

Matthew B. Miles • A. Michael Huberman • Johnny Saldaña

Qualitative Data Analysis

EDITION
4

A METHODS SOURCEBOOK



Sara Miller McCune founded SAGE Publishing in 1965 to support the dissemination of usable knowledge and educate a global community. SAGE publishes more than 1000 journals and over 800 new books each year, spanning a wide range of subject areas. Our growing selection of library products includes archives, data, case studies and video. SAGE remains majority owned by our founder and after her lifetime will become owned by a charitable trust that secures the company's continued independence.

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne

Qualitative Data Analysis

A Methods Sourcebook

Fourth Edition

Matthew B. Miles
A. Michael Huberman
Johnny Saldaña
Arizona State University



Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne



FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

Acquisitions Editor: Helen Salmon
Content Development Editor: Chelsea Neve
Editorial Assistant: Megan O'Heffernan
Production Editor: Andrew Olson
Copy Editor: Liann Lech
Typesetter: C&M Digital (P) Ltd.
Proofreader: Caryne Brown
Indexer: Jeanne Busemeyer
Cover Designer: Scott Van Atta
Marketing Manager: Susannah Goldes

Fourth Edition Copyright © 2020 by SAGE Publications, Inc. First and Second Edition Copyright © 1994 by Matthew B. Miles and A. Michael Huberman. Third Edition Copyright © 2014 SAGE Publications, Inc.

All rights reserved. Except as permitted by U.S. copyright law, no part of this work may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without permission in writing from the publisher.

All third party trademarks referenced or depicted herein are included solely for the purpose of illustration and are the property of their respective owners. Reference to these trademarks in no way indicates any relationship with, or endorsement by, the trademark owner.

Printed in the United States of America

ISBN: 978-1-5063-5307-4

This book is printed on acid-free paper.

19 20 21 22 23 10 9 8 7 6 5 4 3 2 1

DETAILED CONTENTS

List of Displays	xiii
Preface to the Fourth Edition	xvii
A Note on This Revision	xvii
Acknowledgments	xix
About the Authors	xxi

PART I • THE SUBSTANTIVE START

Chapter 1 • Introduction	3
The Purpose of This Book	3
The Nature of This Book	4
Audiences	4
Our Orientation	5
An Approach to Qualitative Data Analysis	6
Analytic Methods: Some Common Features	6
The Nature of Qualitative Data	7
General Nature	7
Strengths of Qualitative Data	7
Our View of Qualitative Data Analysis	8
Data Condensation	8
Data Display	9
Drawing and Verifying Conclusions	9
Suggestions for Readers	10
Students and Other Novice Researchers	11
Experienced Researchers	11
Teachers of Qualitative Research Methods Courses	12
Closure and Transition	12
Chapter 2 • Research Design and Data Management	13
Introduction	14
Loose Versus Tight Research Designs	14
Displaying the Conceptual Framework	15
Description and Rationale	15
Example	15
Other Examples	18
Conceptual Framework Advice	20
Methodologies (Genres) of Qualitative Research	20
Formulating Research Questions	22
Description and Rationale	22
Example	22
Research Question Advice	24
Defining the Case	24
Description and Rationale	24
Examples	24
Case Advice	26

Sampling: Bounding the Collection of Data	26
Description and Rationale	26
Key Features of Qualitative Sampling	27
General Sampling Strategies	27
Within-Case Sampling	28
Multiple-Case Sampling	29
Sampling Advice	30
Instrumentation	30
Description and Rationale	30
Example	33
Instrumentation Advice	34
Linking Qualitative and Quantitative Data	35
Approaches to Mixed Methods Designs	35
Examples	36
Mixed Methods Advice	39
Data Management Issues Bearing on Analysis	39
Data Management	39
Computer and Software Use	40
CAQDAS	41
Staffing and Time Planning	46
Closure and Transition	47
Chapter 3 • Ethical Issues in Analysis	49
Introduction	49
Agreements With Study Participants	50
Agreement Advice	51
Ethical Issues	53
Worthiness of the Project	53
Competence	53
Informed Consent	53
Benefits, Costs, and Reciprocity	54
Harm and Risk	55
Honesty and Trust	55
Privacy, Confidentiality, and Anonymity	56
Intervention and Advocacy	57
Research Integrity and Quality	57
Ownership of Data and Conclusions	58
Use and Misuse of Results	59
Ethical Conflicts and Dilemmas	59
Ethics Advice	60
Closure and Transition	60
Chapter 4 • Fundamentals of Qualitative Data Analysis	61
Introduction	62
First Cycle Codes and Coding	62
Description	62
Rationale	64
First Cycle Coding Examples	64
Descriptive Coding	65
In Vivo Coding	65
Process Coding	66
Concept Coding	66
Emotion Coding	67
Values Coding	67

Evaluation Coding	68
Dramaturgical Coding	68
Holistic Coding	69
Provisional Coding	69
Hypothesis Coding	70
Protocol Coding	70
Causation Coding	70
Attribute Coding	71
Magnitude Coding	71
Subcoding	72
Simultaneous Coding	72
Themeing the Data	73
Coding Craft and Design	74
Levels of Coding Detail	74
A Priori Codes	74
Revising Codes	75
Structure and Unity in Code Lists	75
Definitions of Codes	77
Second Cycle Coding—Pattern Codes	79
Description	79
Applications	79
Examples	80
From Codes to Patterns	82
Narrative Description	83
Matrix Display	83
Network Display	84
Graphic Display	85
Coding Advice	86
Jottings	86
Analytic Memoing	88
Description and Rationale	88
Examples	89
Memoing Advice	90
On Visual Data	91
Hypotheses, Assertions, and Propositions	93
Within-Case and Cross-Case Analysis	95
Purposes of Cross-Case Analysis	95
A Key Distinction: Variables Versus Cases	95
Strategies for Cross-Case Analysis	96
Closure and Transition	99

PART II • DISPLAYING THE DATA

Chapter 5 • Designing Matrix, Network, and Graphic Displays	103
Introduction	103
Display Format Options	104
Matrices	105
Networks	109
Graphics	110
Timing of Display Design	115
Entering Matrix, Network, and Graphic Data	116
Making Inferences and Drawing Conclusions From Matrix, Network, and Graphic Displays	116

The Methods Profiles	118
Closure and Transition	118
Chapter 6 • Methods of Exploring	119
Introduction	119
Exploring Field Work in Progress	120
Data Accounting Log	120
Contact Summary Form	122
Case Analysis Meeting	125
Interim Case Summary	128
Partially Ordered Meta-Matrix	131
Exploring Variables	135
Code Frequency Table	135
Checklist Matrix	137
Content-Analytic Summary Table	142
Contrast Table	144
Two-Variable Case-Ordered Matrix	146
Exploring Reports in Progress	148
Prestructured Case	148
Sequential Analyses	151
Closure and Transition	153
Chapter 7 • Methods of Describing	155
Introduction	155
Describing Participants	156
Demographics and Attributes Table	156
Role-Ordered Matrix	157
Context Chart	161
Positionality Graphic	164
Describing Variability	166
Concept Table	166
Conceptually Clustered Matrix	168
Folk Taxonomy	173
Thematic Array	176
Describing Action	180
Vignettes	180
Poetic Display	183
Cognitive Networks	185
Closure and Transition	188
Chapter 8 • Methods of Ordering	189
Introduction	189
Ordering by Time	190
Qualitative Timeline	190
Event Listing Matrix	192
Growth Gradient	195
Time-Ordered Matrix	198
Ordering Processes	202
Decision Modeling	202
Process Mapping	204
Event-State Network	207
Composite Sequence Analysis	209
Ordering by Cases	212

Case-Process Comparison	212
Case-Ordered Descriptive Meta-Matrix	215
Closure and Transition	220
Chapter 9 • Methods of Explaining	221
Introduction	221
Antecedent Conditions, Mediating Variables, and Outcomes	222
On Explanation and Causation	224
Explaining Interrelationship	225
Variable-by-Variable Matrix	225
Explaining Change	230
Explanatory Matrix	230
Outcomes Matrix	232
Case Dynamics Matrix	234
Explaining Causation	237
Causal Chains	237
Causal Network: Within-Case Analysis	238
Causal Network: Cross-Case Analysis	248
Closure and Transition	254
Chapter 10 • Methods of Predicting	255
Introduction	255
Methods of Predicting	256
Making and Testing Predictions	256
Predictor–Outcome–Consequences Matrix	261
Causal–Prediction Model	265
Closure and Transition	270
PART III • MAKING GOOD SENSE	
Chapter 11 • Drawing and Verifying Conclusions	273
Introduction	274
Tactics for Generating Meaning	274
1. Noting Patterns, Themes	274
2. Seeing Plausibility	275
3. Clustering	276
4. Making Metaphors	277
5. Counting	279
6. Making Contrasts, Comparisons	280
7. Partitioning Variables	281
8. Subsuming Particulars Into the General	281
9. Factoring	282
10. Noting the Relations Between Variables	283
11. Finding Mediating Variables	285
12. Building a Logical Chain of Evidence	286
13. Making Conceptual/Theoretical Coherence	287
Tactics for Testing or Confirming Findings	289
1. Checking for Representativeness	290
2. Checking for Researcher Effects	291
3. Triangulating	293
4. Weighting the Evidence	294
5. Checking the Meaning of Outliers	296
6. Using Extreme Cases	297

7. Following Up Surprises	297
8. Looking for Negative Evidence	298
9. Making If-Then Tests	298
10. Ruling Out Spurious Relations	299
11. Replicating a Finding	301
12. Checking Out Rival Explanations	302
13. Getting Feedback From Participants	303
Standards for the Quality of Conclusions	304
Objectivity/Confirmability	305
Reliability/Dependability/Auditability	305
Internal Validity/Credibility/Authenticity	306
External Validity/Transferability/Fittingness	307
Utilization/Application/Action Orientation	308
Analytic Documentation	308
The Problem	308
Examples	310
Closure and Transition	314
Chapter 12 • Writing About Qualitative Research	315
Introduction	315
Audiences and Effects	316
The Reader and the Writer	316
Types of Effects	317
Voices and Styles	318
Writing Examples and Recommendations	319
Writing the Abstract	319
Writing About Methodology and Methods	321
“Being There”	322
Analytic Storytelling With Displays	324
Traditional Presentation Modes	326
Progressive Presentation Modes	327
On Theses and Dissertations	328
Support and Supervision	328
Writing Fundamentals	329
Closure and Transition	330
Chapter 13 • Closure	331
Qualitative Analysis at a Glance	331
Reflections	333
Final Advice	334
Appendix: An Annotated Bibliography of Qualitative Research Methods Resources	337
References	359
Index	365

LIST OF DISPLAYS

Display 1.1	Components of Qualitative Data Analysis: Interactive Model	10
Display 2.1	A First-Draft Conceptual Framework for a Case Study Teacher and the Influences on Her Practice	16
Display 2.2	Major Influences on a Language Arts Teacher's Practice	17
Display 2.3	A Conceptual Model of Students' Intention to Pursue a Sales Career	19
Display 2.4	Conceptual Model That Presents HCWs' Perceived Limitations in the Application of MTB Infection Control Measures in Clinical Practice, Influencing the Decision-Making Process and Resulting in the "Knowledge-Action" Gap	19
Display 2.5	General and Special Research Questions Relating to the Adoption Decision (School Improvement Study)	23
Display 2.6	The Case as the Unit of Analysis	25
Display 2.7	Prior Instrumentation: Key Decision Factors	31
Display 2.8	Excerpts From Interview Guide, School Improvement Study	34
Display 2.9	Mixed Methods Data-Analytic Procedures	38
Display 2.10	An Excel Spreadsheet With Qualitative and Quantitative Data	42
Display 2.11	Network View Produced by ATLAS.ti	42
Display 2.12	A Model of Lifelong Confidence From High School Speech and Theatre	43
Display 2.13	Word Clouds of Top 30 Most Frequent Words in the Mission (top), Vision (middle), and Values (bottom) Statements of the S&P 500 Companies	45
Display 2.14	A Screenshot of Quirkos CAQDAS Software	45
Display 2.15	A Screenshot of Transana CAQDAS Software	46
Display 3.1	Sample Informed Consent Letter	51
Display 4.1	Details of the Coding System	75
Display 4.2	Abbreviated Codebook	77
Display 4.3	Matrix of Smoking Cessation Patterns at Months 1 and 6.	84
Display 4.4	A Network Model of Smoking Cessation Loss Transformation	84
Display 4.5	Arts-Based Graphic Representation of Smoking Cessation at Months 1 and 6	85
Display 4.6	Interview Transcript With Jottings	87
Display 4.7	A Residential Exterior From the West Side of Chicago, 2018	92
Display 4.8	The Hypothesis Process	94
Display 4.9	Categories of Acute Precipitants of Physical Elder Abuse	98
Display 5.1	Effects Matrix: Assistance Location and Types (Masepa Case)	106
Display 5.2	A Network Model of "Lifelong Impact" From High School Speech Participation	109

Display 5.3	Percentage of 3,416 References Coded in NVivo, Attributed to the Three Components of Sustainability	111
Display 5.4	A Sustainability Definition: (A) Classic Triple Bottom-Line Approach and (B) Proposed Representation Emphasizing Time and Trade-Offs	111
Display 5.5	Impact of Correctional Environment Upon Individual Women	112
Display 5.6	Transformative Process of Prison Inmate Hospice Volunteer Service	113
Display 5.7	Axial Coding of the Six Primary SPELIT Environments of Donald Trump's Campaign Speech	113
Display 5.8	Characteristics of Systems Thinking Among Middle School Leaders	114
Display 5.9	Trajectory of Learning to Use Professional Judgment	115
Display 6.1	Data Accounting Log	121
Display 6.2	Contact Summary Form: Illustration (Excerpts)	122
Display 6.3	Contact Summary Form: Illustration With Themes/Aspects (Excerpts)	124
Display 6.4	Case Analysis Meeting Form	125
Display 6.5	Case Analysis Form: Exhibit With Data	127
Display 6.6	Summary-Aided Approach to Analysis	128
Display 6.7	Interim Case Summary Outline: Illustration	129
Display 6.8	Case-Level Display for Partially Ordered Meta-Matrix (Format)	131
Display 6.9	Case-Level Display for Partially Ordered Meta-Matrix: Users' Second Year of Implementation at Lido	132
Display 6.10	Partially Ordered Meta-Matrix: Users' Feelings/Concerns and Other Variables (Format)	133
Display 6.11	Partially Ordered Meta-Matrix: Users' Feelings/Concerns and Other Variables (Lido Data)	133
Display 6.12	Time-Ordered Meta-Matrix (Format)	134
Display 6.13	Summary Table: Individual and Institutional Concerns During Later Implementation	135
Display 6.14	Summary of Memorable Messages of Social Support	136
Display 6.15	Code Landscaping of "Lifelong Impact" Survey Responses	137
Display 6.16	Checklist Matrix: Conditions Supporting Preparedness at Smithson School, Banestown Case	138
Display 6.17	Checklist Matrix on Preparedness (Alternative Format 1)	141
Display 6.18	Checklist Matrix on Preparedness (Alternative Format 2)	142
Display 6.19	Checklist Matrix on Preparedness (Alternative Format 3)	142
Display 6.20	Content-Analytic Summary Table: The Content of Organization Changes	143
Display 6.21	Contrast Table: Summary of Contextual Factors Shaping Political Styles	145
Display 6.22	Two-Variable Case-Ordered Matrix: Relationships Between User Practice Stabilization and Local Continuation	147
Display 6.23	Prestructured Case Outline: Abbreviated Version	149
Display 6.24	Traditional Analysis Sequence Compared With Prestructured Case	151
Display 7.1	Participant Family Profiles	156
Display 7.2	Role-Ordered Matrix: First Reactions to the Innovation	158
Display 7.3	Context Chart for Tindale East High School and District	162
Display 7.4	Predominant Response Types of Cases	165
Display 7.5	Lifelong Impact: Variability of Influence	167

Display 7.6	Conceptually Clustered Matrix: Motives and Attitudes (Format)	169
Display 7.7	Conceptually Clustered Matrix: Motives and Attitudes of Users, Nonusers, and Administrators at Masepa	171
Display 7.8	A Folk Taxonomy of the Ways Children Oppress Each Other	174
Display 7.9	A Principal's Experiences and Views on Creating Collaborative Learning Among Teachers	177
Display 7.10	Nursing Professional Dignity Perceived as an Achievement	178
Display 7.11	Theory of Influences on Women Early Career Psychologists' [WECP] Career Adjustment	179
Display 7.12	Social Media Seesaw: Positive and Negative Affect Influences	179
Display 7.13	A Model in Progress of HIV Test Anxiety	185
Display 8.1	River School District [RSD] Timeline Showing 18 Years of Linear Change.	190
Display 8.2	Event Listing, Banestown Case	193
Display 8.3	Event History of a Case Study	196
Display 8.4	Growth Gradient for ECRI Innovation, Masepa Case	197
Display 8.5	Time-Ordered Matrix: Changes in the CARED Innovation (a Work Experience Program)	199
Display 8.6	Summary Tables for Verifying and Interpreting a Time-Ordered Matrix: Changes in the CARED Innovation	201
Display 8.7	A Decision Model for Evening Dinner Choices	203
Display 8.8	A Process Map of a Chaos Narrative About Cancer	205
Display 8.9	Event-State Network, Banestown Case (Excerpt)	208
Display 8.10	Composite Sequence Analysis: Career Trajectory Data for 11 Cases	210
Display 8.11	Model of Instructional Decision Making for Two Physics Instructors	213
Display 8.12	Ordered Meta-Matrix: Format for Student Impact Data	216
Display 8.13	Case-Ordered Descriptive Meta-Matrix (Excerpt): Program Objectives and Student Impact (Direct, Meta-Level, and Side Effects)	218
Display 9.1	The Antecedents, Contingencies, and Follower Outcomes of the Interplay of the Emotion-Related Behaviors of Nonprofit Leaders and Nonprofit Organizational Display Norms	224
Display 9.2	Variable-by-Variable Matrix: Coping Strategies and Problems, by Case	226
Display 9.3	Summary Table: Typical Consequences of Coping, by Case	229
Display 9.4	Explanatory Matrix: Ongoing Assistance	231
Display 9.5	Outcomes Matrix: Organizational Changes After Implementation of the ECRI Program	232
Display 9.6	Case Dynamics Matrix: The IPA Innovation as a Force for Organizational Change in the District and Its Schools	235
Display 9.7	Causal Chain: Illustration	237
Display 9.8	Causal Chain: Illustration	238
Display 9.9	Causal Fragment: Mastery of a New Educational Practice	239
Display 9.10	Excerpt From a Causal Network: Perry-Parkdale School	242
Display 9.11	Excerpt From an Event-State Network: Perry-Parkdale School	242
Display 9.12	List of Antecedent Conditions, Mediating Variables, and Outcomes: School Improvement Study	243
Display 9.13	Causal Network for Perry-Parkdale CARED Program	245

Display 9.14	Narrative for Causal Network: Perry-Parkdale CARED Program	247
Display 9.15	Subnetwork: Variable Streams Leading to High Job Mobility, Perry-Parkdale Case	249
Display 9.16	Subnetwork for Job Mobility, Calston Case	251
Display 9.17	Subnetwork for Job Mobility, Banestown Case	252
Display 9.18	Subnetwork for Job Mobility, Plummet Case	252
Display 10.1	Prediction Feedback Form	256
Display 10.2	Factors Supporting “Institutionalization” Prediction	259
Display 10.3	Factors Working Against “Institutionalization” Prediction	259
Display 10.4	Filled-Out Response Form From Case Informant for “Institutionalization” Prediction	260
Display 10.5	Predictor–Outcome–Consequences Matrix: Antecedents and Consequences of Assistance	263
Display 10.6	Causal–Prediction Model Tracing User Practice Changes	266
Display 10.7	Predictor–Outcome Matrix: Predictors of Magnitude of User Practice Change	267
Display 11.1	Reasons Given for Adoption by Users	279
Display 11.2	Two-Variable Relationship	285
Display 11.3	Two-Variable Relationship With Mediating Variables	285
Display 11.4	Example of a Chain of Evidence Supporting an Observed Outcome	286
Display 11.5	Possible Explanation of Spurious Relationship	300
Display 11.6	Display for Testing Explanations in Display 11.5	300
Display 11.7	Analysis of Data, Categories, and Emergent Themes	313
Display 12.1	Elements That Influence Nursing’s Professional Respect: Summary of the Analysis	325
Display 12.2	Professional Respect for Nurses in Hospital and Community Settings	326
Display 13.1	Overview of Qualitative Data Analysis Processes	332

PREFACE TO THE FOURTH EDITION

Johnny Saldaña

This fourth edition of Matthew B. Miles and A. Michael Huberman's classic SAGE Publishing text, *Qualitative Data Analysis*, continues to update the late authors' unique work for new generations of qualitative researchers as well as for the dedicated followers of their methods for almost 40 years. I am honored to join them, in spirit, as the third author of this revised text.

This book offers its readers practical guidance in recommended methods for assembling and analyzing primarily text-based data in visual arrays. *Qualitative Data Analysis: A Methods Sourcebook* is designed for researchers in virtually all fields and disciplines that honor what their human participants have to say, treasure the products and artifacts they create, and respect the complexity of social action as it happens all around us. It is intended primarily for students in graduate degree programs who are learning how to investigate the human condition through qualitative research coursework, and for established scholars and practitioners continuing their professional development by reading the literature on current methods.

A website at study.sagepub.com/miles4e includes a selected bibliography compiled by author Johnny Saldaña, plus downloadable versions of the book's Appendix and display templates.

A NOTE ON THIS REVISION

My charge as third coauthor was to *adapt* Miles and Huberman's work, not to write my own book on qualitative data analysis. For this fourth edition, I have continued to maintain the general spirit and integrity of the core contents of the original 1994 authoritative work, while making the text more accessible and relevant to contemporary researchers. I have added graphics as a display modality in addition to matrices and networks, and have included several new displays to the collection in Chapters 2, 4–9, and 11–12, based on the work of fellow scholars. The chapters on “Research Design and Data Management;” “Fundamentals of Qualitative Data Analysis;” “Designing Matrix, Network, and Graphic Displays;” “Methods of Describing;” “Methods of Ordering;” and “Writing About Qualitative Research” have been substantially revised.

Peer reviewers of the third edition publication and this edition's manuscript offered constructive recommendations for revision, and I have attempted to incorporate several of those ideas into this work. I must counter, however, that just one book alone cannot address absolutely all matters related to qualitative data analysis, lest it become a massive tome or multivolume series. There are excellent works devoted exclusively to related specialized topics such as mixed methods research and digital technology, and I refer my readers to those titles for supplemental information (see the extensive annotated bibliography in the Appendix for recommended resources).

Finally, despite the open-ended leeway I was granted to revise and update Miles and Huberman's classic work for two editions, I still do not feel that this is “my” book. Thus, they remain listed as first and second coauthors. I am a privileged guest in the academic house they've built. I'm responsible for the home's traditions and upkeep, but with generous allowances for my own decorative flair in selected rooms.

ACKNOWLEDGMENTS

I am honored that Helen Salmon, acquisitions editor of SAGE Publications' College Division, commissioned me to adapt Miles and Huberman's text for its third and fourth editions. My initial contact with SAGE began with their former London office editor, Patrick Brindle, who encouraged me to develop *The Coding Manual for Qualitative Researchers*, and for his welcoming invitation to this distinguished publishing house, I am truly grateful. I also thank Chelsea Neve, Susannah Goldes, Andrew Olson, Bennie Clark Allen, Scott Van Atta, Jeanne Busemeyer, Caryne Brown, and Megan O'Heffernan at SAGE Publishing for prepublication guidance, and Liann Lech for meticulous copyediting of this manuscript.

My own qualitative research methods professors at Arizona State University significantly influenced my growth as a scholar and writer. I am indebted to Tom Barone, Mary Lee Smith, Amira De la Garza, and Sarah J. Tracy for their life-changing impact on my academic career. For their collegial friendship, mentorship, and support of my work, I also thank Ann Whitehouse, Coleman A. Jennings, Lin Wright, Harry F. Wolcott, Norman K. Denzin, Yvonna S. Lincoln, Mitchell Allen, Jai Seaman, Ray Maietta, Ronald J. Chenail, Leo A. Mallette, Linda Essig, Kakali Bhattacharya, Patricia Leavy, Liora Bresler, Joe Norris, Laura A. McCammon, Angela Hines, Matt Omasta, and Jim Simpson.

For this particular edition, I also thank Daniel Turner of Quirkos Software and David K. Woods of Transana for screenshot permissions, Taylor & Francis Ltd for figure reproduction permission, Paul J. Mack for a photograph permission, and all the scholars whose work is included in this book for their exemplary qualitative research.

My final thanks go to Matthew B. Miles and A. Michael Huberman themselves. To my knowledge, I never met them or heard them speak at professional conferences, but their data-analytic methods, which I learned intimately in my qualitative research courses, have been part of my work ever since. I always strive to meet their rigorous standards, and I frequently quote their now-classic advice to "think display." I owe much of my career trajectory to the legacy of scholars before me whose methods books and articles helped shape my own ways of working as a qualitative researcher and data analyst. Miles and Huberman are two of those esteemed scholars, and I am honored to be connected with them in this way. I hope that the fourth edition of their book pays proper tribute and homage to their significant level of scholarship.

SAGE Publishing and Johnny Saldaña are grateful for feedback on the third edition and draft manuscript of the fourth edition from the following reviewers:

Sheila J. Foley, Bay Path University

Bronwynne Evans, Arizona State University

Caroline S. Hackerott, Arkansas Tech University

Casey B. Nixon, Piedmont College

Guofang Li, University of British Columbia

Jennifer Morrison, Johns Hopkins University

Karthigeyan Subramaniam, University of North Texas

ABOUT THE AUTHORS

Matthew B. Miles, a social psychologist, had a career-long interest in strategies for educational reform. His work focused on planned change in education, group and organizational studies, and the dissemination and implementation of research findings. His first major research project was a 4-year study of six new, innovative public schools.

A. Michael Huberman's long-term interests were in scientific epistemology, and adult cognition and knowledge use. His work focused on education policy, school reform, and the practical translation of research knowledge into effective school practice. His first extensive project was a 4-year study of an experimental elementary school's implementation of Piagetian theories in classroom practice.

Johnny Saldaña, a theatre educator, has explored qualitative research methodologies ranging from coding to ethnodrama. His work focuses on young people's and adults' participation in and transformation through the arts. His first qualitative research project was a 2½-year ethnography of a beginning teacher's experiences at an inner-city fine arts magnet school.

Miles and Huberman collaborated for many years on studies of the dissemination of educational innovations and school improvement. In the course of a multisite study sponsored by the National Institute of Education, they began a study of the procedures they were using to analyze data, which provided the intellectual stimulus and the content of much of this book. Miles and Huberman led seminars based on the book in the United States, Canada, western Europe, and Australia. Their work together and with other colleagues included the study of school change agents; innovation in urban high schools; the social architecture of schools; teachers' life cycles; the use of scientific knowledge in vocational education; and educational reform in Bangladesh, Ethiopia, New Zealand, and Colombia.

Saldaña's studies have included child audience responses to theatre; teachers' perceptions of achievement standards; a survey of adults reflecting on their high school experiences; longitudinal qualitative research methodology; coding and qualitative data analysis methods; and arts-based research methodologies, including ethnodrama and ethnotheatrical play production. His methods works have been cited and referenced in more than 10,000 research studies conducted in over 130 countries. Johnny Saldaña is Professor Emeritus from Arizona State University's School of Film, Dance, and Theatre, and is the author of SAGE Publications' *The Coding Manual for Qualitative Researchers* (3rd ed.); *Thinking Qualitatively: Methods of Mind*; and, with co-author Matt Omasta, *Qualitative Research: Analyzing Life*.

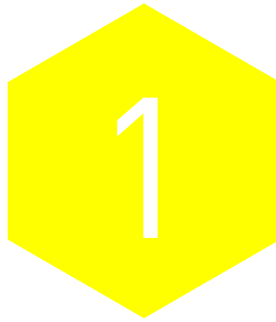
THE SUBSTANTIVE START

Chapter 1 Introduction

Chapter 2 Research Design and Data Management

Chapter 3 Ethical Issues in Analysis

Chapter 4 Fundamentals of Qualitative Data Analysis



INTRODUCTION

CHAPTER SUMMARY

This chapter introduces the reader to the coauthors' perspectives about the nature of qualitative data and the interrelated components of qualitative data analysis. We offer recommended guidance for the various audiences of this book and establish the parameters of what we cover in the text.

CONTENTS

The Purpose of This Book	Our View of Qualitative Data Analysis
The Nature of This Book	Data Condensation
Audiences	Data Display
Our Orientation	Drawing and Verifying Conclusions
An Approach to Qualitative Data Analysis	Suggestions for Readers
Analytic Methods: Some Common Features	Students and Other Novice Researchers
The Nature of Qualitative Data	Experienced Researchers
General Nature	Teachers of Qualitative Research Methods
Strengths of Qualitative Data	Courses
	Closure and Transition

THE PURPOSE OF THIS BOOK

We wrote this book to address a need faced by researchers in all disciplines and fields of human inquiry: The need to construct credible and trustworthy meaning from qualitative data through analytic methods that are practical and will generate knowledge that we and others can rely on.

Qualitative data are a source of well-grounded, rich descriptions of social processes. With qualitative data, one can preserve chronological flow to derive plausible explanations for outcomes. Good qualitative data also lead to serendipitous findings and interrelationships. They help researchers get beyond initial conceptions and generate new understandings. Finally, the findings from well-analyzed qualitative studies have a quality of “undeniability.” Words and images, especially organized

into vignettes, stories, and evocative visual representations, have a concrete, vivid, and meaningful flavor that often proves far more persuasive to a reader than pages of summary statistics.

The expansion of qualitative inquiry from the 1970s onward has been phenomenal. There are now numerous handbooks (e.g., Denzin & Lincoln's *The SAGE Handbook of Qualitative Research* [2018] and Gubrium, Holstein, Marvasti, & McKinney's *The SAGE Handbook of Interview Research* [2012]); introductory textbooks (e.g., Charmaz's *Constructing Grounded Theory* [2014] and Saldaña & Omasta's *Qualitative Research: Analyzing Life* [2018]); prestigious peer-reviewed journals (*Qualitative Inquiry*, *Qualitative Health Research*); online newsletters and forums (SAGE Publications' Methodspace, *The Qualitative Report*); annual conferences (International Congress of Qualitative Inquiry, International Institute for Qualitative Methodology); and qualitative special-interest groups in several major professional associations (American Educational Research Association, National Communication Association).

Although many researchers, from graduate students writing their theses and dissertations to experienced researchers, work alone on their projects and often focus on single cases, qualitative work is becoming more complex. Increasingly, we see mixed methods studies that combine qualitative and quantitative inquiry, carried out by research teams working collaboratively through specialized analytic software. And exciting new approaches to qualitative inquiry draw their inspiration from poetry, dramatic literature, visual art, dance, film, and other creative genres to present and represent social life in fresh, evocative ways.

Yet, in the flurry of this activity, we should be mindful of some pervasive issues that have not gone away. These issues include the labor intensiveness (and extensiveness over months or years) of data collection; frequent data overload; the time demands of processing and analyzing data; the adequacy of sampling when only a few cases can be managed; the generalizability and transferability of findings; the credibility, trustworthiness, and quality of conclusions; and their utility in the world of policy and action.

This book is written in the belief that, as qualitative researchers, we need to keep sharing our *craft and art*—that is, the explicit, systematic, and sometimes creative methods we use to analyze data. This is the need our book addresses. We hope that our experiences will be helpful to our colleagues, as theirs have been to us.

THE NATURE OF THIS BOOK

This is a practical methods sourcebook for all researchers who make use of qualitative data. But it is not intended as a comprehensive text; we do not address matters such as how to gain entry into a field site or how to facilitate participant interviews. Others have dealt with these issues elsewhere and well; we cite their work along the way and refer you to the appendix's annotated bibliography for more information.

For each of the methods outlined, we give specific examples with enough detail so that the reader can see how things work; try the method; and, most important, revise the method in future work. These analytic approaches are manageable and straightforward, though some displays may, at first, appear daunting. Don't let them intimidate you; they are *examples*, not exemplars or standards of practice. Each method of data analysis and display offers practical suggestions for the researcher's use and adaptation with accessible software and basic programs.

Audiences

This book is for practicing researchers in all fields whose work involves actual qualitative data analysis issues. An important subset of that audience is the *beginning researcher*—a graduate student or

early-career faculty member—working with qualitative data. We have encountered many students who launch into qualitative theses, dissertations, or research projects who sometimes feel overwhelmed and undertrained. With them in mind, we keep the language accessible and supportive, and we offer suggestions for using the book in qualitative research methods courses.

Many examples used in the book are drawn from educational research, both ours and others'. But we also include relevant discussion for other disciplines—nursing, anthropology, sociology, psychology, business, the arts, organization studies, sports, social media, political science, criminology, tourism, ethnic studies, communication, family studies, program evaluation, and policy research—to underline that the methods are generic, not field limited.

Some of the methods reported here grew out of multiple case studies of organizations carried out by a research team. But do not despair if you are working alone, if your study has just one case, or if you are focusing at the small-group level. There are many relevant examples for you, along with targeted advice.

OUR ORIENTATION

It is worthwhile, we think, for researchers to make their positionality and standpoint clear. To know how a researcher construes the shape of the social world and aims to give us a credible account of it is to know our conversational partner.

We label ourselves *pragmatic realists*. We believe that social phenomena exist not only in the mind but also in the world—and that some reasonably stable relationships can be found among the idiosyncratic messiness of life. There are regularities and sequences that link together phenomena. From these patterns, we can derive the concepts that underlie individual and social life. The fact that most of those concepts—such as *negotiating*, *possible selves*, *emotional labor*, and so on—are invisible to the human eye does not make them invalid.

Human relationships and societies have unique peculiarities and inconsistencies that make a realist approach to understanding them more complex—but not impossible. We must contend with the institutions, structures, practices, and conventions that people produce, reproduce, and transform. Human meanings and intentions are worked out within the frameworks of these social structures—structures that are invisible but nonetheless real. In other words, social phenomena such as language, culture, and power exist in the world and exert strong influences over human activities because most people construe them in common ways. Things that are believed become real and can be inquired into.

We agree with interpretivists who point out that knowledge is socially constructed. We affirm the existence and importance of the subjective, the phenomenological, and the meaning making at the center of social life. Our goal is to register and transcend these processes by making assertions and building theories to account for a real world that is both bounded and perceptually laden—and to test these assertions and theories in our various disciplines.

Our tests do not use the deductive logic of classical positivism. Rather, our explanations flow from an account of how human actions produced the events we observed. We want to account for events rather than simply document their sequence. We look for an individual or a social process at the core of events that can be captured to provide a causal description of the most likely forces at work.

The paradigms for conducting social research are always shifting beneath our feet. Our view is that it is possible to develop practical methods for judging the goodness of our findings, discoveries, and conclusions. To us, research is actually more a creative craft than a slavish adherence to

methodological rules. No study conforms exactly to a standard methodology; each one calls for the researcher to adapt the methodology and methods to the uniqueness of the setting or case.

Readers looking at the methods in this sourcebook will find them to be orderly ones with some degree of formalization. Many colleagues prefer more relaxed and open-ended voyages through their data, and we wish them well. We have opted for thoroughness and explicitness in this book, not just because it suits us but because vague descriptions are of little practical use to others. The overall structure of this text allows for some techniques to be used and others to be left aside. We advise you to look behind any apparent formalism and seek out what will be *useful* in your own work.

AN APPROACH TO QUALITATIVE DATA ANALYSIS

When you've been doing qualitative research as long as we have, the methodologies or genres start to blur. As pragmatic realists, we no longer adhere slavishly to one school of thought, or practice solely within the boundaries of one particular methodological approach. The data-analytic methods and techniques we've employed over the past few decades have been "a little bit of this and a little bit of that," used on an "as needed" basis. This is not to suggest that we were being improvisationally foolhardy as we figured out what to do next. On the contrary, we have been quite deliberate and diligent in our analytic methodologies and work. We resonate with ethnographer Amanda Coffey's (2018) insight that analysis "involves an ongoing dialogue with and between data and ideas" (p. 25).

Over time, the methods included in this book have become almost signature works—the "Miles and Huberman" way of analyzing qualitative data. But if you ask established qualitative researchers to describe Miles and Huberman's methods, you'll hear diverse opinions ranging from positive descriptors such as "systematic," "rigorous," and "authoritative," to negative descriptors such as "old fashioned," "confusing," and (a personal favorite) "positivists in sheep's clothing." Add Saldaña as a coauthor to the third and fourth editions of this text, and the evolution of these methods continues.

The three of us do not subscribe to any one particular genre of qualitative research—we are "shamelessly eclectic," as the popular saying goes. But our analytic sequence depicted throughout the book is probably a hybrid blend of ethnographic methods and grounded theory. It moves from one inductive inference to another by selectively collecting data, comparing and contrasting this material in the quest for patterns or regularities, seeking out more data to support or qualify these emerging clusters, and then gradually drawing inferences from the links between other new data segments and the cumulative set of conceptualizations. In the past few decades, we've found that refining and developing analysis methods on new projects had a clear payoff. Our confidence in our findings was greater, and credibility for our research, practice, and policy audiences was enhanced.

Analytic Methods: Some Common Features

We've observed features that recur in many established approaches to qualitative analysis. Here is a fairly classic set of analytic moves arranged in sequence:

1. Assigning codes or themes to a set of field notes, interview transcripts, documents, and/or visual data
2. Sorting and sifting through these coded materials to identify similar phrases, relationships between variables, patterns, categories, themes, distinct differences between subgroups, and common sequences

3. Isolating these patterns and processes, and commonalities and differences, and taking them out to the field in the next wave of data collection
4. Noting reflections or other remarks in jottings, analytic memos, and/or journals
5. Gradually elaborating a refined set of assertions, propositions, categories, themes, concepts, and generalizations that cover the consistencies discerned in the database
6. Comparing those generalizations with a formalized body of knowledge in the form of concepts or theories

The analytic challenge for all qualitative researchers is finding coherent descriptions and explanations that still include all of the gaps, inconsistencies, and contradictions inherent in personal and social life. The risk is in forcing the logic, the order, and the plausibility that constitute theory-making on the uneven, sometimes random, nature of social life.

We'll return to recurring features such as these, while acknowledging the diversity of analytic approaches now in use. Next, however, we need to take a step back to ask, what kind of data are we actually faced with in qualitative studies?

THE NATURE OF QUALITATIVE DATA

General Nature

In this book, we focus primarily on data in the form of words—that is, language in the form of extended text. Qualitative data also can appear as still or moving images, and we address these when necessary to the analytic task at hand (see the appendix for recommended titles and guidance for visual data).

The words we collect and analyze are based on *observations*, *interviews*, *documents*, *media*, and *artifacts*. These data collection activities typically are carried out in close proximity to a local setting for a sustained period of time. Such data are not usually immediately accessible for analysis but require some type of processing: Raw field notes need to be expanded and typed up, audio recordings need to be transcribed and corrected, and photographs need to be documented and analyzed.

The words we choose to document what we see and hear in the field can never truly be “objective”; they can only be our *interpretations* of what we experience. Similarly, transcription of audio recordings can be done in many ways that will produce rather different texts. And the influence of the researcher's personal *values*, *attitudes*, and *beliefs* from and toward fieldwork cannot be avoided.

To put it another way, qualitative data are not so much about behavior as they are about *actions* (which carry with them intentions and meanings and lead to outcomes or consequences). Some actions are relatively straightforward; others involve “impression management”—how people want others, including the researcher, to see them. Furthermore, those actions always occur in specific situations within a social and historical context, which deeply influence how they are interpreted by both insiders and the researcher as outsider. Thus, the apparent simplicity of qualitative data masks a good deal of complexity, requiring plenty of care, self-awareness, and reflection on the part of the researcher.

Strengths of Qualitative Data

One major feature of well-collected qualitative data is that they focus on *naturally occurring*, *ordinary events in natural settings*, so that we have a strong handle on what “real life” is like. Qualitative data,

with their emphasis on people's *lived experiences*, are fundamentally well suited to locating the *meanings* people place on the events, processes, and structures of their lives and for connecting these meanings to the *social world* around them. There is *local groundedness*—that is, data that were collected in close proximity to a specific situation. The emphasis is on a specific *case*, a focused and bounded phenomenon embedded in its context. The influences of the *local context* are not stripped away but are taken into account. The possibility for understanding latent, underlying, or nonobvious issues is strong.

Another feature of qualitative data is their *richness and holism*, with strong potential for revealing complexity. Such data provide “thick descriptions” (Geertz, 1973) that are vivid, are nested in a real context, and have a ring of truth that has a strong impact on the reader. Furthermore, the fact that such data are typically collected over a *sustained period* makes them powerful for studying any process (including history). We can go far beyond snapshots of “what?” or “how many?” to just *how* and *why* things happen as they do—and even *assess causation* as it actually plays out in a particular setting.

We make three other claims for the power of qualitative data, to which we will return in later chapters: as *a strategy for discovery*, for *exploring new ideas*, and for *developing hypotheses*. In addition, we underline their potential for *testing hypotheses*, seeing whether specific predictions hold up. Finally, qualitative data are useful when one needs to supplement, validate, or illuminate *quantitative* data gathered from the same setting for mixed methods studies.

OUR VIEW OF QUALITATIVE DATA ANALYSIS

We see analysis as three concurrent flows of activity: data condensation, data display, and conclusion drawing/verification. We will explore each of these components in more depth as we proceed through the book. For now, we make only some overall comments.

Data Condensation

Data condensation refers to the process of selecting, focusing, simplifying, abstracting, and/or transforming the data that appear in the full corpus (body) of written-up field notes, interview transcripts, documents, and other empirical materials. By condensing, we're making data *stronger*. (We stay away from data “reduction” as a term because that implies we're weakening or losing something in the process.)

As we see it, data condensation occurs continuously throughout the life of any qualitatively oriented project. Even before the data are actually collected, anticipatory data condensation is occurring as the researcher decides which conceptual framework, which research questions, which cases, and which data collection methods to choose. As data collection proceeds, further episodes of data condensation occur: writing summaries, coding, generating categories, developing themes, and writing analytic memos. The data condensing/transforming process continues after the fieldwork is over, until a final report is completed.

Data condensation is not something separate from analysis. It is a *part* of analysis. The researcher's decisions—which data units to code and which to pull out, which category labels best summarize a number of codes, which evolving story to tell—are *all analytic choices*. Data condensation is a form of analysis that sharpens, sorts, focuses, discards, and organizes data in such a way that “final” conclusions can be drawn and verified.

By data condensation, we do not necessarily mean quantification. Yet occasionally, it may be helpful to convert the data into *magnitudes* (e.g., the analyst decides that the program being looked at has a “high” or “low” degree of effectiveness), but this is not always necessary.

Data Display

The second major flow of analysis activity is data display. Generically, a *display* is an organized, condensed assembly of information that allows analytic reflection and action. In daily life, displays vary from smartphone apps to newspaper advertisements to Facebook status updates. Looking at displays helps us understand what is happening and to do something—either analyze further or take action—based on that understanding.

The most frequent form of display for qualitative data in the past has been *extended text*. As we will note later, text (in the form of, say, 1,000 pages of field notes) is terribly cumbersome. It is dispersed, sequential rather than simultaneous, poorly structured, and extremely bulky. Using only extended text, a researcher may find it easy to jump to hasty, partial, and unfounded conclusions. Human beings are not very powerful as processors of large amounts of information or big data, as it's currently termed. Extended text overloads our information-processing capabilities and preys on our tendencies to find simplifying patterns. Or we drastically overweight vivid information, such as the exciting event that jumps out of page 124 of the field notes after a long, “boring” passage.

In the course of our work, we have become convinced that good displays are a major avenue to robust qualitative analysis. The displays discussed and illustrated in this book include many types of matrices, networks, and graphics. All are designed to assemble organized information into an immediately accessible, compact form so that the analyst can see what is happening and either draw justified conclusions or move on to the next step of analysis that the display suggests may be useful.

As with data condensation, the creation and use of displays is not separate from analysis—it is a *part* of analysis. Designing a display—deciding on the rows and columns of a matrix for qualitative data and deciding which data, in which form, should be entered in the cells—are analytic activities. In this book, we advocate more systematic, powerful displays and urge a more inventive, self-conscious, and iterative stance toward their generation and use. As we've said in our previous writings, “You know what you display.”

Drawing and Verifying Conclusions

The third stream of analysis activity is conclusion drawing and verification. From the start of data collection, the qualitative analyst interprets what things mean by noting patterns, assertions, propositions, explanations, and causal flows. The competent researcher holds these conclusions lightly, maintaining openness and skepticism, but the conclusions are still there, vague at first, then increasingly explicit and grounded. “Final” conclusions may not appear until data collection is over, depending on the size of the corpus of field notes; the coding, storage, and retrieval methods used; the analytic prowess of the researcher; and any necessary deadlines to be met.

Conclusion drawing, in our view, is only half of the picture. Conclusions are also *verified* as the analyst proceeds. Verification may be as brief as a fleeting second thought crossing the analyst's mind during writing, with a short excursion back to the field notes, or it may be thorough and elaborate, with lengthy argumentation and review among colleagues to develop “intersubjective consensus,” or with extensive efforts to replicate a finding in another data set. The meanings emerging from the data have to be tested for their plausibility, sturdiness, and confirmability—that is, their validity. Otherwise, we are left with interesting stories about what happened but of unknown truth and utility.

Some researchers take issue with the term “conclusions” because the nature of contemporary social life, they believe, is too uncertain, tenuous, and fluid for any notion of finality. Other preferred terms such as “findings,” “discoveries,” and “closure” smooth the presumptive edge of authoritative certainty. Admittedly, we too subscribe to this open-ended perspective, yet we use “conclusions”

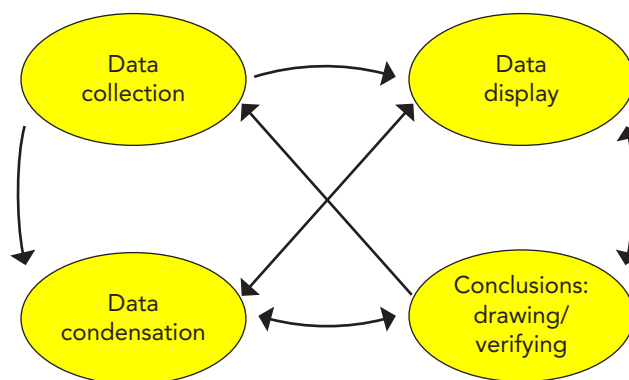
simply because it is a well-established term (and expectation) of research methodology. Also, a report that does not provide readers with *some* form of take-aways at the end serves little value to its audiences and suggests that the researcher has not fulfilled his or her analytic responsibilities.

We have presented these three streams—data condensation, data display, and conclusion drawing/verification—as interwoven before, during, and after data collection in parallel form, to make up the general activity called “analysis.” The three streams can also be represented as shown in Display 1.1—our first network display. In this view, the three types of analysis activity and the activity of data collection itself form an interactive, cyclical process. The researcher steadily moves among these nodes during data collection and then shuttles among condensing, displaying, and conclusion drawing/verifying for the remainder of the study.

The coding of data, for example (*data condensation*), leads to new ideas on what should go into a matrix (*data display*). Entering the data requires further data condensation. As the matrix fills up, preliminary *conclusions* are drawn, but they lead to the decision, for example, to add another column to the matrix to test and *verify* the conclusion.

In this view, qualitative data analysis is a continuous, iterative enterprise. Issues of data condensation, display, and conclusion drawing/verification come into play successively as analysis episodes follow each other. Such a process is actually no more complex, conceptually speaking, than the analysis modes quantitative researchers use. Like their qualitative colleagues, they must be preoccupied with data condensation (calculating means, standard deviations); display (correlation tables, regression printouts); and conclusion drawing/verification (significance levels, experimental/control group differences). But their activities are carried out through well-defined, familiar methods; are guided by canons; and are usually more sequential than iterative or cyclical. Qualitative researchers are in a more fluid and more humanistic position of meaning making.

DISPLAY 1.1 ■ Components of Qualitative Data Analysis: Interactive Model



Source: Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

SUGGESTIONS FOR READERS

Recommendations for what a reader should do with any particular book are often presumptuous, mistaken, or both. Authors have no control over who reads their books or what readers may find useful. Nevertheless, we offer a few suggestions for different types of users.

Students and Other Novice Researchers

We give some direct advice here, keeping in mind that you will often be working alone, usually on a single case, and may be feeling worried about the quality of your study—thesis/dissertation or not. We emphasize that this book is not written for newcomers to qualitative inquiry; readers should have some introductory background knowledge of fundamentals such as basic fieldwork research terminology, conducting a literature review, participant observation, and interviewing participants.

1. This book focuses on *data analysis*. Use other, introductory books to help with the basics of fieldwork (see the appendix for recommended titles and resources).
2. Learn by doing. Use your own study (whether it is in the planning stage or under way) as a vehicle and apply it to relevant methods in each chapter.
3. Compensate for the problem of having to work alone by finding someone to be a critical friend or mentor to respond to your work as you proceed.
4. Keep an informal journal or analytic memo set of what you are running up against. This tactic will help your learning and will be useful when you write up your study.
5. Don't worry about the jargon-like names of particular displays; the issue is what a display can do for you.
6. The biggest enemy of your learning is the gnawing worry that you're not "doing it right." Thesis and dissertation work tends to encourage that. But any given analytic problem can be approached in many useful ways. Creative heuristics—that is, inventing your way out of a problem—is definitely the better stance.

Experienced Researchers

This is a sourcebook. Colleagues have told us that they have used it in several ways:

1. *Browsing*: The book contains a wide range of material, so simply exploring it in an unstructured way can be fruitful.
2. *Problem solving*: Anyone opening the book comes to it with more or less specifically defined problems in doing qualitative data analysis. The index has been designed to be "problem sensitive" to permit easy access to appropriate sections of the book. The table of contents can also be used in this way.
3. *"A to Z"*: Some readers prefer to go through a book sequentially, from start to finish. We have organized the book so that it makes sense that way.
4. *Operational use*: For readers conducting an ongoing qualitative research project, either alone or with colleagues, it's useful to read particular sections focusing on upcoming analysis tasks (e.g., coding, time-ordered displays); then discuss them with available colleagues; and finally plan the next steps in the project, revising the methods outlined here or developing new ones.
5. *Research consulting*: The book can be used by people with an advisory or consulting role in the start-up and ongoing life of research projects. Assuming good problem identification, a research consultant can work with the client in either a problem-solving or a direct-training capacity to aid in thoughtful project design and coping with early problems.

Teachers of Qualitative Research Methods Courses

Some colleagues have used this book as a primary text, others as a supplementary one. In either case, our advice is to engage students in active data collection and analysis. For each topic, we have used a learning approach like this, carried out by individuals or working pairs, who stay together throughout an extended workshop:

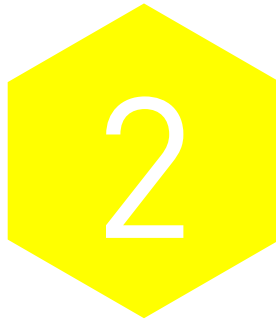
1. Introductory lecture and/or reading to clarify the main conceptual points of the section
2. A brief learning task (e.g., coding a data excerpt, designing a matrix template, drawing a network or graphic, interpreting a filled-out matrix, or writing an initial analysis)
3. Comparing the products of individuals or pairs, drawing generalizations, and discussing future applications of the method

The same general principles apply when the book is used in a semester-long course, although the coverage will be deeper. Interim exercises focusing on actual research tasks, critiqued in class, are particularly productive. Active, reflective self-documentation through personal journals or analytic memos is also beneficial.

CLOSURE AND TRANSITION

This introduction provided some brief groundwork and foundations for the rest of the book. Analysis is doing, so let's proceed to the next chapter of Part One to review research design decisions that will later play important roles in qualitative data analysis. Following that chapter is a brief discussion on ethics, then an expanded chapter on analysis fundamentals.

Part Two's chapters are the heart of this book—methods of displaying data for exploring, describing, ordering, explaining, and predicting about the social worlds we study. Part Three's chapters examine how we verify and document the conclusions we draw through systematic thinking and good writing.



RESEARCH DESIGN AND DATA MANAGEMENT

CHAPTER SUMMARY

This chapter discusses how qualitative research design decisions influence and affect later data analysis. Such core topics include conceptual frameworks, methodologies, research questions, instrumentation, sampling, consideration of mixed methods, and data management with computers and software.

CONTENTS

Introduction	General Sampling Strategies
Loose Versus Tight Research Designs	Within-Case Sampling
Displaying the Conceptual Framework	Multiple-Case Sampling
Description and Rationale	Sampling Advice
Example	Instrumentation
Other Examples	Description and Rationale
Conceptual Framework Advice	Example
Methodologies (Genres) of Qualitative Research	Instrumentation Advice
Formulating Research Questions	Linking Qualitative and Quantitative Data
Description and Rationale	Approaches to Mixed Methods Designs
Example	Examples
Research Question Advice	Mixed Methods Advice
Defining the Case	Data Management Issues Bearing on Analysis
Description and Rationale	Data Management
Examples	Computer and Software Use
Case Advice	CAQDAS
Sampling: Bounding the Collection of Data	Staffing and Time Planning
Description and Rationale	Closure and Transition
Key Features of Qualitative Sampling	

INTRODUCTION

This book is about qualitative data analysis, but first we need to discuss how initial research design decisions may influence and affect the forthcoming analytic components of the study. Research design choices such as the conceptual framework, research methodology, research questions, data collection methods, participant sampling, and so on are analytic acts themselves. They are a sort of *anticipatory data condensation* because they rule out certain factors and relationships and attend to others. Even technical management tasks such as how data will be stored, organized, and processed, and what computer software may be used, both support and affect analysis.

We cannot deal thoroughly here with qualitative research design; see the appendix for recommended titles on the topic. In this chapter, we discuss the *analytic issues* that arise as a study is conceived and implemented. We provide specific examples but want to emphasize that these issues must be dealt with uniquely and flexibly in any particular study. Initial design decisions nearly always lead to a redesign.

LOOSE VERSUS TIGHT RESEARCH DESIGNS

Prior to fieldwork, how much shape and structure should a qualitative research design have? Some researchers keep prestructured designs to a minimum. They consider social processes too complex, too contextual, too elusive, or too fluid to be approached with explicit conceptual frameworks or standardized data collection instruments. They prefer a more loosely structured, emergent, inductively grounded approach for gathering data. Their conceptual frameworks will tend to emerge from the field during the course of the study. The important research questions will become clear only gradually; meaningful settings and participants will not be selected prior to fieldwork but only after initial orientation to the site.

Highly inductive, loosely designed studies make good sense when experienced researchers have plenty of time to explore unfamiliar cultures, understudied phenomena, or very complex social processes. But if you're new to qualitative research and are looking at a better understood phenomenon within a familiar culture or subculture, a loose, inductive design may be a waste of time. Months of fieldwork and voluminous data may yield only a few analytic banalities. Also, fieldwork may well involve multiple-case research rather than single-case studies. If different fieldworkers on a research team are operating inductively with no common framework or instrumentation, they are bound to end up with the double dilemma of data overload and lack of comparability across cases.

Tighter designs provide clarity and focus for beginning researchers worried about protocols and data overload. Tighter designs are also a wiser course for researchers working with well-delineated research questions or who wish to test or further explicate hypotheses. A standardized protocol for data collection across multiple sites gathered by multiple fieldworkers permits better cross-case comparison and analysis.

Yet in multiple-case research, the looser the initial framework, the more each researcher can be receptive to local idiosyncrasies—but cross-case comparability will be hard to get, and the costs and the information load for analysis will be colossal. Tightly coordinated designs face the opposite dilemma: They yield more economical, comparable, and potentially generalizable findings, but they are less case sensitive and may entail bending data out of contextual shape to answer a cross-case analytic question.

Then, too, we should not forget why we are out in the field in the first place: to describe and analyze a pattern of interrelationships. Starting with them (deductively) or getting gradually to them

(inductively) are both possible. In the life of a study, we need both approaches to pull a mass of facts and findings into a wide-ranging, coherent set of conclusions and generalizations.

A case can be made for tight, prestructured qualitative designs and for loose, emergent ones. Whether you choose one end of the continuum or the other for your particular study must be your decision. But a midway point between the extremes is yet another position, and perhaps the one most qualitative researchers take.

DISPLAYING THE CONCEPTUAL FRAMEWORK

Description and Rationale

A conceptual framework explains, graphically and/or in narrative form, the main things to be studied—for example, the key factors, variables, phenomena, concepts, participants—and the presumed interrelationships among them—as a *network*. Conceptual frameworks are simply the current and evolving version of the researcher’s “map” of the qualitative territory being investigated. As the explorer’s knowledge of the terrain improves, the map becomes correspondingly more differentiated and integrated. Thus, conceptual frameworks are developed at the beginning of a study and evolve as the study progresses. This framework also becomes a template, of sorts, for how data collection and analysis begin and proceed during and after fieldwork.

A conceptual framework forces you to be selective—to decide which things are most important; which relationships are likely to be most meaningful; and, as a consequence, what information should be collected and analyzed—at least at the outset. If multiple researchers are involved, the framework helps them study the same phenomena in ways that will permit an eventual cross-case analysis.

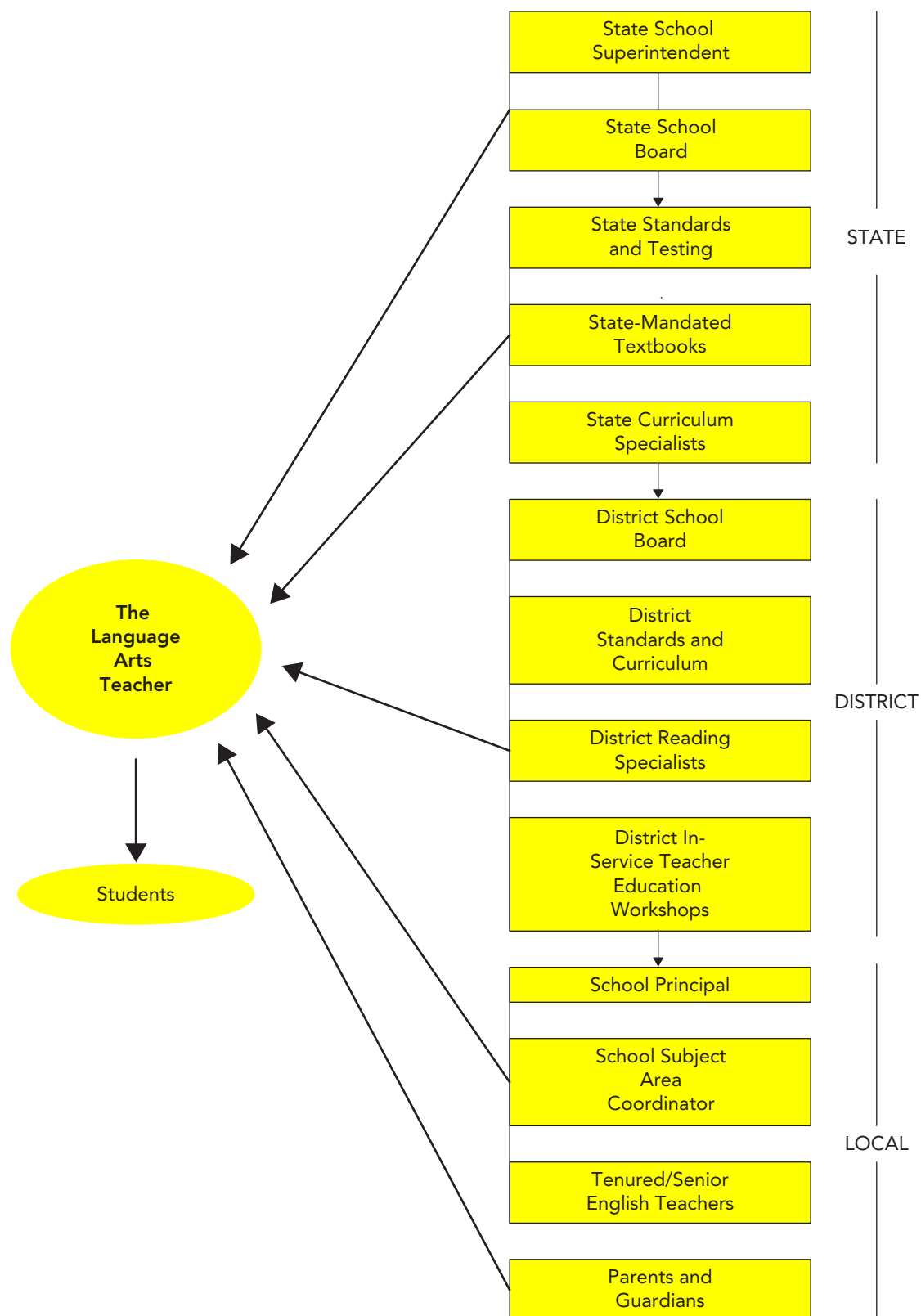
Conceptual frameworks can be simple or elaborate, descriptive or causal, commonsensical or theory driven. Theory relies on a few general concepts that subsume a mountain of particulars. *Culture*, *social intelligence*, *coping*, and *identity* are the labels we put on intellectual “bins” containing constituent actions, experiences, variables, categories, processes, and events. Bins come from theory, from personal experience, and often from the general objectives of the study envisioned. Setting out bins, naming them, and getting clearer about their interrelationships help lead you toward a conceptual framework. It is both the process and product of analytic reflection.

A note on terminology: Some use the terms *conceptual framework* and *theoretical framework* interchangeably, but some methodologists note that there are distinct differences between the two, depending on whose definitions you read. To us, a theoretical framework utilizes theory/theories and their constituent elements as the presumed “working model” that drives the investigation and analysis of a social phenomenon. But a conceptual framework is a more inductively derived and evolutionary model that can certainly include aspects of the theoretical, but primarily incorporates case- or site-specific variables, concepts, contexts, participants, and so on. In other words, a conceptual framework grounds itself in the local elements of a particular, unique study; a theoretical framework abstracts a study’s ideas based on the literature.

Example

A conceptual framework can specify who and what will (and will not) be studied. Display 2.1 is a *first-cut, first-draft* attempt at listing, in graphic form, the myriad influences on a single classroom teacher. This illustration identifies the people (state school superintendent, district reading specialists, etc.); organizations (state school board, district school board, etc.); and “things” or

DISPLAY 2.1 A First-Draft Conceptual Framework for a Case Study Teacher and the Influences on Her Practice



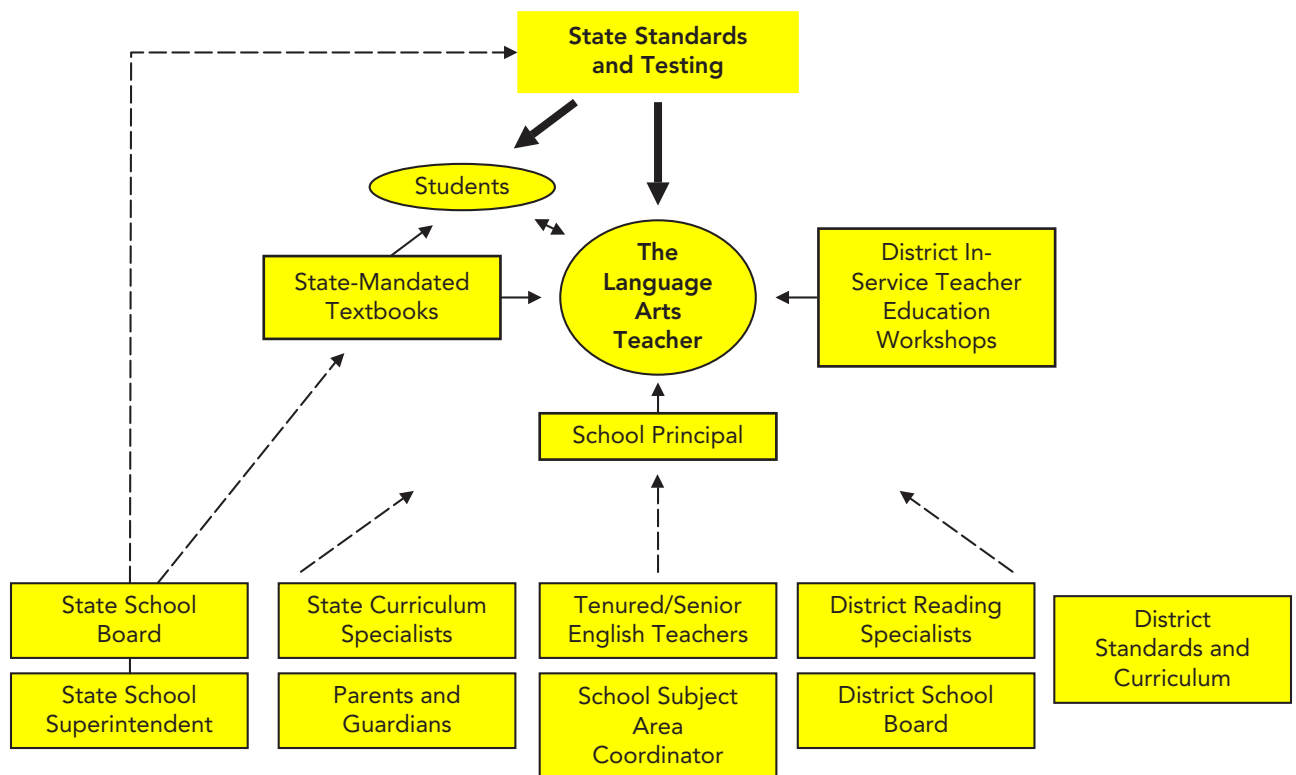
official documents/policies (state-mandated textbooks, district standards and curriculum, etc.) the researcher identified as influential on a classroom teacher's instructional practices with her students. There is a general clustering of *whos* and *whats* by level—state, district, and local—and an implied hierarchy of supervision/authority and distribution of power from top to bottom.

We see here the focusing function of a conceptual framework. Some, not all, social actors in this long list are going to be studied, along with some, not all, aspects of their activity. For example, the researcher, due to accessibility, may be unable to directly interview the state school superintendent or state school board members. But he can certainly access public documents and records of their official business from published minutes of state school board meetings. Once all permissions have been obtained, the language arts teacher herself will be interviewed several times and observed teaching in several of her classrooms. Documents such as her lesson plans, textbooks, and samples of graded student work will be reviewed. Only some relationships in this conceptual framework will be explored, certain kinds of processes documented, and certain analyses made—at least at the outset.

Now for a slightly more complex (i.e., “messy”) conceptual framework using some of the same aspects of this study as they evolved. Display 2.2 includes the exact same bins and labels as Display 2.1, but their arrangement and the arrows of influence are different. There is less symmetry (support that the network reflects the “asymmetrical” social world we live in), a new hierarchy, and a different set of interrelationships.

