The Coding Manual for Qualitative Researchers

4E

Johnny Saldaña

online resources

The Coding Manual for Qualitative Researchers

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

The Coding Manual for Qualitative Researchers

Johnny Saldaña

SAGE

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

Los Angeles | London | New Delhi
Singapore | Washington DC

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

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

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

SAGE Publications Asia-Pacific Pte Ltd
3 Church Street
#10-04 Samsung Hub
Singapore 049483

This edition published 2021

Editor: Jai Seaman
Assistant editor: Charlotte Bush
Production editor: Ian Antcliff
Copyeditor: Neville Hankins
Proofreader: Emily Ayers
Indexer: Adam Pozner
Marketing manager: Ben Griffin-Sherwood
Cover design: Shaun Mercier
Typeset by: C&M Digitals (P) Ltd, Chennai, India
Printed in the UK

Library of Congress Control Number: 2020941916

British Library Cataloguing in Publication data

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

ISBN 978-1-5297-3175-0
ISBN 978-1-5297-3174-3 (pbk)

For Kathy Charmaz – colleague, mentor, and friend

Contents

Detailed Contents

List of Figures

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, Ian Antcliff, Ben Griffin-Sherwood and Charlotte Bush of SAGE Publications UK, Neville Hankins, Emily Ayers, Adam Pozner, and the anonymous manuscript reviewers for their pre-publication guidance throughout this project;

My qualitative research instructors at Arizona State University: Tom Barone, Mary Lee Smith, 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, Ray Maietta, Paul Mihas, Jeff M. Petruzzelli, Ronald J. Chenail, Leo A. Mallette, Patricia Leavy, Steven A. Harvey, Coleman A. Jennings, Kakali Bhattacharya, Joe Norris, Liora Bresler, Hannah Shakespeare, Eric D. Teman, 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 3.3 and 3.4 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 3.5 permission;

David K. Woods (Transana), Elizabeth Jost (VERBI Software/MAXQDA), Eli Lieber (Dedoose/SocioCultural Research Consultants, LLC), and Daniel Turner (Quirkos Software) for their assistance with CAQDAS screenshots and permissions;

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

Hannah Shakespeare of Routledge, Melanie Birks, Jane Mills, AltaMira Press, Oxford University Press, and *The Qualitative Report* for permissions to reprint extended excerpts from selected works;

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

Jonothan 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 the right coding tools for the right analytic job.

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); 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 of *Qualitative Data Analysis: A Methods Sourcebook* (4th ed., Sage Publications, 2020); co-author with Matt Omasta of *Qualitative Research: Analyzing Life* (Sage Publications, 2018); 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.

His methods works have been cited and referenced in more than 18,000 research studies conducted in over 130 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 most popular journal article, "Blue-Collar Qualitative Research: A Rant" (*Qualitative Inquiry*, 2014), has been downloaded by over 4,000 readers, according to ResearchGate.

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 Qualitative Research Special Interest Group, 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 Fourth Edition

The fourth edition of *The Coding Manual for Qualitative Researchers* features a reformatting of its contents into 15 chapters for easier sectional reference. Two new first cycle coding methods, Metaphor Coding and Themeing the Data: Categorically, join the 33 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 Holistic Coding, Provisional Coding, Themeing the Data: Categorically, and Themeing the Data: Phenomenologically. I have enhanced the discussions in the coding methods profiles' Applications and Analysis sections. Additional examples of how to code and possible analytic outcomes have been added.

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 analytic methods.

Google Scholar and ResearchGate updates, conferences, and e-mail correspondence with students and colleagues have informed me how the first three editions of *The Coding Manual for Qualitative Researchers* have been utilized in a variety of studies internationally in disciplines such as K–12 and higher education, the fine arts, government and social services, business, technology and social media, communication, sport, human development, interpersonal relationships, the social sciences, health care, and medicine. I am both humbled and honored that *The Coding Manual for Qualitative Researchers* and its cousin, *Qualitative Data Analysis: A Methods Sourcebook* (third and fourth editions co-authored with the late Matthew B. Miles and A. Michael Huberman), have been cited and referenced in more than 18,000 research studies conducted in over 130 countries.

