

The
Coding
Manual for
Qualitative
Researchers

»»» 5E «««

Johnny
Saldaña

S Sage



1 Oliver's Yard
55 City Road
London EC1Y 1SP

2455 Teller Road
Thousand Oaks
California 91320

Unit No 323-333, Third Floor, F-Block
International Trade Tower
Nehru Place, New Delhi – 110 019

8 Marina View Suite 43-053
Asia Square Tower 1
Singapore 018960

Editor: Jai Seaman
Assistant editor: Becky Oliver
Production editor: Ian Antcliff
Copyeditor: Richard Walshe
Proofreader: Brian McDowell
Marketing manager: Ruslana Khatagova
Cover design: Shaun Mercier
Typeset by: C&M Digitals (P) Ltd, Chennai, India
Printed in the UK

© 2025 Johnny Saldaña

Apart from any fair dealing for the purposes of research, private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may not be reproduced, stored or transmitted in any form, or by any means, without the prior permission in writing of the publisher, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publisher.

Library of Congress Control Number: 2024944456

British Library Cataloguing in Publication data

A catalogue record for this book is available from the British Library

ISBN 978-1-5296-800-10
ISBN 978-1-5296-8000-3 (pbk)

For
Norman K. Denzin,
who opened the door and invited me in ...



Brief Contents

<i>Detailed Contents</i>	ix
<i>List of Figures</i>	xix
<i>Acknowledgements</i>	xxiii
<i>About the Author</i>	xxv
<i>Preface to the Fifth Edition</i>	xxvii
<i>Online Resources</i>	xxix
Part I Coding Foundations	1
1 An Introduction to Codes and Coding	3
2 Fundamental Coding Methods and Techniques	27
3 Writing Analytic Memos About Narrative and Visual Data	65
Part II First Cycle Coding Methods	97
4 Selecting First Cycle Coding Methods	99
5 Grammatical Coding Methods	125
6 Elemental Coding Methods	145
7 Affective Coding Methods	183
8 Literary and Language Coding Methods	211
9 Exploratory Coding Methods	245
10 Procedural Coding Methods	263
11 Methods of Themeing the Data	293
Part III Second Cycle Coding Methods	315
12 Transitioning From First to Second Cycle Coding Methods	317
13 Second Cycle Grounded Theory Coding Methods	343
14 Second Cycle Cumulative Coding Methods	365
15 After First and Second Cycle Coding Methods	387
<i>Appendix A: A Glossary of Coding Methods</i>	409
<i>Appendix B: A Glossary of Analytic Recommendations</i>	419
<i>Appendix C: Forms for Additional Coding Methods</i>	425
<i>References</i>	429
<i>Index</i>	455



Detailed Contents

<i>List of Figures</i>	xix
<i>Acknowledgements</i>	xxiii
<i>About the Author</i>	xxv
<i>Preface to the Fifth Edition</i>	xxvii
<i>Online Resources</i>	xxix

Part I Coding Foundations 1

1 An Introduction to Codes and Coding 3

Chapter Summary	3
Purposes of the Manual	4
What Is a Code?	5
Coding examples	6
Coding for patterns	8
Coding lenses, filters, and angles	10
Coding as a heuristic	12
Codifying and Categorizing	13
From codes to categories	13
Recoding and recategorizing	15
From codes and categories to theory	18
The differences between codes and themes	21
Necessary Personal Attributes for Coding	22
Critiques Against Coding	23

2 Fundamental Coding Methods and Techniques 27

Chapter Summary	27
Coding as Craft	28
What Gets Coded?	28
The coded researcher	28
Amounts of data to code	29
Preliminary Coding Techniques	30
Data layout	30
Pre-coding	32
Preliminary jottings	33
Questions to consider as you code	34
Coding contrasting data	35
The Numbers of Codes	35

“Lumping” and “splitting” the data	36
Code proliferation	37
The quantities of qualities	40
“Quantitizing” the qualitative	41
Deductive and/or Inductive Coding	43
The Codebook or Code List	44
Manual and CAQDAS Coding	48
Coding manually	48
Coding electronically	50
Data formatting for CAQDAS	53
Coding capabilities with CAQDAS	53
Searches and queries with CAQDAS	55
Solo and Team Coding	57
Coding solo	57
Coding collaboratively	57
Multiple coders	59
On AI and ChatGPT	60
On Translated Data and Coding	62
On Method	63

3 Writing Analytic Memos About Narrative and Visual Data	65
Chapter Summary	65
The Purposes of Analytic Memo Writing	66
What Is an Analytic Memo?	66
Examples of Analytic Memos	67
Additional analytic memo topics and prompts	79
Reflection and refraction	80
Coding and Categorizing Analytic Memos	81
Analytic memos generate codes, categories, themes, and concepts	81
Grounded Theory and Its Coding Canon	82
Analyzing Visual Data	82
Photographs	84
Documents and artifacts	89
Live and video-recorded action	91
Recommended guidance	95

Part II First Cycle Coding Methods 97

4 Selecting First Cycle Coding Methods	99
Chapter Summary	99
The Coding Cycles	100
Selecting the Appropriate Coding Method(s)	100
Various perspectives on coding decisions	102

Research question alignment	103
Paradigmatic, conceptual, and methodological considerations	104
Coding and a priori goals	104
Coding in mixed methods studies	106
Exploratory coding	107
Descriptive Coding as a common default	107
“Generic” coding methods	109
New and hybrid coding schemes	110
General criteria for coding decisions	111
On overwhelming fear	113
A First and Second Cycle Coding Example	113
Overview of First Cycle Methods	121
The Coding Methods Profiles	122
Sources	122
Description	122
Applications	122
Example	123
Analysis	124
Notes	124

5 Grammatical Coding Methods 125

Chapter Summary	125
Attribute Coding	126
Sources	126
Description	126
Applications	126
Examples	126
Analysis	127
Notes	129
Magnitude Coding	129
Sources	129
Description	129
Applications	129
Examples	130
Analysis	134
Notes	137
Subcoding	137
Sources	137
Description	137
Applications	137
Example	138
Analysis	138
Notes	140

Simultaneous Coding	140
Source	140
Description	140
Applications	140
Examples	141
Analysis	142
Notes	143
6 Elemental Coding Methods	145
Chapter Summary	145
Structural Coding	146
Sources	146
Description	146
Applications	146
Example	146
Analysis	148
Notes	149
Descriptive Coding	149
Sources	149
Description	149
Applications	150
Example	150
Notes	153
In Vivo Coding	153
Sources	153
Description	153
Applications	154
Example	154
Analysis	155
Process Coding	160
Sources	160
Description	160
Applications	161
Example	161
Notes	165
Initial Coding	165
Sources	165
Description	166
Applications	166
Example	166
Analysis	168
Notes	170
Icon Coding	170

Sources	170
Description	170
Applications	170
Example	171
Analysis	172
Notes	175
Concept Coding	177
Sources	177
Description	177
Applications	178
Example	178
Analysis	179
Notes	182
7 Affective Coding Methods	183
Chapter Summary	183
Emotion Coding	184
Sources	184
Description	184
Applications	184
Example	185
Analysis	186
Notes	192
Values Coding	192
Sources	192
Description	192
Applications	194
Example	194
Analysis	195
Notes	200
Versus Coding	200
Sources	200
Description	200
Applications	201
Example	201
Analysis	202
Notes	205
Evaluation Coding	205
Sources	205
Description	206
Applications	206
Example	206
Analysis	208
Notes	210

8 Literary and Language Coding Methods	211
Chapter Summary	211
Dramaturgical Coding	212
Sources	212
Description	212
Applications	212
Example	213
Analysis	216
Notes	220
Motif Coding	221
Sources	221
Description	221
Applications	221
Example	222
Notes	225
Narrative Coding	225
Sources	225
Description	225
Applications	226
Example	226
Notes	231
Metaphor Coding	232
Sources	232
Description	232
Applications	232
Example	233
Analysis	234
Notes	237
Verbal Exchange Coding	238
Source	238
Description	238
Applications	239
Example	239
Analysis	241
Notes	244
9 Exploratory Coding Methods	245
Chapter Summary	245
Holistic Coding	246
Source	246
Description	246
Applications	246
Example	246

Analysis	247
Notes	248
Provisional Coding	248
Sources	248
Description	249
Applications	249
Example	249
Analysis	250
Notes	252
Hypothesis Coding	252
Sources	252
Description	252
Applications	252
Example	253
Analysis	254
Notes	255
Eclectic Coding	256
Sources	256
Description	256
Applications	256
Example	256
Analysis	257
Notes	261
10 Procedural Coding Methods	263
Chapter Summary	263
Protocol Coding	264
Sources	264
Description	264
Applications	264
Example	264
Analysis	265
Notes	266
OCM (Outline of Cultural Materials) Coding	267
Sources	267
Description	267
Applications	268
Example	268
Analysis	269
Notes	270
Domain and Taxonomic Coding	270
Sources	270
Description	270

Applications	272
Example	272
Analysis	274
Notes	276
Causation Coding	276
Sources	276
Description	276
Applications	278
Example	280
Analysis	282
Notes	290
11 Methods of Themeing the Data	293
Chapter Summary	293
Themes in Qualitative Data Analysis	294
Themeing the Data: Categorically	296
Sources	296
Description	296
Applications	296
Example	297
Analysis	299
Notes	305
Themeing the Data: Phenomenologically	305
Sources	305
Description	305
Applications	306
Example	306
Analysis	308
Notes	312
Metaethnography, Metasummary, and Metasynthesis	312
Part III Second Cycle Coding Methods	315
12 Transitioning from First to Second Cycle Coding Methods	317
Chapter Summary	317
Post-coding Transitions	318
Code Mapping and Landscaping	319
Code Mapping	319
Code Landscaping	325
Operational Model Diagramming	330
Additional Transition Methods	333
Coding the Codes	333
Code Charting	334
Pie-chart coding	335

Tabletop Categories	336
From Codes to Themes	337
“Shop Talking” Through the Study	337
ChatGPT-4 Follow-Up	338
Transitioning to Second Cycle Coding Methods	340
13 Second Cycle Grounded Theory Coding Methods	343
Chapter Summary	343
On Grounded Theory	344
Focused Coding	346
Source	346
Description	346
Applications	346
Example	346
Analysis	349
Notes	351
Axial Coding	351
Sources	351
Description	351
Applications	352
Example	352
Analysis	352
Notes	357
Theoretical Coding	357
Sources	357
Description	357
Applications	358
Example	359
Analysis	360
Notes	363
14 Second Cycle Cumulative Coding Methods	365
Chapter Summary	365
Pattern Coding	366
Source	366
Description	366
Applications	366
Example	366
Analysis	369
Notes	370
Elaborative Coding	370
Source	370
Description	370
Applications	371