The researcher got to this display after spending some time observing and interviewing the language arts teacher. State standards and testing were the primary factors that influenced both her practice

DISPLAY 2.2 ● Major Influences on a Language Arts Teacher's Practice



and the students' classroom experiences. Of all the players from Display 2.1 that were hierarchically listed, some were actually quite peripheral *according to the perspective of the language arts teacher*. These are listed at the bottom of the display, with dashed arrows toward the major bins suggesting minimal or foregrounded influence.

The other three major factors the researcher learned and interpreted as most salient were the teacher's school principal and her "obsession" with raising the school's standardized test scores, the district in-service teacher education workshops (which the teacher praised and found "extremely useful for me and my kids"), and the state-mandated textbooks (which heavily emphasized writing over literature appreciation for language arts).

A display such as this might be considered partly *confirmatory* if the researcher is able to visit other language arts teachers within the same school and within other schools outside the district to determine if their spheres of influence also include the same factors. A *working hypothesis* for field testing that the researcher developed after studying and analyzing this display is this: *The mandates of education manifest themselves primarily as prescriptive products imposed on educators and their students*. Overall, what Displays 2.1 and 2.2 illustrate is that conceptual frameworks *evolve* as a study continues and the bigger picture becomes clearer. Display 2.2 is the product of fieldwork, initial data analysis, and a processual model for testing as additional data and analysis continue.

The conceptual framework is your first analytic *display*. It is a visual representation of your main conceptual ideas about a study and how they interact and interplay with each other. Also remember that conceptual frameworks tell a *story* of some kind. They are evolving, one-page representations of your research in progress, comparable to the storyboards of filmmakers, who first draw on paper or render with software what they will eventually document on digital video.

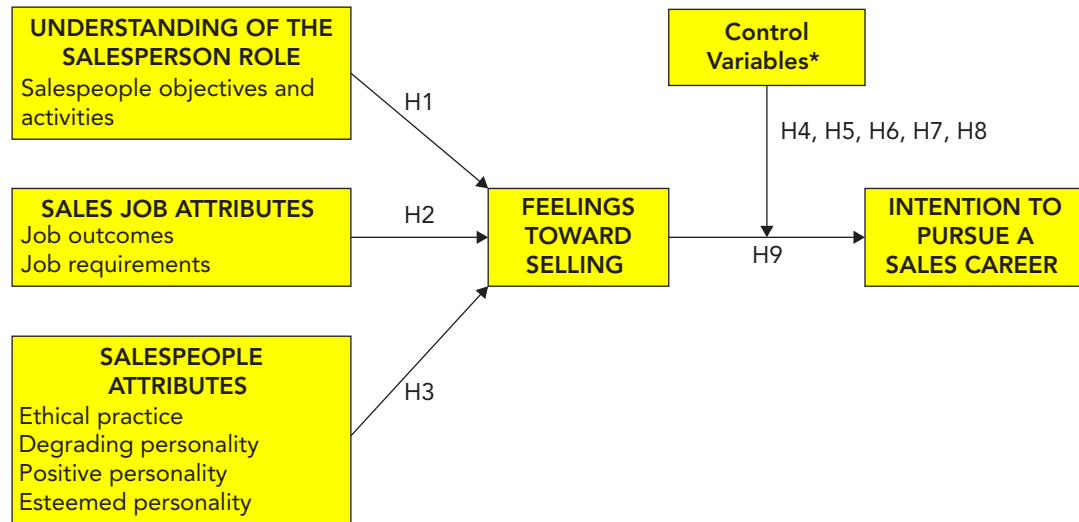
Other Examples

We now provide two examples from published studies that illustrate the display possibilities for conceptual frameworks. Space does not permit us to discuss each one in depth, but we hope you can get the gist of the researchers' approaches by examining the bins and their proposed interaction patterns and trajectories.

Display 2.3, from Ballestra, Cardinali, Palanga, and Pacelli (2017), portrays a conceptual model of teenage students' intentions to pursue a sales career. Notice that the overall design follows a linear trajectory. But also notice that each arrow has one or more hypotheses (H1, H2, H3, etc.) attached to it. Hypothesis 1 in this study is as follows: "Understanding the changes to the salesperson's role will have a significant impact on students' feelings toward selling as a career" (p. 178). Hypothesis 8 is the following: "Studying business/economic subjects during high school has a significant impact on students' later intent to pursue a sales career" (p. 179). This conceptual framework includes nine specific hypotheses, developed from the coresearchers' literature review, to test as qualitative and quantitative data are collected and analyzed. Not only is their display a preliminary or *exploratory* model of a young adult's sales career process, it is also an analytic game plan to examine whether their participants' experiences support the theories of the field.

Display 2.4, from Chapman et al. (2017), examines the role of powerlessness among health care workers in tuberculosis infection control. This model, however, is not representative of the beginning of the study but rather the *end* of it, after the data have been collected and analyzed:

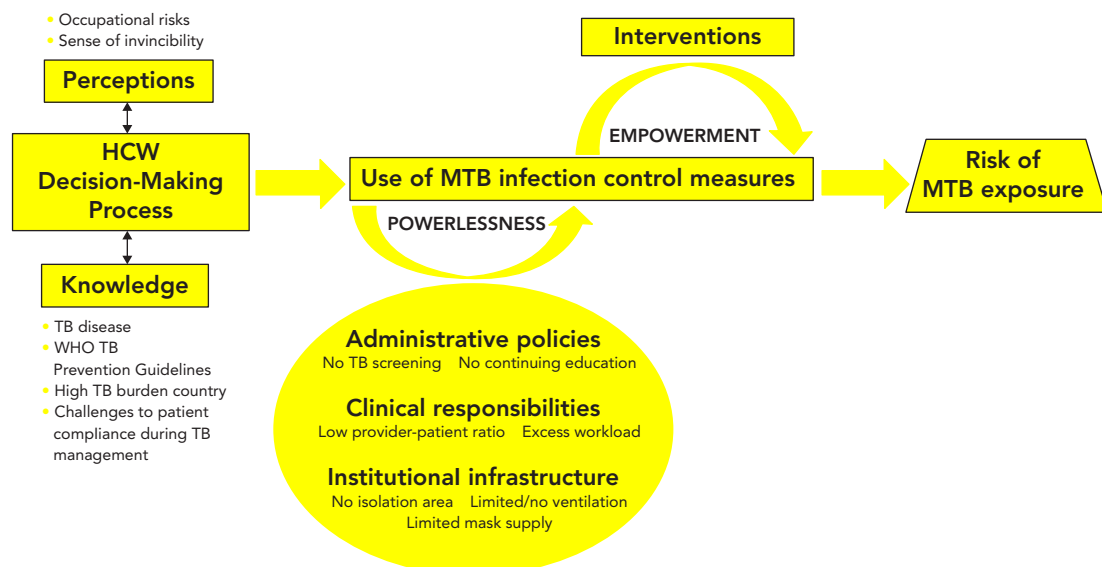
The first cycle describes the context of powerlessness, while the second cycle represents how empowerment through evidence-based interventions can be applied in the future.

DISPLAY 2.3 ● A Conceptual Model of Students' Intention to Pursue a Sales Career

Source: Ballestra, Cardinali, Palanga, & Pacelli (2017), p. 181.

[Display 2.4] presents the HCWs' [Health Care Workers'] perceived limitations in the application of *M. tuberculosis* infection control measures in clinical practice, influencing the decision-making process and resulting in the knowledge-action gap. (Chapman et al., 2017, p. 2119)

Like Display 2.3, this too is a linear model, but notice how the bins in Display 2.4 are supported with bullet-pointed factors that detail the analytic findings. This conceptual model elegantly maps the researchers' conclusions in story form with an accompanying narrative that explains the participants' dilemmas and experiences.

DISPLAY 2.4 ● Conceptual Model That Presents HCWs' Perceived Limitations in the Application of MTB Infection Control Measures in Clinical Practice, Influencing the Decision-Making Process and Resulting in the "Knowledge-Action" Gap

Source: Chapman et al. (2017), p. 2120.

Conceptual Framework Advice

Here are some suggestions that summarize and extend what has been reviewed in this section:

1. Conceptual frameworks are best done graphically, at first, to support the accompanying narrative. Having to get the entire framework on a single page obliges you to specify the bins that hold discrete phenomena, map likely interrelationships, divide variables that are conceptually or functionally distinct, and work with all of the information at once.
2. Expect to do several versions right from the outset. There are probably as many ways of representing the main variables as there are variables to represent, but some—typically later cuts—are more elegant than others.
3. If your study has more than one researcher, have each field researcher do a cut at a framework early on and then compare the several versions. This procedure will show, literally, where everyone's head is. It usually leads to an explication of contentious or foggy areas that otherwise would have surfaced later on.
4. Avoid the no-risk framework—that is, one that defines variables at a very global level and has two-directional arrows everywhere. This avoidance amounts essentially to making no focusing decisions and is little better than the strategy of going indiscriminately into the field to see what the site has to tell. However, you can begin with a generic framework as a way of getting to a more selective and specific one.
5. Prior theorizing and empirical research are, of course, important inputs. It helps to lay out your own orienting frame and then map onto it the variables and relationships from the literature available, to see where the overlaps, contradictions, qualifications, and refinements are.

For an extended discussion of conceptual frameworks and how they influence and affect all aspects of research, see Ravitch and Riggan (2016).

METHODOLOGIES (GENRES) OF QUALITATIVE RESEARCH

Qualitative research can be conducted in dozens of ways, many approaches with long traditions behind them. Saldaña (2011b) describes more than 20 different qualitative research genres (i.e., methodologies or types) out of many more available to investigators, ranging from well-established traditions such as ethnography, grounded theory, phenomenology, case study, and content analysis, to more progressive genres of qualitative research such as poetic inquiry, narrative inquiry, ethno-drama, and autoethnography. To do them all justice is impossible here. For our purposes, the questions are as follows: What do selected genres of qualitative research have to say about *analysis*? And can we see some common themes and practices among them?

Ideally, the conceptual framework you develop for a study is the impetus for selecting the most appropriate methodological approach to investigate its major components. For example, if you conceptualize that depression is a cognitive and emotional *experience*, then a *phenomenological study* with its emphasis on *thematic analysis* may be an appropriate choice. If you conceptualize that depression is an intimate, personal, and contextual state of being, then perhaps an exemplary *case study* of an individual living with depression, holistically analyzed through *narrative vignettes*, may be in order.

Yet sometimes the methodology we select is based not on our conceptual framework but rather on our *research questions* of interest (discussed next). If we wish to get the best answers to our particular set of inquiries, then we should select the best research genre *and its accompanying methods* that will produce them. If we wish to know the details of *how* people cope with depression in their daily lives, then a series of *interviews* with participants may be more revealing than participant observation of their day-to-day routines. Interview transcripts would then be carefully *coded* to develop major *categories* of coping with mental illness (e.g., psychotropic medication, professional therapy, self-care strategies).

Some researchers, however, have a particular affinity for just one genre of research and dedicate their professional careers to studying social life through a particular framework. They may label themselves “grounded theorists,” “phenomenologists,” or “autoethnographers,” and thus investigate the world through a specialized methodological lens and design their projects to accommodate their preferred ways of working. Over time, they also develop nuanced expertise in the genre itself and have much to offer other researchers about a particular methodology and its methods.

The point of this discussion is that the particular research methodology you select for your study will determine what types of data collection methods you employ and thus what types of analysis might be undertaken. Grounded theory, for example, uses a series of cumulative coding cycles and reflective analytic memoing to develop a core category for theory generation. Phenomenology tends to look at data thematically to extract essences and essentials of participant meanings. Mixed methods research integrates both qualitative and quantitative data and analyses for a more multidimensional approach to inquiry. Poetic inquiry, narrative inquiry, and ethnodrama adopt and adapt the conventions of fictional literature to render nonfictional participant experiences in poetic, prosaic, and dramatic forms, as opposed to the traditional and conventional formats of scholarly/academic writing.

The purpose of this section is not to describe every single methodology of qualitative research available to you but to focus on some common features that occur in most genres of qualitative inquiry. We list some of them here, aware that some exemplars are missing:

1. Qualitative research is conducted through intense and/or prolonged contact with participants in a naturalistic setting to investigate the everyday and/or exceptional lives of individuals, groups, organizations, cultures, and/or societies.
2. The researcher’s role is to gain a holistic (systemic, encompassing, and integrated) overview of the context under study: its social arrangement, its ways of working, and its explicit and implicit rules.
3. The researcher attempts to capture data on the perceptions of local participants from the inside through a process of deep attentiveness, empathetic understanding, and suspension or bracketing of preconceptions about the topics under discussion.
4. Relatively little standardized instrumentation is used, though audio-recorded interviews with participants tend to be used most often. The researcher himself or herself is essentially the main instrument in the study.
5. Most of the analysis is done with words. The words can be assembled, subclustered, or broken into segments. They can be reorganized to permit the researcher to compare, contrast, analyze, and construct patterns out of them for analytic outcomes such as extended narratives, categories, themes, assertions, propositions, and/or theories.
6. The main task is to describe and explain the ways people in particular settings come to understand, account for, take action, and otherwise manage their day-to-day situations.

These features may be more relevant for naturalistic, ethnographic studies, but they are configured and used with slight variation in many social research traditions.

FORMULATING RESEARCH QUESTIONS

Description and Rationale

Research questions represent the facets of inquiry that the researcher most wants to explore. Research questions may be general or particular, descriptive or explanatory. The formulation of research questions may precede, follow, or happen concurrently with the development of a conceptual framework and/or the chosen methodological approach. They also may be formulated at the outset or later on and may be revised or reformulated during the course of fieldwork.

