preface

We started with the premise that survey researchers should be thinking always about the future. That has never been more true than now. Beset by problems not of our own making, the survey research discipline faces unprecedented challenges as a result of declining data quality stemming from, for example, falling response rates, inadequate sampling frames, and approaches and tools that have not adapted to the rapid pace of technological change, especially the changes in the way human beings (our respondents) communicate with each other. Conducting a survey is, at its core, a social interaction between a researcher (represented by an interviewer, or, increasingly, a computer screen) and a (potential) respondent. Yet, the current pace of technological change—and the way people communicate with each other—threatens the upheaval of survey research as we know it because people expect modern communication to take place differently than it did when we developed the current set of best practices for survey research.

Thus, survey researchers should—and must—search for ways to improve the manner in which research is conducted. Survey researchers should—and must—constantly scan the landscape of technological and social change to look for new methods and tools to employ. In this spirit, we have been somewhat jealously watching the explosion of “social media.” Social media is, no doubt, having a profound impact on communication styles and expectations, and thus, more than likely, will have an equally large impact on the way we conduct social science.

We have organized this book around the idea of the “sociality hierarchy”—that is, there are three “levels” of sociality inherent in the current (and future) state of person-to-person interactions using computing devices: (1) Broadcast, (2) Conversational, and (3) Community. Survey researchers should recognize these levels when attempting to apply new social media tools to survey research. We show examples of how this can be done and, perhaps more importantly, how survey researchers should think about applying these in the future as a complement to “traditional” survey research.

Chapter 1 discusses the advent of social media in its many and varied forms and defines it from the perspective of a survey researcher. We also show why survey researchers should be interested in, and vigilant about, social media—and the data it produces. We introduce the concept of the sociality hierarchy for social media and show examples of each level or category.

Chapters 2 and 3 examine broadcast-level social media—the first level in the sociality hierarchy. In Chapter 2, Haney provides a handbook for sentiment analysis, paying close attention to the pitfalls that await researchers who do not consider carefully their research question. In Chapter 3, Kim et al. perform a case study of Tweets on health-care reform to determine whether such analysis could ever replace opinion polls on the topic.

Chapters 4 to 7 present examples of use of the conversational properties of social media for survey research, which is the second level of the sociality hierarchy. In Chapter 4, Sage describes using a Facebook application to build a sample of respondents and collect survey data by conversing with them in the special world of Facebook. In Chapter 5, Dean et al. demonstrate how researchers can use the virtual world Second Life and Skype videoconferencing software to conduct cognitive interviews (a conversation between researcher and participant) with a geographically dispersed population. In Chapter 6, Richards et al. use Second Life as a survey laboratory to test comprehension and compliance with the randomized response technique, a method of increasing the privacy of sensitive questions in surveys. In Chapter 7, Roe et al. describe the processes involved and decisions made in building a mobile survey panel that will, again, enable direct one-to-one “conversations with a purpose” between researchers and respondents.

Chapters 8 to 11 examine the community level of the sociality hierarchy. In Chapter 8, Keating describes how crowdsourcing techniques can be used to supplement survey research. In Chapter 9, Richards et al. present a method for using Twitter to collect diary data from specific Twitter-friendly communities. In Chapter 10, Haque et al. use extant social networks in Second Life to recruit and interview subjects with chronic medical conditions. Finally, in Chapter 11, Puleston describes methods for gamifying surveys—making them more interactive, interesting, and fun for respondents and, in effect, building communities of eager survey participants.

In the last chapter (Chapter 12), we use the sociality hierarchy to think of ways to improve the survey research of the future.

We consider ourselves fortunate to be employed by RTI International—an institute that considers innovation to be one of its core values. To foster this innovation (in survey research and all branches of science), RTI makes available internal funding designed to advance the science in the many disciplines under its roof. The great majority of the chapters in this book are tests, ideas, and experiments funded through this program, and we are eternally grateful to RTI for that opportunity.

Further examples and insights from this research are routinely shared in our blog SurveyPost (<http://blogs.rti.org/surveypost>), where readers can review, comment, and discuss topics germane to the future of survey and social research. One advantage of SurveyPost, of course, is that it allows us to respond quickly to new developments in technology and communications and survey research. This book has taken a comparatively longer length of time and, as a result, may contain some references that are out-of-date before it is printed, solely because of the speed of change in the social media and communications world.