Example	371
Analysis	372
Notes	375
Longitudinal Coding	375
Sources	375
Description	375
Applications	379
Example	379
Analysis	381
Notes	385
15 After First and Second Cycle Coding Methods	387
Chapter Summary	387
Post-coding and Pre-writing Transitions	388
Focusing Strategies	388
The “Top 10” List	389
The Study’s “Trinity”	389
Codeweaving	390
The “touch test”	390
Findings “at a Glance”	391
From Coding to Theorizing	392
Elements of a Theory	392
Categories to Concepts	393
Category Relationships	396
Analytic Memos as Sources for Theory	397
Formatting Matters	398
Rich Text Emphasis	398
Headings and Subheadings	398
Writing About Coding	399
Ordering and Reordering	402
Analytic Storylining	402
One Thing at a Time	403
Begin with the Conclusion	404
Assistance from Others	404
Peer and Online Support	404
Searching for “Buried Treasure”	405
Closure	405
<i>Appendix A: A Glossary of Coding Methods</i>	409
<i>Appendix B: A Glossary of Analytic Recommendations</i>	419
<i>Appendix C: Forms for Additional Coding Methods</i>	425
<i>References</i>	429
<i>Index</i>	455

List of Figures

1.1	A streamlined codes-to-theory model for qualitative inquiry	18
1.2	From data to themes (Source: Naeem et al., 2023, p. 3)	20
2.1	An ATLAS.ti code frequency screenshot (courtesy of Ivana Radivojevic, atlasti.com)	36
2.2	Individual In Vivo Codes on separate slips of paper (photo by the study's researcher, anonymity requested)	38
2.3	A sample codebook (Source: Saldaña & Omasta, 2018, pp. 228–9)	46
2.4	A Microsoft Word field notes document with codes in a right-margin text box (courtesy of Teresa Minarsich)	49
2.5	A Microsoft Excel spreadsheet with mixed methods data and codes in its cells	50
2.6	An ATLAS.ti word cloud (courtesy of Ivana Radivojevic, atlasti.com)	52
2.7	A QDA Miner coding co-occurrences 3D map (courtesy of Normand Péladeau, Provalis Research, provalisresearch.com)	52
2.8	A screenshot from Quirkos software (courtesy of Daniel Turner, quirkos.com; original in color)	54
2.9	A Dedoose Code Co-occurrence matrix (courtesy of Sara E. Grummert, dedoose.com)	55
2.10	NVivo 14 coded interview comparison (courtesy of Pattie Hall and Silvana di Gregorio, used with permission from Lumivero, lumivero.com/products/nvivo/)	56
2.11	Category comparison after Values Coding	61
2.12	Thematic analysis comparison	62
3.1	An analytic memo sketch	74
3.2	An elemental model for developing “classic” grounded theory	83
3.3	NVivo 14 image coding by region (courtesy of Pattie Hall and Silvana di Gregorio, used with permission from Lumivero, lumivero.com/products/nvivo/)	85
3.4	webQDA image coding by region (courtesy of António Pedro Costa, webqda.net)	85
3.5	Downtown suburban renovation (original in color; photo by Eliana Watson)	86
3.6	A university student in a library (original in black and white; photo by Eliana Watson)	87

3.7	A mask representing a service member's frustrations and existential reflections with military experiences (original in color) (Source: Walker et al. (2017). The masks were made by military service members during art therapy treatment at the National Intrepid Center of Excellence, Walter Reed National Military Medical Center, Bethesda, MD)	91
3.8	A screenshot from Transana software (courtesy of David K. Woods, transana.com)	94
3.9	A screenshot from INTERACT software (courtesy of Reinhard Grassl, Mangold International, mangold-international.com)	94
4.1	First cycle and second cycle coding methods (see Appendix A for descriptions)	101
4.2	An In Vivo coded excerpt (Source: Terkel, 1972, p. xxxi)	114
4.3	The interrelationship between three Focused Codes	120
4.4	An arts-based model of interrelationship	120
6.1	An illustrated process for spreading rumors	163
6.2	Emoji coding with Quirkos software (courtesy of Daniel Turner, quirkos.com)	176
6.3	Circular Plot Type: Mina (Source: Maldonado et al., 2023, p. 9)	176
7.1	An emotional arc for a participant's account	187
7.2	An integrated system of values, attitudes, and beliefs	193
7.3	A fourth-grade teacher's values, attitudes, and beliefs	196
7.4	An ATLAS.ti Sankey flow diagram of a values system (courtesy of Ivana Radivojevic, atlasti.com)	198
8.1	Social media see-saw: positive and negative affect influences (Source: Weinstein, 2018, p. 3617)	236
10.1	Excerpts from a taxonomic tree diagram of the ways children oppress others	275
10.2	A causation model of speech classes as LifeForce Preparation	286
10.3	A gear graphic template for Causation Coding triplets	286
11.1	A thematic map	301
11.2	An arts-based thematic representation (Source: www.freepik.com/free-photo/close-up-elderly-man-holding-asthma-inhaler_10851281.htm#)	313
12.1	Codes, subcodes, and numeric frequencies in Delve (courtesy of LaiYee Ho, Delve qualitative analysis software, www.delvetool.com)	325
12.2	A word cloud graphic of an interview transcript (created with tagcrowd.com)	326

12.3	Code landscaping of a major category and its subcategories and sub-subcategories	327
12.4	A MAXQDA code matrix browser (courtesy of Maren Heise, VERBI Software/MAXQDA, maxqda.com)	329
12.5	A QDA Miner heatmap (courtesy of Normand Péladeau, Provalis Research, provalisresearch.com)	329
12.6	An operational model diagram of an urban teacher's cultural shock and survival processes	331
12.7	A QDA Miner correspondence plot (courtesy of Normand Péladeau, Provalis Research, provalisresearch.com)	332
12.8	An ATLAS.ti network diagram (courtesy of Ivana Radivojevic, atlasti.com)	333
12.9	A data and codes summary table	335
12.10	A pie-chart representation of a teacher's cultural shock and adaptation	336
12.11	An arrangement of tabletop categories	336
12.12	A ChatGPT-4o thematic illustration of Sam's interview (original in color)	339
12.13	A streamlined codes-to-theory model for qualitative inquiry	341
13.1	An elemental model for developing "classic" grounded theory	345
13.2	A tree diagram from categories and subcategories	349
13.3	Two Axial Codes and their related categories	353
13.4	A simple properties and dimensions table derived from Axial Coding	356
13.5	A diagram for a central/core category and its major processes	360
13.6	A flowchart matrix of Focused Codes to a grounded theory	361
14.1	Codes assembled to determine their Pattern Codes	368
14.2	A longitudinal qualitative data summary matrix (Source: Saldaña, 2003, courtesy of Rowman and Littlefield Publishing Group/AltaMira Press)	377
14.3	A sample handwritten longitudinal matrix with qualitative data	382
15.1	A trinity of concepts as a Venn diagram (based on Soklaridis, 2009)	390
15.2	NVivo 14 concept map (courtesy of Pattie Hall and Silvana di Gregorio, used with permission from Lumivero, lumivero.com/products/nvivo/)	395
15.3	Coding process diagram (Source: Kelly, Harbin, & Schwartz, 2023, p. 5)	402



Acknowledgements

Thanks are extended to:

Patrick Brindle, former editor of Sage Publications UK, who initially encouraged me to create the first edition of this manual; and Helen Salmon, Acquisitions Editor of Sage Publications USA, for her devoted support of my work;

Jai Seaman, Becky Oliver, and Ian Antcliff of Sage Publications UK for their pre-publication guidance throughout this project;

My qualitative research instructors at Arizona State University: Tom Barone, Mary Lee Smith, Sarah Amira De la Garza, and Sarah J. Tracy, for introducing me to the methods and literature of qualitative inquiry;

Harry F. Wolcott, Norman K. Denzin, Yvonna S. Lincoln, Mitch Allen, Kathy Charmaz, Hannah Shakespeare, Ray Maietta, Paul Mihas, Jeff M. Petruzzelli, Ronald J. Chenail, Adam Rosenthal, Coleman A. Jennings, Linda Essig, Eric D. Teman, Leo A. Mallette, Kakali Bhattacharya, Liora Bresler, Patricia Leavy, Steven A. Harvey, Joe Norris, Charles Vanover, Ann Whitehouse, Carol Burns, Gary R. Burns, Laura A. McCammon, and Matt Omasta for their inspiration, mentorship, and collegial support for my own research endeavors;

Patricia Sachs Chess for Figure 1.1 revision ideas;

Eliana Watson for Figures 3.5 and 3.6 photography;

Melissa S. Walker and Thomas J. DeGraba of the National Intrepid Center of Excellence, Walter Reed National Military Medical Center, Bethesda, MD, and Girija Kaimal, Adele M. L. Gonzaga, and Katherine A. Myers-Coffman of the College of Nursing and Health Professions, Drexel University, Philadelphia, PA, for Figure 3.7 permission;

David K. Woods (Transana), Maren Heise (VERBI Software/MAXQDA), Ivana Radivojevic (ATLAS.ti), Normand Péladeau (QDA Miner), Sara E. Grummert (Dedoose), Pattie Hall and Silvana di Gregorio (Lumivero/NVivo), António Pedro Costa (webQDA), Reinhard Grassl (INTERACT/Mangold International), LaiYee Ho (Delve), and Daniel Turner (Quirkos) for their assistance with CAQDAS screenshots and permissions;

Pixabay artists for icon and emoji designs: Raphael Silva, Ali Uyar, Pete Linforth, JS Inspire Studio, Clker-Free-Vector-Images, CryptoSkylark, and 13724641;

Poem in Chapter 6 reproduced with permission of Oxford University Press through PLSclear;

Melanie Birks, Jane Mills, AltaMira Press, and Oxford University Press for permissions to reprint extended excerpts from selected works;

Cody Goulder, Angie Hines, Lisa A. Kramer, Laura A. McCammon, Teresa Minarsich, Matt Omasta, Abdul Mohamed, and Asantewa Sunni-Ali for selected data contributions and analytic displays from fieldwork;

Jonathan Neelands and Tony Goode for the manual's formatting ideas;

Jim Simpson for the "shop talk" and support; and

All writers cited and referenced in this manual, whose works, brought together in this collection, provide us with rich ideas for qualitative data analysis.