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 and coding. But the amounts of books and e-resources on the subject have

increased exponentially over the past decade, 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 fourth 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

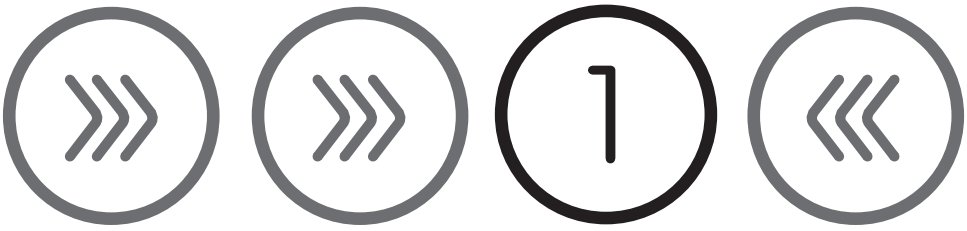
A companion website accompanies this manual with additional resources such as SAGE journal article links, exercises and activities, classroom ancillary materials, and expanded data sets for coding and qualitative data analytic skills development: https://study.sagepub.com/saldanacoding4e

“Wake up and smell the coding.”

(SAGE Campus Marketing)

PART ONE

CODING FOUNDATIONS



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 collection and analytic 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 manual serves as a reference to supplement existing works in qualitative research design and fieldwork. It focuses exclusively on codes and coding and how they play a role in the qualitative data analytic process. For newcomers to qualitative inquiry it presents a repertoire of coding methods in illustrated detail. Additional information and extended discussion of the methods can be found in most of the cited sources.

The manual does not subscribe to any one specific 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 for clarity or flexibility's sake. 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. For the record, here are mine, from *Fundamentals of Qualitative Research*:

> Qualitative research has evolved into a multidisciplinary enterprise, ranging from social science to art form. Yet many instructors of research methods vary in their allegiances, preferences, and prescriptions for how to conduct fieldwork and how to write about it. I myself take a pragmatic stance toward human inquiry and leave myself open to choosing

> the right tool for the right job. Sometimes a poem says it best; sometimes a data matrix does. Sometimes words say it best; sometimes numbers do. The more well versed you are in the field's eclectic methods of investigation, the better your ability to understand the diverse patterns and complex meanings of social life. (Saldaña, 2011b, pp. 177–8)

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.

This manual supplements introductory works in qualitative research because most textbooks 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 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 References.

This book serves primarily as a manual—a collection of useful information. It is not necessarily meant to be read from cover to cover, but it certainly can be if you wish to acquaint yourself with all 35 coding methods' profiles and their analytic possibilities. If you choose to review all the contents, read selected sections at a time, not all of them in one sitting, otherwise it can overwhelm you. If you scan the manual to explore which coding method(s) might be appropriate for your particular study, read the profiles' Description and Applications sections to determine whether further reading of the profile is merited, or check the glossary in Appendix A. I doubt you will use every coding method included in this 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?

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 data 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, academic and fictional literature, and so on. The portion of data coded during 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 collection 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, 2018).

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, even 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*, *It's a Wonderful Life*, "Do Not Go Gentle Into That Good Night"), 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 is a Descriptive Code, which summarizes the primary topic of the excerpt that follows the same superscript number:

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:

Code example 1.2

[1] He cares about me. He has never told me but he does. [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.	[1] SENSE OF SELF-WORTH [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 (2017) notes that such analytic work does not diminish but "value adds" to the research story (p. 103).

Also, some may wonder, "Where did those three codes 'come from'?" They came from my thinking, my background knowledge, 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, 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:

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 HELL

[2] TEACHER'S PET

[3] BAD INFLUENCES

[4] TWEEN ANGST

[5] THE LOST BOY

Note that when we reflect on a passage of data to decipher its core meaning, we are *de*coding; when we determine its appropriate code and label it, we are *en*coding. For ease of reference throughout this manual, *coding* will be the sole term used. Simply understand that coding is the transitional process between data collection 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, 2018). 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

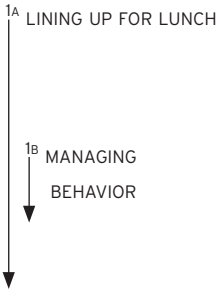
[2] MANAGING BEHAVIOR

[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):

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.



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.

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 (2002) offers that you think of patterns not just as stable regularities but as varying forms. A pattern can be characterized by:

- similarity (things happen 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. 155).