Our intended audience for this book is the survey research community. However, our own backgrounds are quite diverse, having come to survey research from several different vectors and disciplines, so we hope that the book has broad appeal and finds interest among sociologists, political scientists, and psychologists, as well as those from the communications field, human–computer interaction researchers, market researchers, and all interested in the conduct of social science, both now and in the future.

CRAIG A. HILL
ELIZABETH DEAN
JOE MURPHY

April 2013

Acknowledgments

We would like to thank and acknowledge the following people for their contributions to this book: Ian Conlon, for review and comment on draft chapters; Laura Small, for copyediting support; John Holloway, for graphics coordination; Rebecca Powell, for data management and data analysis assistance; and Jodi Swicegood, for literature review assistance. We also want to thank RTI International for its support.

Tayo Jolaoso (RTI) developed the sociality hierarchy graphic in Chapter 1. In Chapter 3, the Kaiser Family Foundation provided public opinion poll data. Sarah Cho, Liz Hamel, and Mollyann Brodie of the Kaiser Family Foundation provided valuable feedback on an early draft of the chapter. The authors of Chapter 5 would like to acknowledge contributions made by Fred Conrad, of the Joint Program in Survey Methodology, to the avatar similarity instrument content and cognitive interview study design. The authors of Chapter 6 would like to acknowledge the vital work of John Holloway, who designed and developed the coin toss mechanism that was used in *Second Life*.

Contents

List of Figures	xiii
List of Tables	xvii
Contributors	xix
Preface	xxi
Acknowledgments	xxv

1. Social Media, Sociality, and Survey Research **1**

Joe Murphy, Craig A. Hill, and Elizabeth Dean

What Is Social Media? 2

Social Media Origins 6

Social Networking Sites and Platforms 6

Blogs 8

Twitter 8

Facebook 9

LinkedIn 9

Second Life 9

*Other Social Networking Platforms and
Functionalities* 10

Why Should Survey Researchers Be Interested
in Social Media? 11

The Current State of Survey Research 11

Falling Response Rates 11

Frame Coverage Errors 13

The Coming Age of Ubiquity 14

Public vs. Private Data 17

Social Media Interaction: Next Wave (or Subwave)? 18

Adding Social Media to the Survey Research Toolbox 21

Toward Using the Concept of Sociality in Survey

Research of the Future 22

How Can Survey Researchers Use Social Media Data? 26

References 28

2. Sentiment Analysis: Providing Categorical Insight into Unstructured Textual Data	35
<i>Carol Haney</i>	
Describing Emotional or Subjective Feeling in Textual Data	36
Definition of Machine-Augmented Sentiment Analysis	37
<i>How Sentiment Analysis Is Used with Text Data</i>	38
Different Ways of Representing Sentiment	42
<i>Ordinal Scales</i>	42
<i>Nominal Emotion Classification</i>	43
<i>Neutral Sentiment</i>	44
Techniques for Determining Sentiment	44
<i>Precursors to Analysis</i>	44
<i>Harvesting</i>	46
<i>Structure and Understand</i>	50
Approaches to Determining Sentiment	51
<i>Machine-Coded Sentiment Analysis</i>	51
<i>Human-Coded Sentiment Analysis</i>	53
Sentiment Analysis as a Subset of Text Analytics	54
Current Limitations of Sentiment Analysis	57
References	59
 3. Can Tweets Replace Polls? A U.S. Health-Care Reform Case Study	 61
<i>Annice Kim, Joe Murphy, Ashley Richards, Heather Hansen, Rebecca Powell, and Carol Haney</i>	
Methods	64
<i>Twitter Data</i>	64
Public Opinion About Health-Care Reform: Kaiser Health Tracking Poll	70
Analysis	70
Results	71
<i>RQ1: To What Extent Was Health-Care Reform Discussed on Twitter?</i>	71
<i>RQ2: What Is the Distribution of Sentiment of Health-Care Reform Tweets?</i>	74
<i>RQ3: Do Trends in the Sentiment of Tweets About Health-Care Reform Correlate with Observed Trends</i>	