About the Author

Johnny Saldaña is Professor Emeritus from Arizona State University's (ASU) School of Film, Dance, and Theatre in the Herberger Institute for Design and the Arts, where he taught from 1981 to 2014. He received his BFA in Drama and English Education in 1976, and MFA in Drama Education in 1979 from the University of Texas at Austin.

He is the author of *Longitudinal Qualitative Research: Analyzing Change through Time* (AltaMira Press, 2003); *Fundamentals of Qualitative Research* (Oxford University Press, 2011); *Ethnotheatre: Research from Page to Stage* (Left Coast Press, 2011); *Thinking Qualitatively: Methods of Mind* (Sage Publications, 2015); *Developing Theory Through Qualitative Inquiry* (Sage Publications, 2025); a commissioned title for Routledge's World Library of Educationalists Series, *Writing Qualitatively: The Selected Works of Johnny Saldaña* (Routledge, 2018); co-author with the late Matthew B. Miles and A. Michael Huberman for *Qualitative Data Analysis: A Methods Sourcebook* (4th ed., Sage Publications, 2020); co-author with Matt Omasta for *Qualitative Research: Analyzing Life* (2nd ed., Sage Publications, 2022); and the editor of *Ethnodrama: An Anthology of Reality Theatre* (AltaMira Press, 2005). Previous editions of *The Coding Manual for Qualitative Researchers* have been translated into Korean, Turkish, and Chinese-Simplified.

Saldaña's methods works have been cited and referenced in more than 50,000 research studies conducted in over 135 countries in disciplines such as K-12 and higher education, medicine and health care, technology and social media, business and economics, government and social services, the fine arts, the social sciences, human development, and communication. He has published a wide range of research articles in journals such as *Research in Drama Education*, *The Qualitative Report*, *Multicultural Perspectives*, *Youth Theatre Journal*, *Journal of Curriculum and Pedagogy*, *Teaching Theatre*, *Research Studies in Music Education*, *Cultural Studies* ↔ *Critical Methodologies*, the *International Journal of Qualitative Methods*, the *International Review of Qualitative Research*, and *Qualitative Inquiry*, and has contributed several chapters to research methods handbooks.

His research in qualitative inquiry, data analysis, and performance ethnography has received awards from the American Alliance for Theatre & Education, the National Communication Association—Ethnography Division, the American Educational Research Association's (AERA) Arts-Based Educational Research Special Interest Group (SIG), the AERA Qualitative Research SIG, New York University's Program in Educational Theatre, the Children's Theatre Foundation of America, and the ASU Herberger Institute for Design and the Arts.



Preface to the Fifth Edition

This may or may not be the last edition published of *The Coding Manual for Qualitative Researchers*. At the time of this writing (fall 2024), Artificial Intelligence (AI), ChatGPT, and comparable platforms have entered the analytic landscape, providing researchers with new heuristics for data entry and initial analyses. My own experiments with ChatGPT-4 (now in versions 4o and 4o1) showed me that the program can summarize and synthesize original qualitative data inputs quickly with surprisingly descriptive detail. And when I prompted the app to provide me with the data's patterns and themes, I found the displayed results quite credible and somewhat comparable to my own human-generated analyses.

But like virtually all digital programs for qualitative data analysis, ChatGPT-4 can only go so far at this stage in its development. Missing from the program's analyses are background contexts, detailed literature reviews, interpretive insights, higher-level abstractions, creative vocabulary use, evaluative judgment, cultural nuance, ethnic/racial literacy, critical capacities, reflexive introspection, experiential intuition, data privacy/security, theory development, and astute emotional intelligence. Qualitative researchers experimenting with AI generally assert that these programs can assist and augment, but not replace, the human researcher.

As several methods profiles in this manual will later explain, ChatGPT-4 can be "taught" to code data with specific, carefully phrased prompts. And if you ask what the program already knows about a particular analytic method or methodology such as grounded theory or longitudinal analysis, more often than not, it will possess that knowledge before you input your data. Nevertheless, the purpose and utility of coding as an analytic approach must be questioned when AI programs need not apply such measures to quickly generate categories and themes—two primary outcomes of coding qualitative data. So, in the AI era, is coding still relevant or even necessary?

It would be foolish if not irresponsible to accept ChatGPT-4's data analyses at face value without the researcher's critical scrutiny of the program's results. Even the program itself includes a permanently pinned message at the bottom of its home page: "ChatGPT can make mistakes. Consider checking important information." To me, a trustworthy qualitative researcher knows the legacies of their discipline's analytic methods, which include the varieties of codes and coding. Such knowledge informs the investigator's epistemological pathways with their generated data and serves the analyst's nuanced review of a digital research assistant's work. To readers and audiences, we are accountable not only for our findings but for how we found them in the first place. AI structures its answers through quantitative algorithms. But humans structure their answers through a myriad of qualitative heuristics and methods as occasions necessitate.

And sometimes those occasions require an In Vivo Code, an Emotion Code, a Dramaturgical Code, a Causation Code, a Theoretical Code, and so on.

Digital technology's evolutionary future is unpredictable, and so is its integration with qualitative data analysis. Rather than shy away from the unknown, researchers need to explore and experiment with AI's possibilities as well as its limitations. I have enhanced the discussions in this edition's coding methods profiles with recommendations for ChatGPT-4 applications. Use these heuristics as supplements, not substitutes, for your own analytic work and deep reflection.

In the fifth edition of this manual, a new first cycle coding method, Icon Coding, joins the 35 others in the collection. Analytic software screenshots and academic references have been updated. Several new figures have been added throughout the manual. Revised examples and analyses are provided for analytic memos, photograph and video analysis, metaethnography, Magnitude Coding, Structural Coding, Emotion Coding, Dramaturgical Coding, and Pattern Coding. The online resources now include a digital photograph and social media data in addition to interview transcripts, participant observation fieldnotes, and document data for analytic exercises.

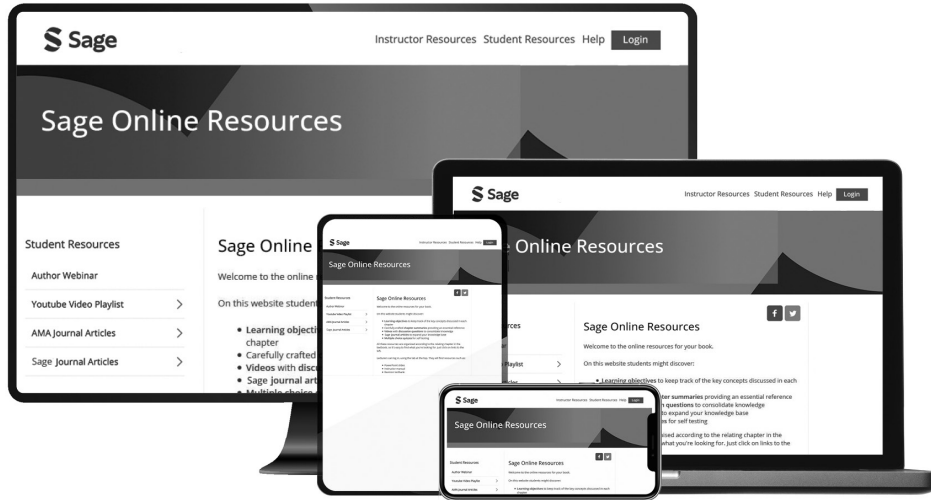
I stress at the beginning and ending of this book that coding is just *one* way, not *the* way to analyze qualitative data. Even if you prefer other analytic approaches such as assertion development, content analysis, or theory-based holistic interpretation, this manual offers guidance for non-coded analytic reflection, along with resources and references for learning more about the field's diverse data analysis methods.

Graduate students and their professors have told me how much they appreciate the manual's extensive citations, clarity, and mentorship tone for their professional development and projects. Yet I must also extend my own thanks and gratitude to the legacy of scholars whose publications provide rich sources for several of the ideas collected in this book. I give credit where credit is due by quoting, citing, and referencing their works through fair use guidelines.

My primary role as author of this manual is to serve as a contemporary archivist of the vast literature on qualitative methods, and to selectively display and explain relevant material about codes, coding, and qualitative data analysis. But the amounts of books and e-resources on the subject has increased exponentially over the past few decades, and I cannot possibly survey everything in the area. I must rely on you to bring your specific disciplinary knowledge base and your rich personal experiences to supplement the material included in this resource. I hope that this expanded fifth edition of *The Coding Manual for Qualitative Researchers* and its companion website offer readers even more pragmatic guidance for qualitative data analysis.

Johnny Saldaña,
Professor Emeritus

Online Resources



The fifth edition of *The Coding Manual for Qualitative Researchers* is supported by a wealth of online resources to aid your study, which are available at <https://study.sagepub.com/saldanacoding5e>

For students

Exercises and Activities to be done by yourself, or with others, help attune you to basic principles of coding, pattern development, categorization and qualitative data analysis.

Sample Data, Code Lists and Transcripts give you the opportunity to practice using qualitative coding and themeing techniques with real data.

CAQDAS Weblinks, along with the author's annotations and recommendations, enable you to make an informed choice about using software packages for qualitative data management and analysis.



“The primary purpose of qualitative coding is to organize and categorize the data, facilitating the extraction of meaningful insights and interpretations.” (ChatGPT-4)



Part I

Coding Foundations

1	An Introduction to Codes and Coding.....	3
2	Fundamental Coding Methods and Techniques.....	27
3	Writing Analytic Memos About Narrative and Visual Data	65



1

An Introduction to Codes and Coding

Chapter Summary

This chapter first presents the purposes and goals of *The Coding Manual for Qualitative Researchers*. It then provides definitions and examples of codes and categories and their roles in qualitative data analysis. Fundamental principles of coding follow, and the chapter concludes with reflections on necessary researcher attributes and the role of method in coding.

Purposes of the Manual

The three primary purposes of this manual are:

- 1 to illustrate the functions of codes, coding, and analytic memo writing during the qualitative data generation and analysis processes;
- 2 to profile a selected yet diverse repertoire of coding methods generally applied in qualitative data analysis; and
- 3 to provide readers with sources, descriptions, recommended applications, examples, and additional resources for coding and further analyzing and synthesizing qualitative data.

This book serves as a reference to supplement works in research design and fieldwork, and assumes that readers have some beginning knowledge of the fundamentals of qualitative inquiry. The manual focuses exclusively on codes and coding methods and how they play a role in the qualitative data analysis process.

The manual does not subscribe to any one specific qualitative research methodology or method. Throughout this book you will read a breadth of perspectives on codes and coding, sometimes purposely juxtaposed to illustrate and highlight diverse opinions among scholars in the field. The following demonstrates just two examples of such professional divergence:

Any researcher who wishes to become proficient at doing qualitative analysis must learn to code well and easily. The excellence of the research rests in large part on the excellence of the coding. (Strauss, 1987, p. 27)

But the strongest objection to coding as a way to analyze qualitative research interviews is not philosophical but the fact that it does not and cannot work. It is impossible in practice. (Packer, 2018, p. 94)

No one, including myself, can claim final authority on the utility of coding or the “best” way to analyze qualitative data. In fact, I take moderate liberty in adapting and even renaming selected prescribed coding methods. I do this not to standardize terminology within the field, but simply to employ consistency throughout this particular resource.

I must also emphasize at the very beginning that *there are times when coding the data is absolutely necessary, and times when it is most inappropriate for the study at hand*. All research questions, conceptual frameworks, methodologies, and fieldwork parameters are context-specific. Also, whether you choose to code or not depends on your individual value, attitude, and belief systems about qualitative inquiry.

Coding is just *one* way of analyzing qualitative data, not *the* way. Be cautious of those who demonize the method outright. And be equally cautious of those who swear unyielding affinity to codes or what has been colloquially labeled “coding fetishism.” I prefer that

you yourself, rather than some presumptive theorist or hardcore methodologist, determine whether coding is appropriate for your particular research project.

Most textbooks in qualitative research limit their discussions about coding to the writer's prescribed, preferred, or signature methods. I wanted to provide in a single resource a selected collection of various coding methods developed by other researchers and myself that provides students and colleagues with a useful reference for classroom exercises and assignments, and for their own independent research for thesis and dissertation fieldwork and future qualitative studies. But by no means is this manual an exhaustive resource. If you need additional information and explanation about the coding methods, check the Appendices and References.

This book is not necessarily meant to be read from cover to cover, but it certainly can be if you wish to acquaint yourself with all 36 coding methods' profiles and their analytic possibilities. I doubt you will use every coding method included in this reference manual for your particular research endeavors throughout your career, but they are available here on an "as-needed" basis for your unique projects. Like an academic curriculum, the sequential order of the profiles has been carefully considered. They do not necessarily progress in a linear manner from simple to complex, but are clustered generally from the fundamental to the intermediate to the advanced.

What Is a Code?

Qualitative research is an umbrella term for a wide variety of approaches to and methods for the study of natural social life. The data generated and analyzed is primarily but not exclusively non-quantitative in character. Qualitative data (also known as empirical materials) can consist of interview transcripts, participant observation field notes, journals, documents, open-ended survey responses, drawings, artifacts, photographs, video, websites, e-mail correspondence, social media texts, academic and fictional literature, and so on. There are several approaches to the management of the typically voluminous amounts of qualitative data generated for a particular research project, and coding is the analytic heuristic (from the Greek, *heuriskein*, meaning "to discover") profiled in this manual.

A code in qualitative analysis is most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data. The portion of data coded during initial first cycle coding processes can range in magnitude from a single word to a full paragraph, an entire page of text, or a stream of moving images.

In second cycle coding processes, the portions coded can be the exact same units, longer passages of text, analytic memos about the data, and even a reconfiguration of the codes themselves developed thus far. Charmaz (2001) describes coding as the "critical link" between data generation and their explanation of meaning. To me, first cycle coding *is* analysis—taking things apart. Second cycle coding is synthesis—putting things together into new assemblages of meaning (Saldaña & Omasta, 2022).

Do not confuse the use of code in qualitative data analysis with the use of code in semiotics, even though slight parallels exist between the two applications. In semiotics, a code relates to the interpretation of symbols in their specific social and cultural contexts. And while some code choices by the analyst may appear metaphoric, most codes are not metaphors (according to the principles established by Lakoff & Johnson, 2003). Also, coding in computer software applications refers to algorithmic program writing and software engineering. That purpose and meaning are not what is covered in this book, though the disciplines of digital technology and software engineering have found the contents of this manual surprisingly useful.

In qualitative data analysis, a code is a researcher-generated interpretation that symbolizes or “translates” data (Vogt, Vogt, Gardner, & Haeffele, 2014, p. 13), and thus attributes meaning to each individual datum for later purposes of pattern detection, categorization, theme, assertion or proposition development, theory building, and other analytic processes. Clarke and Braun (2016) suggest that a well-developed code should stand on its own and solidly represent the data from which it originated, while methodologist Paul Mihas sees codes as “invitations and openings” to further inquiry. Just as a title represents and captures a book, film, or poem’s primary content and essence (e.g., *The Scarlet Letter*, *Black Panther*, “To Live in the Borderlands”), so does a code represent and capture a datum’s primary content and essence.

Throughout this manual, I use a consistent set of rich text features to differentiate symbolic summaries:

- CODES, SUBCODES, AND THEMES ARE SET IN CAPS
- *SUBTHEMES ARE SET IN ITALICIZED CAPS*
- **Categories are set in bold**
- ***Subcategories are set in bold italic.***

Coding Examples

An example of a coded datum, as it is presented in this manual, looks like this when taken from a set of field notes about an urban neighborhood. The one-word capitalized code in the right column, SECURITY, is a Descriptive Code, which summarizes the primary topic of the excerpt that follows the same superscript number. Linking a specific data excerpt with its specific code through comparable numbers has been termed geocoding:

Code example 1.1

¹ I notice that the grand majority of homes have chain link fences in front of them. There are many dogs (mostly German shepherds) with signs on fences that say “Beware of the Dog.”

¹ SECURITY

Here is an example of several codes applied to data from an interview transcript in which a high school senior describes his favorite teacher. The codes are based on what outcomes the student receives from his mentor. Note that one of the codes is taken directly from what the participant himself says and is placed in quotation marks—this is called an In Vivo Code (profiled in Chapter 6):

Code example 1.2

¹ He cares about me. He has never told me but he does.	¹ SENSE OF SELF-WORTH
² He's always been there for me, even when my parents were not. He's one of the few things that I hold as a constant in my life. So it's nice. ³ I really feel comfortable around him.	² STABILITY
	³ "COMFORTABLE"

Did you agree with the codes? Did other words or phrases run through your mind as you read the data? It is all right if your choices differed from mine. Coding is not a precise science; it is primarily an interpretive act. Also be aware that a code can sometimes *summarize*, *distill*, or *condense* data, not *reduce* them. Reduction implies something lost. Madden (2023) notes that such analytic work does not diminish but “value adds” to the research story (p. 10).

Also, some may wonder, “Where did those three codes ‘come from’?” They came from my thinking, my disciplinary background knowledge, my earlier research with the case, and my creativity. One common phrase in the research methods literature is that codes, categories, themes, and so on seem to mysteriously and magically “emerge.” Yes, ideas for labels sometime occur spontaneously in your mind and may seem to appear “out of nowhere.” But codes, categories, themes, and other qualitative data summations are actively *constructed*, *formulated*, *created*, and *revised* by the researcher, not through some elusive process. Ethnographer Steven A. Harvey astutely shared, “We should be more careful with the word ‘emerge’ so that students don’t expect to just twist their transcripts like you would wring out a towel and expect the codes to fall all over the floor like water” (Personal communication, 2020).

The introductory examples above were kept purposely simple and direct. But depending on the researcher’s academic discipline, ontological and epistemological orientations, theoretical and conceptual frameworks, research questions, generated data, and even the choice of coding method itself, some codes can attribute more evocative meanings to data. In the excerpt below, a mother describes her teenage son’s troubled school years. The codes derive from the perspective of middle and junior high school years as a difficult period for most youth. They are abstract and evocative in nature; this is an example of Concept Coding (profiled in Chapter 6):

Code example 1.3

¹ My son, Barry, went through a really tough time about, probably started the end of fifth grade and went into sixth grade. ² When he was growing up young in school he was a people-pleaser and his teachers loved him to death. ³ Two boys in particular that he chose to try to emulate, wouldn't, were not very good for him. ⁴ They were very critical of him, they put him down all the time, and he kind of just took that and really kind of internalized it, I think, for a long time. ⁵ In that time period, in the fifth grade, early sixth grade, they really just kind of shunned him all together, and so his network as he knew it was gone.

¹ MIDDLE-SCHOOL HELL² TEACHER'S PET³ BAD INFLUENCES⁴ TWEEN ANGST⁵ THE LOST BOY

Note that when we reflect on a passage of data to decipher its core meaning, we are *decoding*; when we determine its appropriate code and label it, we are *encoding*. For ease of reference throughout this manual, *coding* will be the sole term used. Simply understand that coding is the transitional process between data generation and more extensive data analysis and synthesis. Coding, in fact, is compatible with the way the human mind naturally thinks. “This ability to digest large amounts of information by breaking it into smaller pieces is how our brains turn information into knowledge” (Duhigg, 2016, pp. 245–6).

Coding for Patterns

A pattern is repetitive, regular, or consistent occurrences of action/data that appear more than twice. “At a basic level, pattern concerns the relation between unity and multiplicity. A pattern suggests a multiplicity of elements gathered into the unity of a particular arrangement” (Stenner, 2014, p. 136). As qualitative researchers, we seek patterns as somewhat stable indicators of humans’ ways of living and working to render the world “more comprehensible, predictable and tractable” (p. 143). They become more trustworthy evidence for our findings since patterns demonstrate habits, salience, and significance in people’s daily lives. They help confirm our descriptions of people’s “five Rs”: routines, rituals, rules, roles, and relationships (Saldaña & Omasta, 2022). Discerning these trends is a way to solidify our observations into concrete instances of meaning.

Bernard (2018) succinctly states that analysis is “the search for patterns in data and for ideas that help explain why those patterns are there in the first place” (p. 355). In the examples presented thus far, each unit of data was assigned its own unique code, due primarily to the short length of the excerpts. In larger and complete data sets, you will find that several to many of the same codes will be used repeatedly throughout. This is both natural and deliberate—natural because there are mostly repetitive patterns of action and consistencies in human affairs, and deliberate because one of the coder’s primary goals is to find these repetitive patterns of action and consistencies in human

affairs as documented in the data. In the example below, note how the same Process Code (a word or phrase which captures action) is used twice during this small unit of elementary school classroom activity:

Code example 1.4

¹ Mrs. Jackson rises from her desk and announces, “OK, you guys, let’s get lined up for lunch. Row One.” Five children seated in the first row of desks rise and walk to the classroom door. Some of the seated children talk to each other. ² Mrs. Jackson looks at them and says, “No talking, save it for the cafeteria. ³ Row Two.” Five children seated in the second row of desks rise and walk to the children already standing in line.