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 100% complete and non-contradicted.

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 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, 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 subjective perspectives, 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 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 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, 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 (Saldaña & Omasta, 2018, pp. 34–6).

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 for a children's literacy study, 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 *Rashōmon*, 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 (from the Greek, *heuriskein*, meaning "to discover"), 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 and 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.

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* and *category* interchangeably when they are, in fact, two 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, interpretive thought, or the crystallization of a new theory. I blithely offer: "Quantitative analysis calculates the mean. Qualitative analysis calculates meaning."

For example, in Harry, Sturges, and Klingner's (2005) ethnographic study of the overrepresentation of minorities 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, 2013, p. 351). 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 (Freeman, 2017, p. 25).

Recoding and recategorizing

Rarely will anyone get coding right the first time. Qualitative inquiry demands meticulous attention to language 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 sub-divide 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 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 collected 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; it was about "force" and "feelings." The three initial categories were eventually reduced 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. As a very basic process, codifying usually follows the ideal and streamlined scheme illustrated in Figure 1.1.

Keep in mind that the actual act of reaching theory is much more complex than illustrated. 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

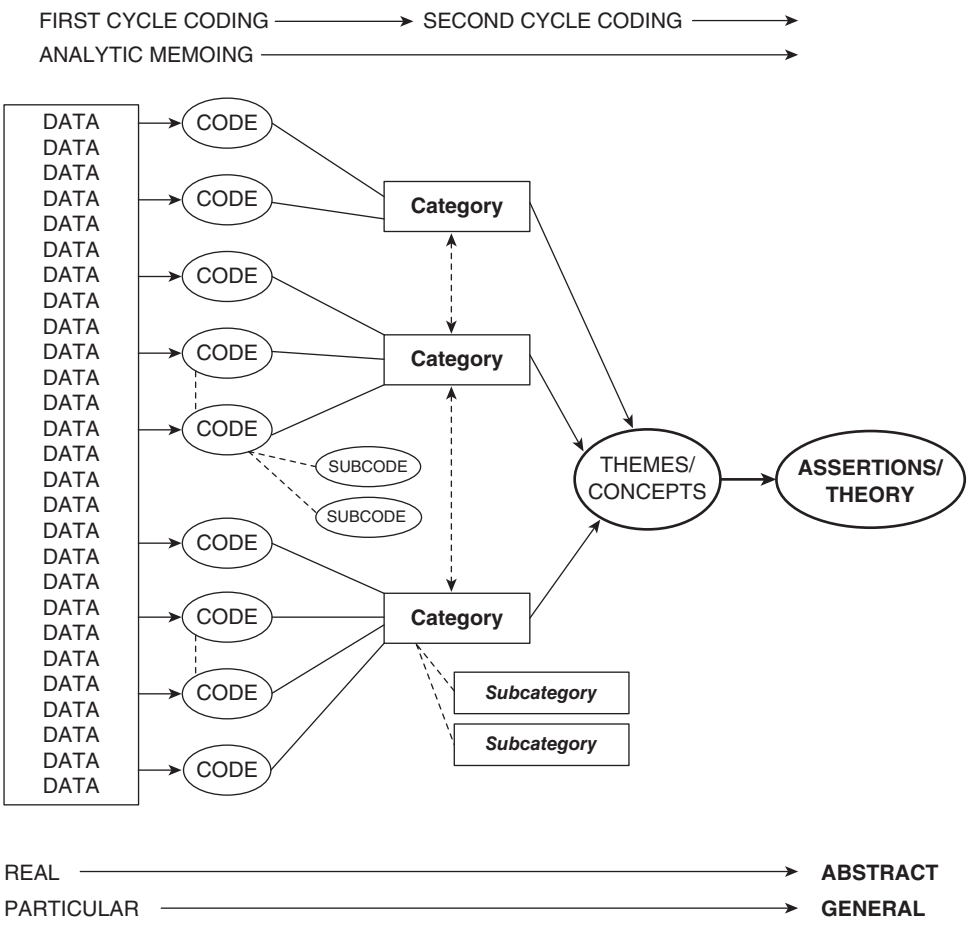


Figure 1.1 A streamlined codes-to-theory model for qualitative inquiry

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 (Mason, 2002).

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 the classrooms. 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: