

Finding Your Social Science Project

The most important step in social science research is the first step – finding a topic. Unfortunately, little guidance on this crucial and difficult challenge is available. Methodological studies and courses tend to focus on theory testing rather than theory generation. This book aims to redress that imbalance. The first part of the book offers an overview of the book’s central concerns. How do social scientists arrive at ideas for their work? What are the different ways in which a study can contribute to knowledge in a field? The second part of the book offers suggestions about how to think creatively, including general strategies for finding a topic and heuristics for discovery. The third part of the book shows how data exploration may assist in generating theories and hypotheses. The fourth part of the book offers suggestions about how to fashion disparate ideas into a theory.

John Gerring is Professor of Government at the University of Texas at Austin. He is the author of several books, most recently: *The Production of Knowledge: Enhancing Progress in Social Science* (Cambridge University Press, 2020; with Colin Elman & James Mahoney), and *Population and Politics: The Impact of Scale* (Cambridge University Press, 2020; with Wouter Veenendaal), along with numerous articles. He is co-editor of *Strategies for Social Inquiry*, a book series at Cambridge University Press.

Jason Seawright is Professor of Political Science at Northwestern University. He is the author of *Party-System Collapse: The Roots of Crisis in Peru and Venezuela* (2012), *Multi-Method Social Science: Combining Qualitative and Quantitative Tools* (Cambridge University Press, 2016), and *Billionaires and Stealth Politics* (2019, with Benjamin I. Page and Matthew Lacombe), along with numerous articles.

Strategies for Social Inquiry

Finding Your Social Science Project

The Research Sandbox

Editors

Colin Elman, *Maxwell School of Syracuse University*

John Gerring, *University of Texas, Austin*

James Mahoney, *Northwestern University*

Editorial Board

Bear Braumoeller, David Collier, Francesco Guala, Peter Hedström, Theodore Hopf, Uskali Maki, Rose McDermott, Charles Ragin, Theda Skocpol, Peter Spiegler, David Waldner, Lisa Wedeen, Christopher Winship

This book series presents texts on a wide range of issues bearing upon the practice of social inquiry. Strategies are construed broadly to embrace the full spectrum of approaches to analysis, as well as relevant issues in philosophy of social science.

Published Titles

Jennifer Widner, Michael Woolcock and Daniel Ortega Nieto (eds.), *The Case for Case Studies: Methods and Applications in International Development*

Colin Elman, John Gerring and James Mahoney (eds.), *The Production of Knowledge: Enhancing Progress in Social Science*

John Boswell, Jack Corbett and R. A. W. Rhodes, *The Art and Craft of Comparison*

John Gerring, *Case Study Research: Principles and Practices, 2nd edition*

Jason Seawright, *Multi-Method Social Science: Combining Qualitative and Quantitative Tools*

Peter Spiegler, *Behind the Model: A Constructive Critique of Economic Modeling*

James Mahoney and Kathleen Thelen (eds.), *Advances in Comparative-Historical Analysis*

Diana Kapiszewski, Lauren M. MacLean and Benjamin L. Read, *Field Research in Political Science: Practices and Principles*

Andrew Bennett and Jeffrey T. Checkel (eds.), *Process Tracing: From Metaphor to Analytic Tool*

Nicholas Weller and Jeb Barnes, *Finding Pathways: Mixed-Method Research for Studying Causal Mechanisms*

Thad Dunning, *Natural Experiments in the Social Sciences: A Design-Based Approach*

Carsten Q. Schneider and Claudius Wagemann, *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*

Michael Coppedge, *Democratization and Research Methods*

John Gerring, *Social Science Methodology: A Unified Framework, 2nd edition*

Finding Your Social Science Project

The Research Sandbox

John Gerring

The University of Texas at Austin

Jason Seawright

Northwestern University



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781009100397

DOI: 10.1017/9781009118620

© John Gerring and Jason Seawright 2022

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2022

A catalogue record for this publication is available from the British Library.

ISBN 978-1-009-10039-7 Hardback

ISBN 978-1-009-11491-2 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

	<i>List of Figures</i>	<i>page</i> xiii
	<i>List of Tables</i>	xiv
	<i>Acknowledgments</i>	xv
	<i>Glossary</i>	xvi
1	Introduction	1
<hr/>		
Part I	Overview	17
2	Current Practices	19
3	Contributions to Knowledge	38
<hr/>		
Part II	Playing with Ideas	71
4	Strategies	73
5	Heuristics	102
<hr/>		
Part III	Playing with Data	133
6	Case Selection	135
7	Soaking and Poking	150

Part IV	Playing with Theories	179
8	Theoretical Frameworks	181
9	Explanatory Challenges	203
10	Tools and Tips for Theorizing	218
Part V	Conclusions	243
11	From Exploration to Appraisal	245
	<i>Appendix A: Surveys and Interviews</i>	261
	<i>References</i>	300
	<i>Index</i>	332

Detailed Contents

1	Introduction	1
	An Eclectic, Holistic View of Scientific Exploration	3
	Putting the Pieces Together	4
	A Sandbox of Options	6
	This Book	7
	Special Features	8
	Surveys	8
	Examples	9
	Causal Diagrams	9
	What Lies Ahead	11
	<i>Coda</i> : Extant Work on Our Subject	13
	A Brief Survey	14
	Conclusions	16
2	Current Practices	19
	Identifying a Topic	20
	Process and Outcome	26
	Unpublished Work	29
	Intellectual Trajectories	34
	One Big Question	34
	Interconnected Questions	35
	Periodic Reinvention	36
	Conclusions	36
3	Contributions to Knowledge	38
	Phenomena	40
	The Present and the Past	41
	Discussion	42
	Concepts	42

	Clarification and Reconstruction	43
	Neologism	44
	Typologies	45
	Discussion	46
	Data and Measures	47
	Data Collection Projects	47
	The Role of New Data in Theoretical Debates	50
	Discussion	51
	Causes	51
	Another Cause of Y	51
	Distal Factors	53
	Discussion	54
	Mechanisms	54
	Discussion	55
	Research Designs	56
	Experiments	57
	Simple Natural Experiments	58
	Regression-Discontinuity Designs	59
	Instrumental-Variable Designs	60
	Discussion	62
	Empirical Refinements and Extensions	62
	Sensitivity	63
	Functional Form	63
	Moderators and Scope-Conditions	65
	Replications	67
	Conclusions	68
4	Strategies	73
	Work and Passion	74
	Ambition	76
	Scholarly Interaction	78
	Pedagogy	78
	Professional Meetings	80
	Networking	81
	Collaboration	81
	The Literature	84
	Unmet Demand	84
	Knowledge Frontiers	86
	Attitudes toward the Forebearers	86
	Intellectual History	87
	Reading Selectively	89

	The State of the World	90
	Translating Outrage into Science	90
	The Times	92
	The Familiar and the Unfamiliar	93
	Foreign Turf	94
	Specializing and Generalizing	96
	Something Old or Something New	98
	Something Old	98
	Something New	99
	Something Very Old and Therefore New	100
	Conclusions	100
5	Heuristics	102
	Thinking	104
	Time to Think	105
	Thinking along Multiple Tracks	105
	Thinking Softly	106
	Asking Questions	107
	Play	108
	Skepticism toward Words and Numbers	111
	Error and Anomaly	113
	Analogies	115
	Intellectual Arbitrage	118
	Thought Experiments	119
	Causal Effects and Causal Mechanisms	123
	Cases and Samples	124
	Hermeneutics	125
	Abstraction	127
	Failures	128
	Conclusions	131
6	Case Selection	135
	Extreme	136
	Index	139
	Deviant	140
	Most-Similar	142
	Diverse	145
	Conclusions	147

x	Detailed Contents	
<hr/>		
7	Soaking and Poking	150
	Ethnography	152
	Archives	159
	Statistics	163
	Conclusions	176
8	Theoretical Frameworks	181
	Motivational Frameworks	182
	Interests	183
	Norms	185
	Psychology	186
	Structural Frameworks	187
	Material Factors	187
	Human Capital/Demography	188
	Networks	189
	Institutions	190
	Interactive Frameworks	191
	Adaptation	191
	Coordination	192
	Diffusion	194
	Path Dependence	195
	Clarifications	196
	An Example: <i>Birth</i>	197
	Conclusions	199
9	Explanatory Challenges	203
	Criteria of a Good Theory	204
	Parsimony v. Impact	204
	Abstraction v. Realism	206
	Precision v. Generality	207
	Conclusions	207
	Causes and Effects	207
	One-to-One Theories	208
	Many-to-One Theories	208
	One-to-Many Theories	209
	Mechanisms	210
	A Cautionary Note	211
	Units and Levels of Analysis	214
	Conclusions	216

10	Tools and Tips for Theorizing	218
	The Spoken Word	218
	The Written Word	219
	Pictures	220
	Graphs	221
	Causal Diagrams	221
	Venn Diagrams	223
	Multidimensional Diagrams	224
	Decision Trees	225
	Maps	226
	Networks	228
	Key Cases	229
	Formal Models	231
	General Guidelines	235
	Hypotheses	237
	Alternate Explanations	238
	Tractability	240
11	From Exploration to Appraisal	245
	Vetting and Refining	248
	Costs, Payoffs, and Risks	249
	Ethics	250
	Market-Testing	251
	Going Public with Exploratory Work	252
	What to Reveal about the Exploratory Process	255
	Toward Appraisal	257
	<i>Appendix A: Surveys and Interviews</i>	261
	Data Sources	261
	Graduate Student Survey	263
	Practitioner Survey	264
	Interviews	265
	Social Desirability Bias	265
	Generalizability	266
	Pre-PhD Questionnaire	267
	Introduction	267
	The Dissertation	268
	Personal Info	273
	Open-Ended Questions	274
	Post-PhD Questionnaire	275

Introduction	275
The Work	276
File Drawer	280
Personal Info	281
Open-Ended Questions	283
Descriptive Information	286
Cluster Analysis	286
Innovation	288
Impact	290
Baseline Factors	292
Processual Factors	296
Methodological Considerations	298
<i>References</i>	300
<i>Index</i>	332

Figures

1.1 A causal diagram	<i>page</i> 10
2.1 Finding a dissertation topic	22
2.2 Activities	23
2.3 Entrée	24
2.4 Inspiration	25
2.5 Motivation	26
2.6 Process	28
2.7 Hunches	29
2.8 Innovation	30
2.9 The size of the file drawer	31
2.10 The status of projects in the file drawer	32
2.11 Obstacles encountered by projects in the file drawer	33
2.12 Archetypal intellectual biographies	34
7.1 Correlation heat map for infant mortality	166
7.2 Trellis scatterplot of democracy and infant mortality	167
9.1 A typology of causes and effects	208
9.2 A typology of mechanisms	211
9.3 The residual status of mechanisms	212
10.1 The demographic transition illustrated	221
10.2 A complex set of interactions in a causal diagram	222
10.3 A Venn diagram of American Indian identities	223
10.4 Cross-cutting cleavages	224
10.5 Decision tree showing reasons for gold mining in the Amazon	225
10.6 Map of cholera outbreaks in London	227
10.7 A Facebook social network	228
10.8 A spatial model of vote choice and party positioning	233
10.9 A branching theory	236

Tables

3.1	Illustration of a matrix typology of regime types	<i>page</i> 46
3.2	Occasions for regression discontinuity designs	60
3.3	Replication types	67
6.1	Exploratory case-selection strategies	136
6.2	Extreme case exemplars	137
6.3	Index case exemplars	139
6.4	Deviant case exemplars	141
6.5	Exploratory most-similar case design	143
6.6	Most-similar (exploratory) case exemplars	144
6.7	Diverse case exemplars	146
6.8	Diverse cases in Moore (1966)	147
7.1	Cluster means	170
7.2	LASSO regression results	173
7.3	Random forest variable importance	174
8.1	Explanatory frameworks	182
8.2	Prisoner's dilemma	193
9.1	Criteria of a good causal theory	205
10.1	Table of key cases	230
11.1	Two stages of research	248

Acknowledgments

This book grew out of several previous projects, from which we draw ideas and material. Chapter 2 draws on Gerring, Seawright, Shen (2020). Chapter 6 draws on Gerring (2017: chapter 5, section 5.1). Chapter 8 draws on Gerring, Christenson (2017: chapter 5). Portions of the book were presented to the research workshop at the Institute of Qualitative and Multimethod Research at Syracuse University, in June 2019.

For feedback on various drafts we owe thanks to Colin Elman, Mike Findley, Laura Garcia, Gary Goertz, Alan Jacobs, Diana Kapiszewski, Markus Kreuzer, Charlie Kurzman, Jennifer Larson, Carsten Schneider, and Steven Wilson. We also benefited from three excellent reviews as part of the formal review by Cambridge University Press, which helped immeasurably in revising the final version of the manuscript.

We are especially grateful to students and faculty who took time out of their busy schedules to share their experiences and their wisdom on the subject of exploratory research. In some cases, this meant filling out a lengthy survey, in other cases, interviews or discussions with the authors. We hope that this book offers some payoff for their efforts.

Glossary

This short lexicon is limited to terms used repeatedly in the book. Other specialized terms are defined in the text, with chapters of special relevance noted in the final column.

		<i>Chap.</i>
CAUSALITY		
Outcome (<i>Y</i>)	The phenomenon to be explained. Synonyms: dependent variable, explanandum.	1
Causal factor (<i>X</i>)	Hypothesized to affect <i>Y</i> . Synonyms: treatment, independent variable.	1
Antecedent cause (<i>A</i>)	Affects <i>X</i> and may serve as an instrument in a two-stage analysis.	1
Common-cause confounder (<i>C</i>)	Affects both <i>X</i> and <i>Y</i> , and thus potentially confounds the relationship.	1
Mechanism (<i>M</i>)	Conduit through which <i>X</i> affects <i>Y</i> . Synonyms: pathway, intermediate variable, mediator.	1
Noise (<i>B</i>)	Cause of <i>Y</i> uncorrelated with <i>X</i> .	1
Moderator (<i>D</i>)	Moderates the relationship between <i>X</i> and <i>Y</i> .	1
Background factor (<i>Z</i>)	A factor of no theoretical interest, e.g., <i>A</i> , <i>C</i> , <i>B</i> , or <i>D</i> .	1
RESEARCH DESIGNS		
Data generating process (DGP)	How the data under investigation was produced.	
Experiment	Where the treatment of theoretical interest is randomized by the researcher across two or more groups.	2
Natural experiment	Where the treatment of theoretical interest is assigned in an as-if random fashion, but not by the researcher. Variants include regression discontinuity and instrumental variables.	2
Observational	Nonexperimental.	2
Case study	One or several cases form the basis for an argument about a larger population.	4
Case selection	Methods of selecting cases for case study research, e.g., extreme, index, deviant, most-similar, diverse.	4

Continued

ARGUMENTS

Theoretical framework	The broadest sort of theory, diffuse to the point of unfalsifiability.	9
Theory	Explanation for an outcome or set of outcomes (if causal).	10
Hypothesis	Observable, and hence testable, implication of a theory.	10

RESEARCH STAGES

Exploratory	Prior to the identification of a research question, theory, and hypothesis. Synonyms: discovery, theory generation.	11
Confirmatory	After a research question, theory, and hypothesis have been identified. Synonyms: appraisal, confirmation, falsification, justification, testing, verification.	11

Cambridge University Press & Assessment
978-1-009-10039-7 — Finding Your Social Science Project
John Gerring , Jason Seawright
Frontmatter
[More Information](#)