It is a direct step from a conceptual framework to research questions. If I have a bin labeled “State-Mandated Textbooks,” as in Display 2.1, with an arrow from that bin directed toward “The Language Arts Teacher,” I am implicitly asking myself some questions about how textbooks influence a teacher’s practice (e.g., *In what ways do state-mandated language arts textbooks shape the language arts teacher’s curriculum?*). If I have a two-way arrow between “The Language Arts Teacher” and “Students,” as in Display 2.2, my question has to do with the reciprocal interrelationship between them and the interpersonal dynamics of education (e.g., *What kinds of teaching-learning methods best prepare students for state-mandated testing in language arts?*).

If my conceptual framework is more constrained, so are my questions. In Display 2.2, the State School Board has little, if any, direct *influence and affect* (our qualitative preference over quantitative research’s “cause and effect”) on the teacher—even though it was the body that mandated state standards and testing. To the teacher, the *tests* are perceived as the primary influence on her practice, not the personnel or administrative overview of the State School Board. There may be one research question about the board to verify its inconsequential impact, but not much time or effort will be spent in pursuing this minimal factor.

What do these questions do for me? They tell me what I want to know most or first; my collection of data will be more focused. I am also beginning to make some implicit sampling decisions. I will look only at *some* participants in *some* contexts dealing with *some* issues. The questions also begin to point me toward data-gathering methods—observations, interviews, and document collection. Finally, the research questions begin to operationalize the conceptual framework and make the initial theoretical assumptions even more explicit.

A conceptual framework’s display shows researchers’ preferred bins and relational arrows as they map and carve up social phenomena. They use these explicitly or implicitly to decide which questions are most important and how they should get the answers. We believe that better research happens when you make your framework (and associated choices of research questions, cases, sampling, instrumentation, etc.) explicit, rather than claiming inductive purity.

Example

Our (Miles and Huberman’s) school improvement study investigated the dissemination of educational innovations carried out from 1979 to 1983. (Many display examples in this text come from that classic study. More detailed information appears in the first, second, and third editions of this book.) The study was nested in a larger study of school improvement, covering 145 school buildings and nearly 4,000 people throughout the United States involved in the implementation of educational innovations. Joined by two colleagues, Beverly Loy Taylor and Jo Anne Goldberg, we

repeatedly visited a stratified sample of 12 field sites across the country throughout the 1979–1980 school year, with follow-up contacts the next year to verify the main findings.

The volume of data collected included 440 interviews, 85 observations, some 259 documents, and 2,713 pages of field notes. We developed a common set of data displays and, for each of the 12 field sites, used them to draw conclusions, resulting in 12 case reports ranging from 70 to 140 pages each, with a common format. Our subsequent cross-case analysis was built from the appropriate sections of the 12 case reports.

A study of this scope necessitated an extensive series of research questions, and we present just one subset of them related to a school's decision to adopt an educational innovation. The procedure we used was to cluster specific research questions under more general ones, as shown in Display 2.5.

DISPLAY 2.5 ■ General and Special Research Questions Relating to the Adoption Decision (School Improvement Study)

How was the adoption decision made?

Who was involved (e.g., principal, users, central office people, school board, outside agencies)?

How was the decision made (top-down, persuasive, consultative, collegial-participative, or delegated styles)?

How much priority and centrality did the new program have at the time of the adoption decision?

How much support and commitment was there from administrators?

How important was it for teachers, seen in relation to their routine, "ordinary" activities, and any other innovations that were being contemplated or attempted?

Realistically, how large did it loom in the scheme of things?

Was it a one-time event or one of a series?

What were the components of the original plan for implementation?

These might have included front-end training, monitoring and debugging/troubleshooting unexpected problems, and ongoing support.

How precise and elaborate was this plan?

Were people satisfied with it at the time?

Did it deal with all of the problems anticipated?

Were the requisite conditions for implementation ensured before it began?

These might have included commitment, understanding, materials and equipment, skills, time allocation, and organizational backup.

Were any important conditions seen as missing? Which were most missing?

Source: Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Notice the choices being made within each topic area. For example, in the first two areas, the main things we want to know about the decision to adopt are who was involved, how the decision was actually made, and how important this project was relative to others. All of the questions seem to be functional, rather than theoretical or descriptive—they have to do with getting something done. When such a research question gets operationalized, an attempt will be made to determine whether these conditions were present or absent at the various field sites and whether that made any

difference in the execution of the project. This is an example of how research questions feed directly into data collection and thus analysis.

Research Question Advice

1. Even if you are in a highly inductive mode, it is a good idea to start with some general research questions. They allow you to get clear about what, in the general domain, is of most interest. They make the implicit explicit, without necessarily freezing or limiting your vision.
2. Formulating more than a dozen or so overarching research questions is looking for trouble. You can easily lose the forest for the trees and fragment the collection of data. Having a large number of questions makes it harder to see emergent links across different parts of the database and to integrate findings. A solution to research question proliferation for a small-scale study is the use of one major question with no more than five related subquestions for clarity and specificity. Studies with larger magnitude (e.g., multiple-case or longitudinal) can generate more research questions.
3. It is sometimes easier to develop a conceptual framework *after* you've made a list of research questions. You look at the list for common themes, common concepts, implicit or explicit relationships, and so on, and then begin to map out the underlying framework joining these pieces. Some researchers operate best in this mode.
4. Once the list of research questions is generated and honed, look it over to ensure that each question is, in fact, researchable. Delete or revise those questions that you or your participants have no real means of answering, or you of measuring (qualitatively or quantitatively).
5. In a multiple-case study, be sure all fieldworkers understand each question and see its importance. Multiple-case studies have to be more explicit, so that several researchers can be aligned as they collect information in the field. Unclear questions or different understandings can make for incomparable data across cases.
6. Keep the research questions in hand and review them during fieldwork. This closeness will focus data collection and analysis. You will think twice before noting down what participants have for lunch or where they park their cars. Unless something has an obvious, direct, or potentially important link to a research question, it should not be entered into your field notes.

DEFINING THE CASE

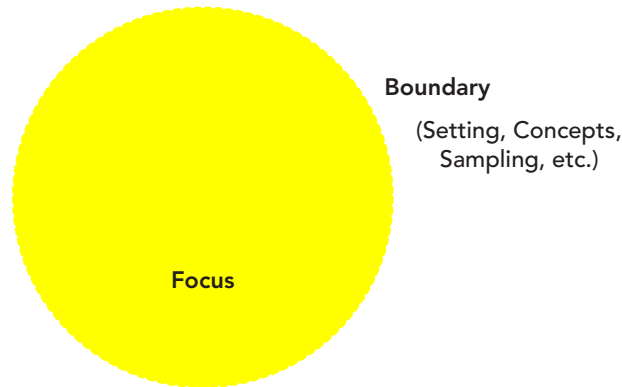
Description and Rationale

Qualitative researchers often struggle with questions of “what my case is” and “where my case leaves off.” Abstractly, we can define a *case* as a phenomenon of some sort occurring in a bounded context. The case is, in effect, your unit of analysis. Studies may be of just one case or of several. Display 2.6 shows this graphically: There is a focus or “heart” of the study, and a somewhat indeterminate boundary defines the edge of the case: what will not be studied.

Examples

What are some examples of cases? Sometimes the “phenomenon” may be an *individual* in a defined context, as suggested by Displays 2.1 and 2.2: a language arts teacher and her series of classes with junior-level high school students during an 18-week spring semester—the same semester her students will take a state-mandated standardized “high-stakes” test in language arts. Note that the

DISPLAY 2.6 The Case as the Unit of Analysis



Source: Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

“heart” here is the teacher. The boundary defines her students and school site as the major contexts. The researcher will not, for example, interview the teacher’s mother or visit the child care facility where the teacher leaves her own child during workdays.

The bounding is also by *time*: No information will be gathered after the spring semester ends in 18 weeks and the standardized test scores have been reported. We can also expect that the boundary will be defined further by *sampling*, which we’ll discuss later. For example, this researcher will not be interviewing the school guidance counselor, only the principal and, if pertinent, other language arts teachers.

A case may also be defined by

1. a *role* (school principal, CEO, nurse supervisor);
2. a *small group* (African American men in an inner-city neighborhood, a college-level rock band, a breast cancer survivor support group);
3. an *organization* (a nursery school, a software engineering company, the American Sociological Association);
4. *space and environment* (a mall for adolescents to “hang out,” visitors at the Vietnam Veterans Memorial Mall in Washington DC, nighttime foot traffic along the Las Vegas strip);
5. a *community or settlement* (the French Quarter of New Orleans, a village in Tanzania, the Tenderloin District of San Francisco);
6. *episodes or encounters* (voting for the first time, a “one-night stand,” bullying incidents on an elementary school playground);
7. an *event* (a search committee meeting, a high school graduation ceremony, New Year’s Eve in New York City’s Times Square);
8. a *period of time* (a day in the life of a firefighter, spring break, how customers use their time between ordering food from a server and having it delivered to their table);
9. a *process* (grocery shopping and meal preparation, organizing and managing an international conference, the adoption and implementation of an innovational education program in a school district);

10. a *culture or subculture* (African American women in academia, Los Angeles drag queens, “skater dudes”); or
11. a *nation* (Greece during the period of its 21st-century economic crisis, America during the 2018 midterm election cycle).

Single cases are the stuff of much qualitative research and can be very vivid and illuminating, especially if they are chosen to be “critical,” “extreme,” or “revelatory,” as Yin (2018, p. 24) suggests. But the cases may not be monolithic; cases may have subcases embedded within them. A case study of a school will contain cases of specific classrooms; a case study of a hospital ward may have cases of specific medical personnel–patient relationships within it.

We suggest that multiple cases offer the researcher an even deeper understanding of the processes and outcomes of cases, the chance to test (not just develop) hypotheses, and a good picture of locally grounded causation. The question of just which cases to include in a sample is discussed below.

A comment on notation: We sometimes prefer—and use here and there in this book—the word *site* because it reminds us that a “case” always occurs in a specified social and physical *setting*; we cannot study individual cases devoid of their site-specific context in the way a quantitative researcher often does.

Case Advice

1. Define the case as early as you can during a study. Given a starting conceptual framework and research questions, it pays to get a bit stern about who and what you are defining as a case; that will help clarify further both the framework and the questions.
2. Start intuitively but remember the focus and build outward. Think of whom and what you will *not* be studying as a way to firm up the boundaries.
3. Attend to several dimensions of the case: its *conceptual* nature, its *social size*, its *physical* location, and its *temporal* extent.
4. Remember that sampling will define the case(s) further.

SAMPLING: BOUNDING THE COLLECTION OF DATA

Description and Rationale

Sampling involves decisions not only about which people to observe and/or interview but also about settings, events, and social processes. Qualitative studies call for continual refocusing and redrawing of study parameters during fieldwork, but some initial selection still is required. A conceptual framework and research questions can help set the foci and boundaries for sampling decisions.

Sampling may look easy, but settings have subsettings (schools have classrooms, classrooms have cliques, cliques have individuals), so deciding where to look is not easy. Within any case, social phenomena proliferate (science lessons, teachers’ questioning techniques, student unruliness, use of innovations); they, too, must be sampled. But as much as you might want to, you cannot study everyone everywhere doing everything. Your choices—whom to look at or talk with, where, when, about what, and why—place limits on the conclusions you can draw and on how confident you and others feel about them. Sampling is crucial for later analysis, so how do we manage it all? We discuss some general principles and suggest useful references for detailed help.

Key Features of Qualitative Sampling

Qualitative researchers usually work with small samples of people, nested in their context and studied in-depth—unlike quantitative researchers, who aim for larger numbers of context-stripped cases and seek statistical significance.

Sampling in qualitative research involves two actions that sometimes pull in different directions. First, you need to set *boundaries*: to define aspects of your case(s) that you can study within the limits of your time and budget, that connect directly to your research questions, and that probably will include examples of what you want to study. Second, at the same time, you need to create a conceptual *frame* to help you uncover, confirm, or qualify the basic processes or concepts that undergird your study. Display 2.2 suggests that the frame of this study is primarily about the pressures and consequences of state-mandated testing on one language arts teacher and her students.

Qualitative samples tend to be *purposive* rather than random. Samples in qualitative studies are usually not wholly prespecified but can evolve once fieldwork begins. The initial choices of participants lead you to similar and different ones, observing one class of events invites comparison with another, and understanding one key relationship in the setting reveals facets to be studied in others. This is conceptually driven *sequential* sampling.

Qualitative sampling is sometimes *theory driven*, either “up front” or progressively, as in a grounded theory mode. Suppose that you were studying how adolescents develop friendships and that you could only manage to look at one high school. At first, that seems very limited. But if you chose a site according to relevant theory, you might choose one that has a wide range and diversity of students in terms of gender, race/ethnicity, socioeconomic class, sexual orientation, religious background, and so on. This will enable you to test the theory that friendships become more selective—that is, discriminating—when multiple choices are available to adolescents.

You would sample *within* each class (freshman, sophomore, junior, senior) for certain developmentally expected processes such as clique formation, orientation of the newcomer, use of technology for friendship maintenance, and so on. You might also find that certain events, such as lunchtime and organized extracurricular sports and arts activities, are unusually rich with socialization actions, and then you would sample more carefully for these. Sampling both within and across cases puts flesh on the bones of general concepts and their relationships. We can see generic processes; our generalizations are not to “all adolescents” but to existing or new theories of how friendship development works.

General Sampling Strategies

Erickson (1986) suggests a generic, funneling sampling sequence, working from the outside in to the core of a setting. For example, in studying schools, he would begin with the school community (census data, a walk around the neighborhood) and then enter the school and the classroom, staying several days to get a sense of the frequency and occurrence of different events. From there, the focus would tighten: specific events, times, and locations. Periodically, however, Erickson would follow lines of influence into the surrounding environment to test the typicality of what was found in a given classroom and to get a better fix on external influences and determinants.

There is a wide range of sampling strategies available to qualitative researchers within a complex case or across cases (Patton, 2008, 2015). They can be selected ahead of time or can evolve during early data collection. It is impossible to prescribe which sampling strategies go best with each type of study, for there are too many unique conditions within each project (specific research questions, specific sites and cases, etc.). But you should be able to provide to your readers justification for why you selected certain types of sampling over others.

Random sampling is a gold standard of quantitative research but is used quite minimally in qualitative research because random sampling can sometimes deal you a nebulous hand. Our sampling tends to be more *strategic* and *purposive* because we are focusing on a case's unique contexts. Admittedly, there are times when we select a case to study because it is accessible to us geographically and immediately—a form of *convenience* sampling.

How do sampling strategies affect analysis? *Maximum variation sampling*, for example, involves looking for outlier cases to see whether the main patterns still hold, while *homogeneous sampling* focuses on people with similar demographic or social characteristics. The *critical case* is the instance that supports or exemplifies the main findings. Searching deliberately for *confirming and disconfirming cases*, *extreme or deviant cases*, and *typical cases* serves to increase confidence in analytic conclusions. Some strategies benefit inductive, theory-building analysis (e.g., *opportunistic* or *snowball sampling*). *Politically important cases* are salient participants who may need to be included (or excluded) because they connect with politically sensitive issues anticipated in the analysis.

Other strategies can be used for selection of participants prior to data collection. For example, Goetz and LeCompte (1984) offer (a) *comprehensive sampling*—examining every case, instance, or element in a given population; (b) *quota selection*—identifying the major subgroups and then taking an arbitrary number from each; (c) *reputational case selection*—instances chosen on the recommendation of an expert or key participant; and (d) *comparable case selection*—selecting individuals, sites, and groups on the same relevant characteristics over time (a replication strategy). Most of these strategies will increase confidence in analytic findings on the grounds of representativeness.

The sampling strategies we've been discussing can be applied both within and across multiple cases. Let's turn to some of the core issues in each of these domains.

Within-Case Sampling

A qualitative case may range widely in definition from individuals to roles, groups, organizations, processes, and cultures. But even when the case is an individual, the qualitative researcher has many within-case sampling decisions: Which *activities*, *processes*, *events*, *times*, *locations*, and *role partners* will I sample?

Within-case sampling is almost always *nested*—for example, studying children within classrooms within schools within neighborhoods, with regular movement up and down that ladder. For a cardiovascular bypass patient, we might want to sample his or her diet and exercise activities; the processes of understanding, taking in, and acting on medical advice; events such as admission and discharge interviews; time periods, including prehospitalization, hospitalization, and posthospitalization (once every 2 weeks); locations, including the recovery room, the ward, and the patient's home; and role partners, including the patient's physician, ward nurses, dietitian, and spouse.

A second major point is that such sampling should be *theoretically* driven—whether the theory is prespecified or emerges as you go, as in traditional grounded theory's "theoretical sampling." Choices of participants, episodes, and interactions should be driven by a conceptual question, not by a concern for representativeness. To get to the concept, such as *negotiation*, we need to see different instances of it at different moments in different places with different people. The prime concern is with the *conditions* under which the concept or theory operates, not with the generalization of the findings to other populations.

The third point is that within-case sampling has an *iterative* or "rolling" quality, working in progressive waves as the study progresses. Sampling is investigative; we observe, talk to people, and examine artifacts and documents. That leads us to new samples of participants and observations,

new documents. At each step along the evidential trail, we are making sampling decisions to clarify the main patterns, see contrasts, identify exceptions or discrepant instances, and uncover negative instances—where the pattern does not hold. Our analytic conclusions depend deeply on the within-case sampling choices we made.

So, within-case sampling helps us see a local configuration in some depth. What can additional cases do for us, and how do we create a sample of cases?

Multiple-Case Sampling

Multiple-case sampling adds *confidence* to findings. By looking at a range of similar and contrasting cases, we can understand a single-case finding, grounding it by specifying *how* and *where* and, if possible, *why* it carries on as it does. We can strengthen the precision, validity, stability, and trustworthiness of the findings. In other words, we are following a *replication* strategy (Yin, 2018). If a finding holds in one setting and, given its profile, also holds in a comparable setting but does not in a contrasting case, the finding is more robust.

With multiple-case studies, does the issue of *generalizability* change? Essentially, no. We are generalizing from one case to the next on the basis of a match to the underlying theory, not to a larger universe. The choice of cases usually is made on *conceptual* grounds, not on representative grounds. The cases often are arrayed on a continuum (e.g., highly gifted to underachieving pupils), with few exemplars of each, or they are contrasted (e.g., assertive and passive adolescents). Other, unique properties may be added (e.g., some assertive adolescents are from cities, some from rural areas). Because case study researchers examine intact settings in such minute detail, they know all too well that each setting has a few properties it shares with *many* others, some properties it shares with *some* others, and some properties it shares with *no* others. Nevertheless, multiple-case sampling gives us confidence that our emerging theory is transferable, because we have seen it work out—or not work out—in predictable ways across different cases.

How many cases should a multiple-case study have? This question is not answerable on statistical grounds, of course. We have to deal with the issue conceptually: How many cases, and in what kind of sampling frame, would give us confidence in our analytic generalizations? It also depends on how rich and complex the within-case sampling is. With high complexity, a study with more than 10 cases or so can become unwieldy. There are too many data to scan visually and too many permutations to account for. And the problems of practical and intellectual coordination among multiple researchers get very large once you are a staff of more than five people. Still, we've seen multiple-case studies in the 20s and 30s; the price is usually thinner data. If we were forced to recommend a specific number, we would suggest five or six richly researched cases as a minimum for multiple-case sampling adequacy. (We have read outstanding qualitative studies that compared just two, three, and four cases, but their authors did not assert any generalizability.)

Questions of practicality also face us. There is a finite amount of time, with variable access to different participants and events, and an abundance of logistical problems. Being selective calls for some restraint in the classes of data you go after. Here we might suggest some guidelines. For example, useful data would (a) identify *new leads* of importance, (b) *extend* the area of information, (c) *relate* or bridge the already existing elements, (d) reinforce the main *trends*, (e) account for *other information* already in hand, (f) exemplify or provide more evidence for an *important theme*, and (g) *qualify* or *refute* existing information.

Key processes can be identified at the outset or gradually—often via pattern codes, analytic memos, and interim summaries (to be described in later chapters). Being explicit about processes and

collecting comparable data on them will foster cross-case comparability and give you easier access to the core underlying concepts as you get deeper into data collection.

Sampling Advice

1. If you're new to qualitative research, rest assured that there is never enough time to do any study. It is probably a good idea to start with a fallback sample of participants and subsettings: the things you have to cover in light of what you know at that point. That sample will change later, but less than you may think.
2. Just thinking in sampling-frame terms is good for your study's health. If you are talking with one kind of participant, you need to consider why this kind of participant is important and, from there, who else should be interviewed or observed.
3. In complex cases, remember that you are sampling people to get at the characteristics of settings, events, and processes. This means watching out for an overreliance on talk or on observation of participants while neglecting sampling for key events, interactions in different settings, and episodes embodying the emerging patterns in the study. The sampling choices at the start of the study may not be the most pertinent or data-rich ones. A systematic review can sharpen the early and later choices.
4. There is a danger of sampling too narrowly. Go to the meatiest, most study-relevant sources. But it is also important to work a bit at the peripheries—to talk to people who are not central to the phenomenon but are neighbors to it, to people no longer actively involved, to dissidents, renegades, and eccentrics. Spending a day in the adjoining village, school, neighborhood, or clinic is also worth the time, even if you don't see the sense at that point. You may learn a lot and obtain contrasting and comparative information that may help you understand the phenomenon at hand by decentering yourself from a particular way of viewing your primary cases.
5. Spend some time checking whether your sampling frame is feasible. Be sure the time is there, the resources are there, the requisite access to people and places is ensured, and the conditions are right for doing a careful job. Plan to study a bit less, rather than more, and “bank” the extra time. If you are done, the time is yours for a wider or deeper pass at the field.
6. Three kinds of instances have great payoff. The first is the apparently “typical” or “representative” instance. If you can find it, try to find another one. The second is the “negative” or “disconfirming” instance; it gives you both the limits of your conclusions and the point of greatest variation. The third is the “exceptional” or “discrepant” instance. This instance will allow you to qualify your findings and to specify the variations or contingencies in the main patterns observed. Going deliberately after negative and atypical instances is also healthy in itself; it may force you to clarify your concepts, and it may tell you that you indeed have sampled too narrowly.

INSTRUMENTATION

Description and Rationale

Instrumentation comprises specific methods for collecting data. They may be focused on qualitatively or quantitatively organized information and may be loosely to tightly structured. Note that the term *instrumentation* may mean little more than some shorthand devices for observing and

recording events. But note, too, that even when the instrumentation is an open-ended interview or fieldwork observation, some technical choices must be made: Will notes be taken? Of what sort? Will the transaction be audio- or video-recorded? Transcribed?

Kvale and Brinkmann (2015) point out that during an open-ended interview, some analysis and interpretation occur along the way. The person describing his or her “life world” may discover new relationships and patterns during the interview. The researcher who occasionally summarizes or reflects what has been heard is, in fact, condensing and interpreting the flow of meaning. Data are not being collected but rather coauthored.

We’ve discussed how conceptual frameworks, research questions, and sampling have a focusing role within a study. They give some direction to the researcher, before and during fieldwork, by clarifying what he or she wants to find out from whom and why. Knowing what you want to find out, at least initially, leads to the question of *how* you will get that information. That question, in turn, later determines the analyses you can conduct. If I want to find out how crime suspects are arrested and booked, I may decide to *interview* the people associated with this activity (police officers, suspects, and attorneys); *observe* bookings; and collect arrest-relevant *documents* (e.g., regulations, interrogation transcripts). If permitted, I may also take digital *photographs or video* of arrests and bookings. But how much of this instrumentation has to be designed prior to going out to the field, and how much structure should such instruments have?

There are several possible answers to how much preplanning and structuring of instrumentation is desirable: “little” (i.e., hardly any prior instrumentation) to “a lot” (of prior instrumentation, well structured) to “it depends” (on the nature of the study). Each view has supporting arguments; let’s review them in capsule form (Display 2.7 is a summary of some of the main issues in deciding on the appropriate amount of front-end instrumentation).

DISPLAY 2.7 ● Prior Instrumentation: Key Decision Factors

Little Prior Instrumentation	“It Depends”	A Lot of Prior Instrumentation
Rich context description needed		Context less crucial
Concepts inductively grounded in local meanings		Concepts defined ahead by researcher
Exploratory, inductive		Confirmatory, theory-driven
Descriptive intent		Explanatory intent
“Basic” research emphasis		Applied, evaluation or policy emphasis
Single case		Multiple cases
Comparability not too important		Comparability important
Simple, manageable, single-level case		Complex, multilevel, overloading case
Generalizing not a concern		Generalizability/representativeness important
Need to avoid researcher impact		Researcher impact of less concern
Qualitative only, freestanding study		Multimethod study, quantitative included

Source: Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Arguments for Little Prior Instrumentation

1. Predesigned and structured instruments blind the researcher to the site. If the most important phenomena or underlying concepts at work in the field are not collected through the instruments, they will be overlooked or misrepresented.
2. Prior instrumentation is usually stripped of context for purposes of universality, uniformity, and comparability. But qualitative research lives and breathes through seeing site-specific contexts; it is the particularities that produce the generalities, not the reverse.
3. Many qualitative studies involve single cases with few people involved. Who needs standardized questionnaires, observation schedules, or tests—whose usual function is to yield economical, comparable, and parametric distributions for large samples?
4. The lion's share of fieldwork consists of taking field notes; recording events (conversations, meetings); interviewing participants; and examining things (documents, products, artifacts). *Instrumentation* is a misnomer; some orienting questions for observations may be all you need at the start.

Arguments for a Lot of Prior Instrumentation

1. If you know what you are after, there is no reason not to plan in advance how to collect the information.
2. If interview protocols or observation schedules are not focused, too much superfluous information will be collected. An overload of data will compromise the efficiency and power of the analysis.
3. Using the same instruments as in prior studies is an efficient way to converse across studies. Otherwise, the work will be noncomparable, except in a very global way. We need common instruments to construct explanations or predictions, and to make recommendations about practice.
4. A biased or uninformed researcher will ask partial questions, take selective notes, make unreliable observations, and skew information. The data will be invalid and unreliable. Using validated instruments well is the best guarantee of dependable and meaningful findings.

Arguments for "It Depends"

1. If you are running an *exploratory*, largely descriptive study, you do not really know the parameters or dynamics of a social setting. So, heavy initial instrumentation or closed-ended devices are inappropriate. If, however, you are doing a *confirmatory* study, with relatively focused research questions and a well-bounded sample of persons, events, and processes, then well-structured instrument designs are a logical choice. Within a given study, there can be both exploratory and confirmatory aspects that call for differential front-end structure, or there can be exploratory and confirmatory *times*, with exploration often called for at the outset and confirmation near the end.
2. A *single-case* study calls for less front-end preparation than does a *multiple-case* study. The latter is looking forward to cross-case comparison, which requires some standardization of instruments so that findings can be laid side by side in the course of analysis. Similarly, a *freestanding* study has fewer constraints than a *multimethod* study. A *basic* study often needs less advance organizing than an *evaluation* or *policy* study. In the latter cases, the focus is tighter and the instrumentation more closely keyed to the variables of interest.