¹ LINING UP FOR LUNCH

² MANAGING BEHAVIOR

³ LINING UP FOR LUNCH

Another way the above passage could be coded is to acknowledge that MANAGING BEHAVIOR is not a separate action or an interruption of the routine that disrupts the flow of LINING UP FOR LUNCH, but to interpret that MANAGING BEHAVIOR is an embedded or interconnected part of the larger social scheme that composes LINING UP FOR LUNCH. The coding might appear thusly, using a method called Simultaneous Coding which applies two or more codes within a single datum (profiled in Chapter 5):

Code example 1.5

^{1a} Mrs. Jackson rises from her desk and announces, “OK, you guys, let’s get lined up for lunch. Row One.” Five children seated in the first row of desks rise and walk to the classroom door. Some of the seated children talk to each other. ^{1b} Mrs. Jackson looks at them and says, “No talking, save it for the cafeteria. Row Two.” Five children seated in the second row of desks rise and walk to the children already standing in line.

^{1a} LINING UP FOR LUNCH
↓
^{1b} MANAGING
BEHAVIOR
↓

Take note of some important caveats when it comes to understanding patterns and regularity: idiosyncrasy is a pattern (Saldaña, 2003, pp. 118–22), there can be patterned variation in data (Agar, 1996, p. 10), and patterns can be demi-regular—that is, not perfectly consistent but generally consistent throughout the data. Sometimes we code and categorize data by what participants talk about. They may all share with you their personal perceptions of school experiences, for example, but their individual experiences and value, attitude, and belief systems about education may vary greatly from being bored and disengaged to being enthusiastic and intrinsically motivated. When you search for patterns in coded data to categorize them, understand that sometimes you may group things together not just because they are exactly alike or very much alike, but because they might also share something in common—even if, paradoxically, that commonality consists of differences, variation, or contradiction (Bouvier & Rasmussen, 2022, p. 88).

For example, each one of us may hold a strong opinion about who should lead our country. The fact that we each have an individual opinion about that issue is what we have in common. As for who we each believe should lead the country, that is where the differences and variations occur. Acknowledge that a confounding property of category construction in qualitative inquiry is that data cannot always be precisely and discretely bounded; they are within “fuzzy” boundaries at best (Tesch, 1990, pp. 135–8). That is why Simultaneous Coding is an option, when needed.

Hatch (2023) offers that you think of patterns not just as stable regularities but as varying forms. A pattern can be characterized by:

- similarity (things happen in the same way)
- difference (they happen in predictably different ways)
- frequency (they happen often or seldom)
- sequence (they happen in a certain order)
- correspondence (they happen in relation to other activities or events)
- causation (one appears to cause another) (p. 194)

Alvesson and Kärreman (2011) caution that a narrow focus on codification for pattern making with qualitative data can oversimplify the analytic process and hamper rich theory development: “Incoherencies, paradoxes, ambiguities, processes, and the like are certainly key aspects of social reality and worth exploring—both as topics in their own right and as a way of getting beyond premature pattern-fixing and the reproduction of taken-for-granted assumptions about specific patterns” (p. 42). Their advice is well taken, for it is not always the regularities of life but its anomalies and deviations that intrigue us, that stimulate us to question and to investigate why they exist concurrently with the mundane and normative—a process called “abductive analysis” (Tavory & Timmermans, 2014). As you code, construct patterns, certainly, but do not let those one or two codes that do not quite seem to fit anywhere frustrate you or stall your analytic work. Use these fragments as stimuli for deep reflection on the reason for their existence, if not their purpose, in the larger social scheme of things. But also take comfort in Braun and Clarke’s (2006) realist perspective that “a pattern in data is rarely, if ever, going to be 100% complete and non-contradicted” (p. 95).

Coding Lenses, Filters, and Angles

Coding requires that you wear your researcher’s analytic lens. But how you perceive and interpret what is happening in the data depends on what type of filter covers that lens and from which angle you view the phenomenon. For example, consider the following statement from an older white rancher living north along the Arizona border with Mexico: “There’s just no place in this country for illegal immigrants. Round them up and send those criminals back to where they came from.” One researcher, a grounded theorist using In Vivo Coding to keep the data rooted in the participant’s own language, might code the datum this way:

Code example 1.6

¹ There's just no place in this country for illegal immigrants.
Round them up and send those criminals back to where they
came from.

¹ "NO PLACE"

A second researcher, a rural ethnographer employing Descriptive Coding to document and categorize the breadth of opinions stated by multiple participants on both sides of the border, might code the same datum this way:

Code example 1.7

¹ There's just no place in this country for illegal immigrants.
Round them up and send those criminals back to where they
came from.

¹ IMMIGRATION ISSUES

And a third researcher, a critical race theorist employing Values Coding to capture and label participants' subjective perspectives about contentious social issues, may code the exact same datum this way:

Code example 1.8

¹ There's just no place in this country for illegal immigrants.
Round them up and send those criminals back to where they
came from.

¹ XENOPHOBIA

The same datum above was coded three different ways by three different data analysts employing three different qualitative research methodologies driving the investigation of three different central research questions designed from three different theoretical perspectives. In other words, each researcher perceives and interprets the data with different and specific sets of lenses, filters, and angles.

The collection of coding methods in this manual offers a repertoire of possible lenses, filters, and angles to consider and apply to your approaches to qualitative inquiry. Researchers' eyes are like camera lenses. The way they perceive social life can be influenced and affected by the investigator's own significant demographic attributes such as gender, age, race/ethnicity, sexual orientation, socioeconomic class, and/or occupation (Behar & Gordon, 1995; Saldaña, 2015; Stanfield & Dennis, 1993). Lenses might also consist of the particular research methodology or disciplinary approach employed for a study (educational, sociological, psychological, grounded theory, phenomenology, etc.).

Cameras also have filters covering their lenses that let certain wavelengths in and keep others out. The filters that cover a researcher's lens might consist of a set of personal values, attitudes, and beliefs about the world, formed by his or her unique personal biography, learned experiences, and individual thinking patterns. Researchers' identities as human beings will influence and affect what they observe in the field site since we tend to interpret others' experiences based on our own. Filters also consist of

particular theoretical perspectives within a discipline—for example, feminist, critical, emancipatory.

Cameras are also placed at particular angles, suggesting not just panoramic and close-up views, but also a researcher's relational positionality and standpoint as a peripheral, active, complete, and/or covert member (Adler & Adler, 1987; Emerson, Fretz, & Shaw, 2011), in addition to the researcher's interpretations of social action he or she sees and hears at the micro- (local and particular), meso- (cultural, national, or mid-range), and/or macro- (conceptual, global, or universal) levels of life. Researchers zoom in and out throughout the course of observations to get varied perspectives of the social scene, varying from insider to outsider, from intimate to distant, or from emotionally invested to neutrally detached.

Merriam (1998) states that “our analysis and interpretation—our study's findings—will reflect the constructs, concepts, language, models, and theories that structured the study in the first place” (p. 48). And it is not just your methodological approach to qualitative inquiry (e.g., case study, ethnography, phenomenology) and ontological and epistemological foundations that influence and affect your coding decisions (Creswell & Poth, 2018; Mason, 2002). Sipe and Ghiso (2004), in their revealing narrative about coding dilemmas, note that “All coding is a judgment call” since we bring “our subjectivities, our personalities, our predispositions, [and] our quirks” to the process (pp. 482–3). Like the characters in director Akira Kurosawa's classic film *Rashomon*, multiple realities exist because we each perceive and interpret social life from different points of view.

Coding as a Heuristic

The majority of qualitative researchers will code their data both during and after collection as an analytic tactic, for coding *is* analysis. Differing perspectives, however, attest that “Coding and analysis are not synonymous, though coding is a crucial aspect of analysis” (Basit, 2003, p. 145). Coding is a heuristic, an exploratory problem-solving technique without specific formulas or algorithms to follow. Codes are significant phrases that “make meaning ..., they are something that happens that make something [else] happen” (Fuller & Goriunova, 2014, p. 168); they initiate a rigorous and evocative analysis and interpretation for a report. Plus, coding is not just labeling, it is *linking*: “It leads you from the data to the idea and from the idea to all the data pertaining to that idea” (Richards & Morse, 2013, p. 154).

Coding is a cyclical act. Rarely is the first pass or first cycle of coding data perfectly attempted. The second cycle (and possibly the third, fourth, etc.) of recoding further manages, filters, highlights, and focuses the salient features of the qualitative data record for generating categories, themes, and concepts, grasping meaning, and/or building theory. Coffey and Atkinson (1996) propose that “coding is usually a mixture of data [summation] and data complication, breaking the data apart in analytically relevant ways in order to lead toward further questions about the data” (pp. 29–31). Locke, Feldman, and Golden-Biddle (2015) conceptualize the coding process as a “live” rather

than inert action. Coding “is organic in which coding, codes and data shape each other; they are interdependent and inseparable” (p. 373). Once a code is applied to a datum during first cycle analysis, it is not a fixed representation but a dynamic and malleable process “through which to consider and interact with further observations and ideas” (p. 6). Indeed, heuristic fluidity is necessary to prioritize insightful qualitative analytic discovery over mere mechanistic validation. To some, coding is comparable to an introspective journey or a rite of passage, a liminal period of mental and spiritual contemplation in order to achieve higher-level meanings.

Dey (1999) critically posits that “With categories we impute meanings, with coding we compute them” (p. 95). To some, *code* is a “dirty four-letter word.” A few research methodologists perceive a code as mere shorthand or an abbreviation for the more important category yet to be discovered. Unfortunately, some use the terms code, category, theme, and concept interchangeably when they are, in fact, four separate components of data analysis. I advocate that qualitative codes are essence-capturing and essential elements of the research story that, when clustered together according to similarity and regularity (i.e., a pattern), actively facilitate the development of categories and thus analysis of their connections. Ultimately, I like one of Charmaz’s (2014) metaphors for the process when she states that coding “generates the bones of your analysis. ... [I]ntegration will assemble those bones into a working skeleton” (p. 113).

Codifying and Categorizing

To codify is to arrange things in a systematic order, to make something part of a system or classification, to categorize. When you apply and reapply codes to qualitative data, you are codifying—a process that permits data to be divided, grouped, reorganized, and linked in order to consolidate meaning and develop explanation (Grbich, 2013). Coding enables you to organize and group similarly coded data into categories or “families” because they share some characteristic—the beginning of a pattern. You use classification reasoning plus your tacit and intuitive senses to determine which data “look alike” and “feel alike” when grouping them together (Lincoln & Guba, 1985, p. 347).

From Codes to Categories

Synthesis combines different things in order to form a new whole, and it is the primary heuristic for transitioning from coding to categorizing (and from categorizing to other analytic syntheses). A quantitative parallel is determining the mean or average of a set of numbers. You take, say, 10 different test scores varying in range from a perfect score of 100 to the lowest achieved score of 62. Add each score (totaling 872), divide by the number of scores (10), and the mean is calculated (87.2). You have synthesized 10 different test scores into one new whole or symbol of meaning. But does qualitative data analysis have a heuristic equivalent? No and yes.

How do you “average” 10 different but somewhat comparable words and phrases to arrive at a category? There is no qualitative algorithm or formula that adds up the codes and calculates their mean. But there are methods for synthesizing the collective, not to arrive at a reduced answer but to move toward *consolidated meaning*. That meaning may take the symbolic form of a category, theme, concept, assertion, or proposition, or set in motion a new line of investigation, hypothesis, interpretive thought, or the crystallization of a new theory. I blithely offer: “Quantitative analysis calculates the mean, but qualitative analysis calculates meaning.”

For example, in Harry, Sturges, and Klingner’s (2005) ethnographic study on the overrepresentation of youth of color in special education programs, data initially coded as classroom MATERIALS, COMPUTERS, and TEXTBOOKS were categorized under the major heading, **Resources**. As their study continued, another major category was constructed labeled **Teacher Skills** with the subcategories **Instructional Skills** and **Management Skills**. The codes subsumed under these subcategories—part of the overall hierarchical “coding scheme” (Silver & Lewins, 2014)—were:

Category: Teacher Skills

Subcategory 1: Instructional Skills

Code: PEDAGOGICAL

Code: SOCIO-EMOTIONAL

Code: STYLE/PERSONAL EXPRESSION

Code: TECHNICAL

Subcategory 2: Management Skills

Code: BEHAVIORIST TECHNIQUES

Code: GROUP MANAGEMENT

Code: SOCIO-EMOTIONAL

Code: STYLE (overlaps with instructional style)

Code: UNWRITTEN CURRICULUM

As another example, Eastman’s (2012) ethnographic study, “Rebel Manhood: The Hegemonic Masculinity of the Southern Rock Music Revival,” employed grounded theory’s Initial, Focused, and Axial Coding to develop categories of “identity work strategies [Southern US] rebel men use to compensate for their lack of the economic resources and authority higher class men use to signify their hegemonic manhood” (p. 195). One major conceptual category was **Rebel Manhood as Protest Masculinity**, with its three subcategories:

- ***Protesting Education and Rejecting Cultural Capital***
- ***Protesting Work and Career***
- ***Protesting Economic Authority***

Another conceptual category was **Compensatory Rebel Manhood Acts**, with its three subcategories:

- ***Drinking Alcohol and Violence***
- ***Drug Use***
- ***Protesting Authority and Risk Taking***

Maykut and Morehouse (1994) refine each category by developing a rule for inclusion in the form of a propositional statement, coupled with sample data. For example, if a category in a case study is labeled **Physical Health**, its rule for inclusion might read:

Physical Health: The participant shares matters related to physical health such as wellness, medication, pain, etc.: “I’m on 25 milligrams of amitriptyline each night”; “I’ve lost ten pounds on this new diet.”

Categories might also evolve as conceptual processes rather than descriptive topics such as:

Inequity: Participants perceive unfair treatment directed toward themselves and favoritism directed toward others: “I’ve been working here for over 25 years and some newcomers are making higher salaries than me.”

The categories’ statements are then compared to each other to discern possible relationships to create an *outcome proposition* based on their combination.

There are exceptions to every rule, however. Harding (2019) promotes that codes can be placed in more than one category or subcategory if you feel that the multiple classification is justified. This tactic is incompatible with analytic methods such as Domain and Taxonomic Coding and analysis (see Chapter 10), but quite logical within the paradigm of “fuzzy sets,” which acknowledges that categories are not always discretely bounded but oftentimes overlap (Bazeley, 2018). I prefer to keep my codes singular and clustered into their most appropriate categories for analysis. Yet it is good to know that, *if and when needed*, a code can get subsumed into more than one category. Too much of this, though, may suggest that the codes and/or the categories may not be as clearly defined as necessary, for there is a big difference between “fuzzy” category boundaries and “vague” ones.

Overall, the purposes of categorizing are to identify how an array of codes belongs in certain groups, to sort codes according to defining attributes, to compare one categorical group to another, and to condense the complexity of the data corpus—the body of data (Freeman, 2017, p. 25).

Recoding and Recategorizing

Rarely will anyone get coding right the first time. Qualitative inquiry demands meticulous attention to language, action, and images, and deep reflection on the researcher-constructed

patterns and meanings of human experience. Recoding can occur with a more attuned perspective using first cycle methods again, while second cycle methods describe those processes that might be employed during the second (and third and possibly fourth, etc.) review of data. Punch (2009), researching childhoods in Bolivia, describes how her codes, categories, and themes (as she defines them) developed and subdivided during her ethnographic fieldwork and concurrent data analysis:

[O]ne of my initial large codes was “home”. Everything relating to life at home was coded under this category and then subdivided into three themes: gender roles; child/adult work roles in the household; power and discipline. On reading through this latter category, I realized not only did it concern adult power over children, but also children’s strategies for counteracting adult power. After reorganizing these two sub-sections, I decided to split up the theme of children’s strategies into different types: avoidance strategies, coping strategies, and negotiation strategies. Finally, on browsing again through the sub-theme of negotiation strategies I found that I could further subdivide it into child-parent negotiations and sibling negotiations. These data then formed the basis for structuring my findings on children’s lives at home. (pp. 94–5)

If you extract the coding scheme described in Punch’s narrative above, and transform it into an outline format or a hierarchical tree, it might appear thusly:

- I HOME
 - A **Gender Roles**
 - B **Child/Adult Work Roles in the Household**
 - C **Power and Discipline**
 - 1 ***Adult Power over Children***
 - 2 ***Children’s Strategies for Counteracting Adult Power***
 - a Avoidance Strategies
 - b Coping Strategies
 - c Negotiation Strategies
 - i *Child/Parent Negotiations*
 - ii *Sibling Negotiations*

As you code and recode, expect—or rather, strive for—your codes and categories to become more refined and, depending on your methodological approach, more conceptual and abstract. Some of your first cycle codes may be later subsumed by other codes, relabeled, or dropped altogether. As you progress toward second cycle coding, you might rearrange and reclassify coded data into different and even new categories. Abbott (2004) cleverly likens the analysis process to “decorating a room; you try it, step back, move a few things, step back again, try a serious reorganization, and so on” (p. 215).

For example, I observed and interviewed fourth- and fifth-grade children to learn the ways they hurt and oppress each other (Saldaña, 2005b). This was preparatory fieldwork

before an action research project that attempted to empower children with strategies, learned through improvised dramatic simulations and role-playing, for dealing with bullying in the school environment. I initially categorized their responses into **Physical** and **Verbal** forms of oppression. Some of the codes that fell under these categories were:

Category: Physical Oppression

Code: PUSHING

Code: FIGHTING

Code: SCRATCHING

Category: Verbal Oppression

Code: NAME-CALLING

Code: THREATENING

Code: LAUGHING AT

As coding continued, I observed that a few oppressions were a combination of both physical and verbal actions. For example, a child can exclude others physically from a game by pushing them away, accompanied with a verbal statement such as “You can’t play with us.” Hence, a third major category was developed: **Physical and Verbal Oppression**.

As the study continued, more data were generated through other methods, and gender differences in children’s perceptions and enactment of oppression became strikingly apparent. To young participants, oppression was not about the body and voice, as the research team labeled; it was about, in their terms, “force” and “feelings.” The three initial categories were eventually condensed to two during second cycle coding, and renamed based on what seemed to resonate with gender-based observations. The new categories and a few sample codes and rearranged subcodes included:

Category: Oppression through Physical Force (primarily but not exclusively by boys)

Code: FIGHTING

Subcode: SCRATCHING

Subcode: PUSHING

Subcode: PUNCHING

Category: Oppression through Hurting Others’ Feelings (primarily but not exclusively by girls)

Code: PUTTING DOWN

Subcode: NAME-CALLING

Subcode: TEASING

Subcode: TRASH TALKING

Also note how the subcodes themselves are specific, observable types of *realistic* actions related to the codes, while the two major categories labeled **Oppression** are more *conceptual* and *abstract* in nature.

See the Domain and Taxonomic Coding profile in Chapter 10 for an extended discussion of this case, the Initial and Focused Coding examples in Chapters 6 and 13 respectively, and the techniques of Code Mapping and Code Landscaping in Chapter 12 to learn how a series of codes gets categorized.

From Codes and Categories to Theory

Some categories may contain clusters of coded data that merit further refinement into subcategories. And when you compare major categories to each other and consolidate them in various ways, you transcend the particular reality of your data and progress toward the thematic, conceptual, and theoretical—that is, more abstract and general analytic meanings. As a very basic process, codifying usually follows the ideal and streamlined scheme illustrated in Figure 1.1.

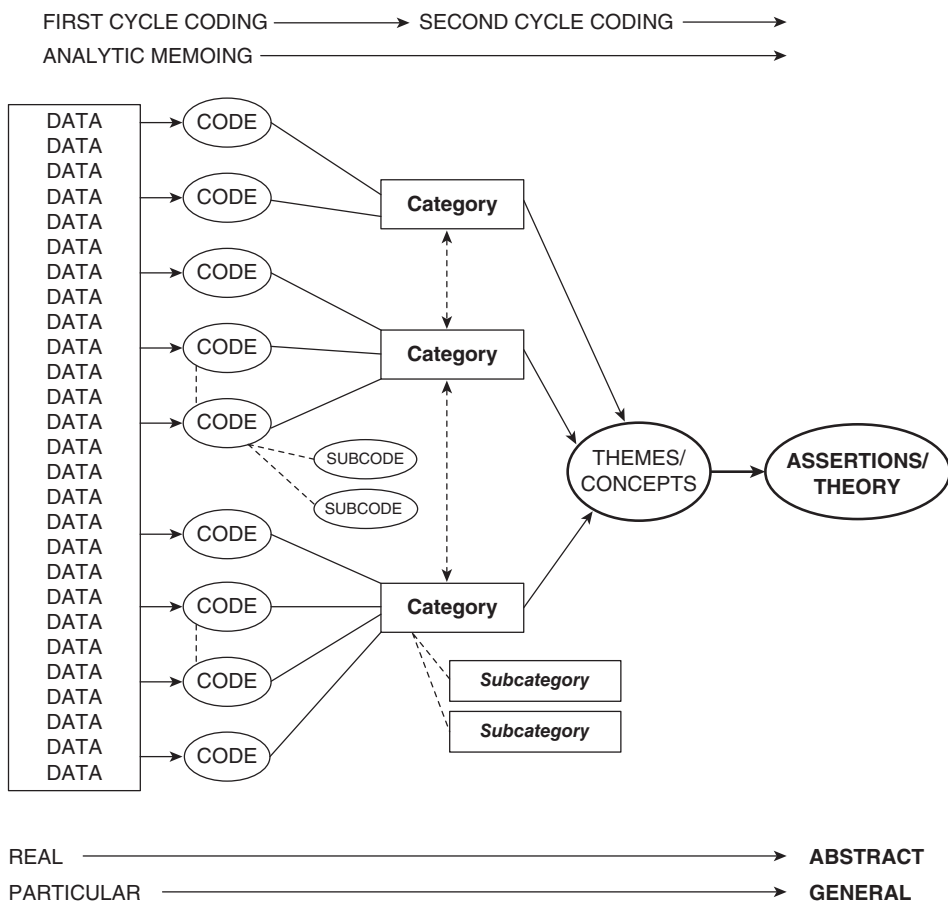


Figure 1.1 A streamlined codes-to-theory model for qualitative inquiry

The large rectangular bin on the left side of the figure represents the data corpus consisting of all generated empirical materials from fieldwork such as interview transcripts, field notes, video recordings, photographs, and so on. The data are then coded (and subcoded, as needed) accordingly and clustered into groups that look alike and feel alike to create categories (and subcategories, as needed), the culminating act of first cycle coding. There are only three category bins included in the diagram for space reasons, but a qualitative study's analysis can have several more. Also notice that the category bins are connected with dashed arrows, suggesting an interrelationship among them rather than isolated, self-standing units.

For some qualitative studies, the analysis can end with category development if this answers the study's primary research questions and provides a satisfactory set of findings. But some studies and methodologies need to go further through second cycle coding processes, and so proceeding to concepts or themes as prerequisites for higher-level theories and/or assertions takes the analytic work to more summative and condensed forms of meaning. The entire journey begins with a particular set of data that culminates in the possibility of generalizable or transferable findings to other contexts. Naeem et al. (2023) illustrate how Figure 1.1's model was adapted to display the thematic analysis results of their study on panic buying during the COVID-19 pandemic. Figure 1.2 progresses from qualitative data on the left side, to the culmination of three primary themes on the right side.

Keep in mind that the actual act of reaching theory is much more complex than illustrated in Figure 1.1. Richards and Morse (2013) clarify that “categorizing is how we get ‘up’ from the diversity of data to the shapes of the data, the sorts of things represented. *Concepts* are how we get up to more general, higher-level, and more abstract constructs” (p. 173). Our ability to show how these themes and concepts systematically interrelate leads toward the development of theory (Corbin & Strauss, 2015), though Layder (1998) contends that pre-established sociological theories can inform, if not drive, the initial coding process itself. The development of an original theory is not always a necessary outcome for qualitative inquiry, but acknowledge that pre-existing theories drive the entire research enterprise, whether you are aware of them or not.

In the example above of children's forms of oppression, I constructed two major categories from the study: **Oppression through Physical Force**, and **Oppression through Hurting Others' Feelings**. So, what major themes or concepts can be developed from these categories? An obvious theme we noticed was that, in later childhood, PEER OPPRESSION IS GENDERED. One higher-level concept we constructed—an attempt to progress from the real to the abstract—was *child stigma*, based on the observation that children frequently label those who are perceived different in various ways “weird,” and thus resort to oppressive actions (Goffman, 1963). We could not, in confidence, formulate a formal theory from this study due to the limited amount of fieldwork time in just a few classrooms at one school. But a key assertion (Erickson, 1986)—a statement that proposes a summative, interpretive observation of the local contexts of a study—that we developed and put forth was:

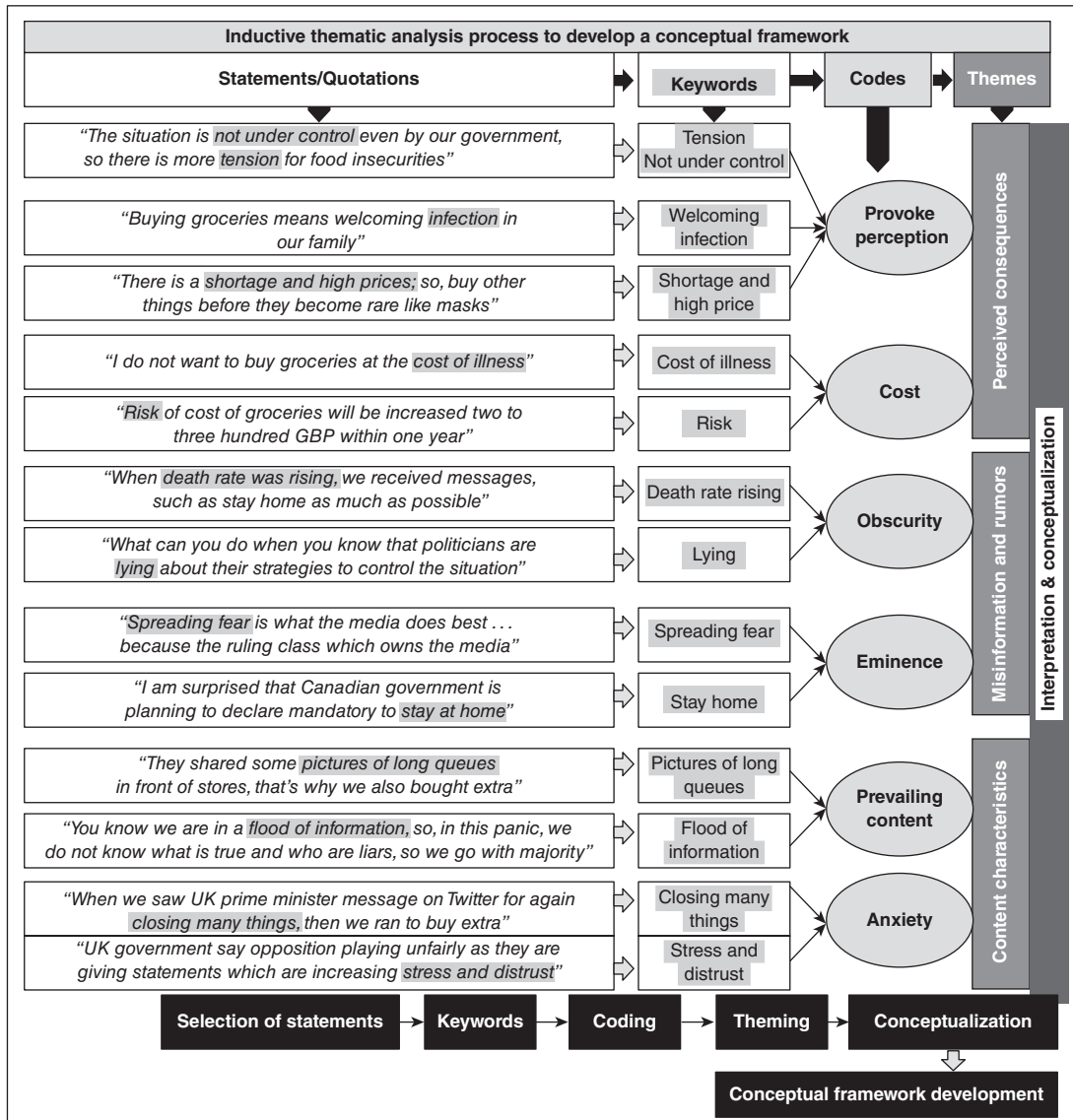


Figure 1.2 From data to themes (Source: Naeem et al., 2023, p. 3)

To artist and activist Augusto Boal, adult participation in theatre for social change is "rehearsal for the revolution." With ages 9–11 children, however, their participation in theatre for social change seems more like an "audition" for preadolescent social interaction. The key assertion of this study is: *Theatre for social change overtly reveals the interpersonal social systems and power hierarchies within an elementary school classroom microculture, because the original dramatic simulations children create authentically reflect their statuses and stigmas.* It diagnostically shows which children are leaders, followers, resisters, and targets; who is influential and who is ignored;

which children may continue to assert dominance in later grade levels; and which children may succumb to those with more authority in later grade levels. (adapted from Saldaña, 2005b, p. 131)

This key assertion, like a theory, attempts to progress from the particular to the general by inferring *transfer*—that what was observed in just six elementary school classrooms at one particular site over the course of one semester may also be observed in comparable elementary school classrooms in other locations. This assertion also progresses from the particular to the general by *predicting patterns* of what may be observed and what may happen in similar present and future contexts.

The Differences Between Codes and Themes

Several qualitative research studies report that the analyst employed “thematic coding.” That, to me, is misleading terminology. A theme can be an *outcome* of coding, categorization, or analytic reflection, but it is not something that is, in itself, coded (that is why there is no “theme coding” method in this manual, but there are references to thematic analysis and a chapter titled “Methods of Themeing the Data”). A datum is initially and, when needed, secondarily coded to discern and label its content and meaning according to the needs of the inquiry. Rossman and Rallis (2003) explain the differences: “think of a category as a *word or phrase* describing some segment of your data that is explicit, whereas a theme is a *phrase or sentence* describing more subtle and tacit processes” (p. 282, emphasis added). As an example, SECURITY can be a category, but DENIAL MEANS A FALSE SENSE OF SECURITY can be a theme.

Qualitative researchers are not algorithmic automatons. If we are carefully reading and reviewing the data before and as we formally code them, we cannot help but notice a theme or two (or a pattern or concept) here and there. Make a note of it in an analytic memo (see Chapter 3) when it happens, for it can sometimes guide your continued coding processes. A set of themes is a good thing to construct from analysis, but at the beginning cycles there are other rich discoveries to be made with specific coding methods that explore phenomena such as participant processes, emotions, and value systems.

The four components of analysis discussed thus far have parallels to the ways our minds naturally work. Our brains

synthesize vast amounts of information into symbolic summary (codes); we make sense of the world by noticing repetition and formulating regularity through cognitive schemata and scripts (patterns); we cluster similar things together through comparison and contrast to formulate bins of stored knowledge (categories, concepts); and we imprint key learnings from extended experiences by creating proverb-like narrative memories (themes, assertions, theories). (adapted from Saldaña, 2015, p. 11)

Necessary Personal Attributes for Coding

Aside from cognitive skills such as logical and critical thinking, there are seven personal attributes all qualitative researchers should possess, particularly for coding processes.

First, you need to be *organized*. This is not a gift that some people have and others do not. Organization is a set of disciplined skills that can be learned and cultivated as habits. A small-scale qualitative study's word count of data will range in the tens and sometimes hundreds of thousands of words. The multiple codes you generate will need a tightly organized framework for qualitative analysis; in fact, organization *is* analysis. And despite the electronic filing systems of hard drives, AI/ChatGPT, and dedicated qualitative data analysis software (discussed in Chapter 2), you will still encounter and manipulate many pages of paper in qualitative work. Date and label all incoming data and keep multiple digital and hard copies as backups.

Second, you need to exercise *perseverance*. Virtually every writer of qualitative research methods literature remarks that coding data is challenging and time-consuming. Some writers also declare how tedious and frustrating it can be. Take breaks from your work when you need to, of course—this will keep you refreshed and alert. But cultivate a personal work ethic and create an environment and schedule that enable you to sustain extended periods of time with analytic tasks requiring your full concentration.

Third, you need to be able to *deal with ambiguity*. Coding and codifying are not precise sciences with specific algorithms or procedures to follow. Yes, occasionally answers may suddenly and serendipitously crystallize out of nowhere, and AI/ChatGPT can generate descriptive summaries of your data corpus in just a few seconds. But at other times, a piece of the analytic puzzle may be missing for days or weeks or even months. Rich ideas need time to formulate in your mind, so have trust and faith in yourself that these may develop in due time. But remember that you can accelerate the process through analytic memo writing.

Fourth, you need to exercise *flexibility*. Coding is a cyclical process that requires you to recode not just once but twice (and sometimes even more). Virtually no one gets it right the first time. If you notice that your initial methods choices may not be working for you or not delivering the answers you need, be flexible with your approach and try a modified or different method altogether. Virtually all researcher-developed coding schemes are never fixed from the beginning—they evolve as analysis progresses.

Fifth, you need to be *creative*. There is a lot of art to social science. Ethnographer Michael H. Agar (1996) asserts that the early stages of analysis depend on “a little bit of data and a lot of right brain” (p. 46). We generally advocate that qualitative researchers remain close to and deeply rooted in their data, but every code and category you construct or select is a choice from a wide range of possible options. Creativity also means the ability to think visually, to think symbolically, to think in metaphors, and to think of as many ways as possible to approach a problem (Saldaña, 2015). Creativity is essential for your data generation, data analysis, and even for your final written report.

Sixth, you need to be *rigorously ethical*. Honesty is perhaps another way to describe this, but I deliberately choose the phrase because it implies that you will always be:

rigorously ethical with your participants and treat them with respect; rigorously ethical with your data and not ignore or delete those seemingly problematic passages of text; and rigorously ethical with your analysis by maintaining a sense of scholarly integrity and working diligently toward the final outcomes.

The seventh and arguably most important skill you need for coding is an *extensive vocabulary*. The precision of quantitative research rests with statistical accuracy. In qualitative research, our precision rests with our word choices. For example, there are subtle interpretive differences between something that “may,” “could,” “can,” “probably,” “possibly,” and “seemingly” happen; and a wide interpretive difference between something that happens “frequently,” “usually,” and “often” (Hakel, 2009). An unabridged dictionary and thesaurus become vital reference tools to find just the right words for your codes, categories, themes, concepts, assertions, propositions, and theories. Explore the origins of key words in an unabridged dictionary to find surprising new meanings (e.g., did you know that the root word of *hypocrite* is “actor”?). A thesaurus review of a key word chosen as a code or category may introduce you to an even better—and more precise—word for your analysis.

For an applied introduction to the cognitive skills and personal attributes necessary for coding and qualitative data analysis, see Saldaña (2015) and the exercises and simulations in this book’s companion website.

Critiques Against Coding

There have been some legitimate critiques against coding, some of them philosophical and some of them methodological. Yet when I hear these criticisms, I am inclined to think that my colleagues’ reservations originate from what used to be earlier, positivist approaches to coding—mechanical and technical procedures that did indeed make the enterprise sheer drudgery and the outcomes often little more than topic-driven lists. Below are some of the most frequent criticisms I have heard against coding and my responses to those (mis)perceptions.

Coding is reductionist. Coding is what you perceive it to be. If you see it as reductionist, then that is what it will be for you. But recall that my definition of coding approaches the analytic act as one that assigns rich symbolic meanings through essence-capturing and/or evocative attributes to data. The 36 coding profiles in this book present an array of methods. And by design or necessity, a few are indeed meant to assist with nothing more complicated than descriptive, topical indexing, and even fewer are formulaic and prescriptive because that is how their developers intended them. But most of these methods can generate discovery of the participant’s voice, processes, emotions, motivations, values, attitudes, beliefs, judgments, conflicts, microcultures, identities, worldview, life course patterns, etc. These are not reductionist outcomes but multidimensional facets about the people we study. Qualitative methodologist Sally Thorne offers in her workshops that coding is an effective way to organize and manage data, but the analyst must move beyond these technical structures toward intellectual meaning-making. As I state later in this book, coding is in service to thinking.

Coding tries to be objective. Somewhat and no. This could become an extended discussion about the ontological, epistemological, and methodological assumptions of inquiry, but let me bypass those in favor of a quick response. A selected number of quantitative researchers believe that, without evidentiary statistics, qualitative analyses are unreliable and invalid. And by default this perspective also negates the qualitative paradigm altogether. Intercoder agreement in team coding (see Chapter 2) does indeed seem as if “objectivity” is the driving analytic force due to the need for two or more researchers to independently corroborate on the meaning of each datum. But in reality, the process is not so much objectivity as it is simply achieving similar results between two or more people.

For the individual researcher, assigning symbolic meanings (i.e., codes) to data is an act of personal signature. And since we each most likely perceive the social world differently, we will therefore experience it differently, interpret it differently, document it differently, code it differently, analyze it differently, and write about it differently. Objectivity has always been an ideal yet contrived and virtually impossible goal to achieve in quantitative research. Qualitative researchers do not claim to be objective because the notion is a false god. Coding requires insightful, emotionally intelligent, and interpretive judgment.

Coding is mechanistic, instrumentalist, and distances you from your data.

If you are doing your job right as a qualitative researcher, nothing could be further from the truth. Coding well requires that you reflect deeply on the meanings of each and every datum. Coding well requires that you read, reread, and reread yet again as you code, recode, and recode yet again. Coding well leads to total immersion in your data corpus with the outcome exponential and intimate familiarity with its details, subtleties, and nuances. When you can quote verbatim by memory what a participant said from your data corpus and remember its accompanying code, I do not understand how that action has “distanced” you from your work.

Coding is nothing more than counting. In traditional, quantitatively-driven content analysis studies, counting the number of times a particular set of codes occurs is indeed an important measure to assess the frequency of items or phenomena. But one of the caveats I propose later in this manual is that frequency of occurrence is not necessarily an indicator of significance. The analytic approaches for most of these coding methods do not ask you to count; they ask you to ponder, to scrutinize, to interrogate, to experiment, to feel, to empathize, to sympathize, to speculate, to assess, to organize, to pattern, to categorize, to connect, to integrate, to synthesize, to reflect, to hypothesize, to assert, to conceptualize, to abstract, and—if you are really good—to theorize. Counting is easy; thinking is hard work.

Coding is “dangerous,” “violent,” and “destructive.” I have difficulty understanding why extremist words such as these have been chosen to describe the act of coding. I associate these words with natural disasters, crime, and war, not with qualitative data analysis. I feel these monikers are sensationalist hyperbole in a culture of fear, and I question their legitimacy and accuracy for describing their critics’ intended concerns. In other

words, these are, to me, poor word choices for an argument. And poor word choosers make bad coders.

Coding is an outdated method for qualitative data analysis. Coding qualitative data has over a half-century of use, and a substantive track record in many disciplines and scholarly publications. The technology needed for the enterprise has most certainly evolved through time, as have the methodologies and methods. But the core process of coding remains to this day a legitimate option for qualitative researchers. It is a tradition that has endured, not out of mindless adherence to established protocols, but due to its successful utility as a purposeful analytic approach to voluminous amounts of data.

There has been a recent trend in some circles of scholarship to discount and “refuse” coding outright as an old-fashioned, postpositivist, colonial approach that does not harmonize with more theory-based analytics (e.g., inspired by Derrida, Deleuze, Foucault). Qualitative methodologist Yvonna S. Lincoln (2015) challenges poststructuralist rejection of coding because it walks away from disciplined, rigorous, and systematic analysis and places too much stock on “long dead French intellectuals” (pp. 203–4).

Coding does not preclude or push theory and theorists aside. You as the analyst can still weave theory into your thinking through analytic memo writing and in the final report itself. Coding is neither a philosophy nor a way of viewing the world; it is simply a heuristic for achieving some sense of clarity about the world from your data and your deep reflections on them.

AI/ChatGPT makes manual coding obsolete. My exploration of qualitative data analysis with ChatGPT-4 suggests that, at this time in the program’s development, I can indeed receive rapid descriptive summaries and even a list of relevant albeit basic categories and themes of the narrative data I submit for processing. But critical scholars of AI/ChatGPT note that the researcher’s experiential knowledge, cognitive input, and critical evaluative skills are essential for substantive analysis and theory development, capacities not yet fully realized by AI (Christou, 2023, p. 2746). Zhang et al. (2023) concur, noting that “The exact nature of ChatGPT’s ‘understanding’ is fundamentally different from human cognition. While it can mimic the process of qualitative analysis, it lacks human intuitions, emotions, and the depth of experience” (p. 22). I perceive ChatGPT outputs as extended narrative “word clouds”—a heuristic that gives the researcher an organized, descriptive, first look “at a glance” overview of the data’s contents, but still requires additional human analysis and interpretation to create more substantive meaning and transferability. Though Morgan (2023) predicts that ChatGPT may one day make coding irrelevant, he advises that the platform is still a “tool” for initial queries needing systematic researcher follow-up (p. 9).

There’s more to data analysis than just coding. I absolutely agree. The 44 analytic approaches documented in Appendix B alone support this perception. This manual advocates that coding is a heuristic—a method of discovery that hopefully stimulates your *thinking* about the data you have been given and have generated. And

in case you forgot two very important principles stated at the beginning of this chapter, I repeat them for emphasis:

- Coding is just *one* way of analyzing qualitative data, not *the* way.
- There are times when coding the data is absolutely necessary, and times when it is most inappropriate for the study at hand.

The next chapter discusses some of the fundamental methods and techniques of coding qualitative data.