Beyond the Numbers

Qualitative Research
from Words to Evidence

VIVEK H. PATIL

MARGIN OF ERROR MEDIA

Table of contents

Copyright 3

Preface 5

Introduction: How Research Works 9

I Asking and Listening 17

1 What Words Can Show 19

2 The Interview 29

II Participants and Trust 39

3 Whom to Talk To 41

4 Consent and Trust 51

III From Text to Themes 61

5 Coding 63

6 From Codes to Themes 73

IV Evidence and Reporting 83

7 Checking the Analysis 85

8 Text Beyond Interviews 97

Index of Tools and Data 109

Beyond the Numbers

For the people who agreed to be interviewed, and for the researchers who owe them an accurate account.

Copyright

Beyond the Numbers: Qualitative Research from Words to Evidence

© 2027 Vivek H. Patil, Ph.D. All rights reserved.

Published by Margin of Error Media LLC, Spokane, Washington.

No part of this book may be reproduced, distributed, or transmitted in any form without the prior written permission of the publisher. Brief quotations in reviews or scholarly work, with full attribution, are the exception.

For permissions, classroom adoption, or bulk purchase, write to patilv@gmail.com.

First edition, 2027.

ISBN, paperback: [ISBN-PENDING]

ISBN, Kindle: [ISBN-PENDING]

The Text in This Book

This book uses two kinds of text, and each passage is labeled.

Synthetic text. The interview transcripts in Chapters 1 through 7 were written for this book. No real person said these words. The 12 housing interviews were modeled on themes in published research on housing instability. The participants' names are invented. Every synthetic passage is marked.

Real text. The consumer complaints in Chapters 7 and 8 are real. Consumers sent them to the Consumer Financial Protection Bureau, and each consumer agreed to have the complaint published. The bureau removes personal details before it publishes a complaint. The complaints are in the public domain. Each one in this book has its complaint number, and a reader can look it up in the bureau's public database. This book has masked most company names in the text. The bureau does not verify the facts that consumers report, and neither does this book.

A Note on AI

I wrote this book with a language model as a working partner. I used Claude, made by Anthropic, to draft and revise passages, to check reasoning, and to build the tools that go with the chapters. The model also wrote the first drafts of the synthetic interviews, which I then edited.

In Chapter 7, the model appears as a subject of study. It assigned codes to 40 real complaints, and the chapter measures how far a reader's codes agree with it. Those codes are labeled as the model's and were not corrected.

The judgment in the book is mine. Where the book is wrong, the fault is mine, and I would like to hear about it at the address above.

Preface

Some questions cannot be answered with a count.

A city knows how many eviction cases were filed last year. It does not know what a parent told a child on the night the notice came. A hospital knows how many patients missed a follow-up visit. It does not know why. The answers to questions like these are in what people say, and a researcher has to go and ask.

Qualitative research is the study of words, and of the meaning people give to their experience. Its data are interviews, group discussions, written comments, and documents. Its findings are patterns in that material, supported by quotations that a reader can inspect.

Many people assume that this kind of research is informal. It has methods, and the methods can be done well or badly. A researcher chooses whom to interview and can explain the choice. The researcher labels what was said by written rules and tests each pattern against the cases that do not fit. This book teaches those methods.

How the Book Is Organized

An introduction comes first. It explains what research is, the steps it follows, and where this book fits. Readers who are new to research should start there.

Eight chapters follow in four parts. The parts follow the order of the work.

1. **Asking and Listening.** Chapter 1 explains when qualitative research fits. Chapter 2 covers the interview.
2. **Participants and Trust.** Chapter 3 covers how to choose participants. Chapter 4 covers consent and the protection of participants.
3. **From Text to Themes.** Chapter 5 covers coding. Chapter 6 covers how codes become themes.
4. **Evidence and Reporting.** Chapter 7 covers how to check an analysis, including one a language model helped with. Chapter 8 covers text that already exists and the written report.

Two Bodies of Text

Qualitative data are hard to share. Real interview transcripts contain details that can identify the people who spoke, and most cannot be published. This book handles the problem in two ways.

The 12 housing interviews are **synthetic**. They were written for this book and are marked wherever they appear. They are short, and they were written to contain the kinds of patterns that real interviews on this topic contain. They let every reader work on the same material. They are not evidence about housing, and the book does not treat them as evidence.

The 1,000 consumer complaints are **real**. People wrote them in their own words and agreed to their publication. They show what unedited text looks like, with its misspellings, its anger, and its gaps.

What Each Chapter Contains

Each chapter opens with a case. Key terms are in bold where they first appear, and a glossary at the end of the chapter defines them. Two kinds of boxes appear in each chapter.

- An **Ethics Moment** describes a choice that affects participants.
- An **AI Reality Check** describes what a language model can and cannot do for that step.

Each chapter also has a section called **Work With the Data**. It appears at the point where the idea is taught and is part of the chapter. In it you use a tool on the chapter's text, with numbered steps and a note on what to look for. In the online edition the tool sits on the page. Readers of the print edition can open it in any web browser.

Each chapter ends with exercises at three levels. *Check Your Understanding* covers concepts. *Apply It* asks you to do a piece of the work. *Think Deeper* asks for judgment.

For Instructors

The eight chapters can fill a short course or form one part of a longer course. An instructor guide is available on request. It has a session plan for each chapter, an activity for each tool, student worksheets, and answers to the exercises. Instructors can write to patilv@gmail.com.

The Companion Volume

Before the Numbers covers research that produces numbers, from the research question to the dataset. The two books can be read alone or together. Chapter 1 of this book explains how the two kinds of research fit together.

Introduction: How Research Works

This book assumes no background in research. This introduction explains what research is, the steps it follows, and where the eight chapters fit. Readers who have conducted a study before can skim it. Readers who are new to research should start here.

Research Starts With a Decision

People in organizations make decisions every day without knowing all the facts. A store owner decides whether to stay open later. A school board decides which program to cut. A county decides where to place a new clinic.

Some of these decisions are small and easy to reverse. Others cost a great deal and affect many people. For the second kind, a guess is a weak basis. The decision is only as good as the evidence behind it.

One example runs through this introduction. It is invented for this book. A community health clinic finds that patients miss 22 percent of their appointments. Each missed visit costs the clinic money and leaves another patient waiting. The director must decide what to do.

She has several ideas. The clinic could send text reminders, add evening hours, or offer rides. Each idea costs money, and she cannot afford all three. Her appointment records show how many visits are missed and by whom. They do not show why. She needs to hear from the patients.

What Research Is

Research is a planned and careful way of producing information to support a decision. Three features separate it from ordinary fact-finding.

- **It is planned.** The researcher decides what to learn and how to learn it before collecting anything.
- **It is systematic.** The same care is applied to every case, so one vivid story does not outweigh the rest.

- **It is open to checking.** The researcher records what was done, so another person can judge the work.

The clinic director could ask the front desk staff why patients miss visits. That is useful, and it is not yet research. The staff hear from the patients who call to cancel. They do not hear from the patients who never call. A planned study would reach both groups.

Many people think of talking with a few people as informal. It becomes research when the three features above are present. The researcher can say why these people were chosen, what they were asked, and how their words were turned into findings.

Research that serves one decision for one organization is **applied research**. The clinic study is an example. Research that aims to add to general knowledge, with no single decision in view, is **basic research**. Both kinds use the same methods, and this book serves both.

The Research Process

Research is a sequence of steps. Textbooks number the steps in slightly different ways, and the logic is the same in all of them. This book uses eight steps, adapted from Churchill and Iacobucci (2015) and from Zikmund and colleagues (2013).

1. **Formulate the problem.** State the decision and the question that the research must answer.
2. **Estimate the value of the information.** Decide whether the answer is worth the cost of getting it.
3. **Choose the research design.** Decide what kind of study fits the question.
4. **Design the method and the forms.** Decide how to gather the data. In an interview study, the main form is the interview guide.
5. **Plan the sample.** Decide whom to study and how to choose them.
6. **Collect the data.** Conduct the interviews and record them.
7. **Prepare, analyze, and interpret the data.** Turn the recordings into text, label the text, and look for patterns.
8. **Write the report.** Tell the decision-maker what was found and how far it can be trusted.

Each step limits the steps that follow. A leading question in step 4 produces answers that no analysis in step 7 can repair. A sample that left out one kind of person in step 5 cannot be completed in step 8.

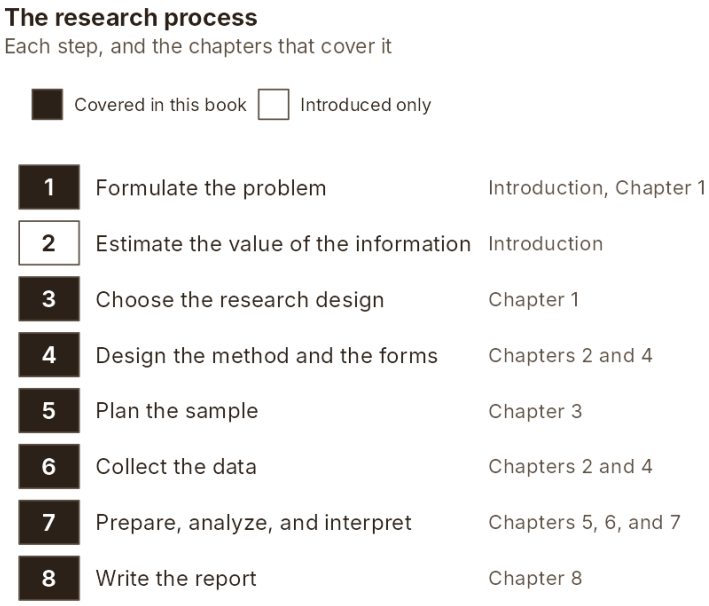


FIGURE 1. *The eight steps of the research process and where this book covers each one.*

In qualitative research the steps overlap more than the list suggests. A researcher often begins to analyze after the first few interviews. What she learns can change whom she interviews next and what she asks. This is a planned part of the method, and Chapter 3 explains how to plan for it.

Figure 1 shows the eight steps and where this book covers each one.

From a Decision to a Research Question

The first step is the one most often rushed. A decision-maker usually arrives with a symptom. The clinic's symptom is the 22 percent of visits that are missed. A symptom is a sign that something needs attention. It does not say what the cause is.

The researcher and the decision-maker work from the symptom toward a question that a study can answer. Two statements come out of that work.

A **decision statement** says what the decision-maker has to choose. A **research objective** says what information would help with that choice. Table 1 shows both for the clinic.

Table 1. *A decision statement and three research objectives for the clinic example.*

Decision statement	Research objectives
Should the clinic spend its budget on text reminders, evening hours, or rides?	Identify the reasons patients give for missing a visit. Describe what happens on the day of a missed visit, in the patient's own account. Learn which kinds of help patients say they would use.

A useful test for an objective is to ask what the decision-maker would do differently with the answer in hand. If no answer would change the decision, the objective can be dropped.

Is the Research Worth Doing

Research costs time and money. Before starting, the researcher asks whether the information is worth its cost. Four considerations help.

- **The stakes.** A costly decision justifies more research than a minor one.
- **The uncertainty.** Research adds the most when the decision-maker does not know which way to go.
- **The reversibility.** A decision that is hard to undo deserves more evidence than one that can be changed next month.
- **The information that already exists.** Someone may have collected the answer already.

Data that someone else collected for another purpose are **secondary data**. Data collected for the study in hand are **primary data**. A careful researcher looks for secondary data first. Text can be secondary data too. Chapter 8 uses 1,000 complaints that consumers sent to a federal agency, and no researcher asked for them.

Three Kinds of Design

A **research design** is the overall plan for a study. The right design depends on how much is already known about the problem. There are three main kinds.

Exploratory research is for a problem that is not yet understood. The researcher does not know which questions to ask. The methods are flexible, and the samples are small. The clinic could begin by interviewing 15 patients who missed visits.

Descriptive research is for a problem that is understood well enough to measure. It answers questions about how many, how often, and who. The clinic could survey 400 patients to learn how common each reason is.

Causal research tests whether one thing produces a change in another. Its main tool is the experiment. The clinic could send reminders to a randomly chosen half of its patients and compare missed visits in the two halves.

Many projects use all three in that order. The interviews show which reasons exist. The survey shows how common each one is. The experiment shows whether a remedy works.

Qualitative research is the usual tool for the first kind. It also has a place after the other two. A survey can show that younger patients miss more visits and leave the reason unexplained. Interviews can then supply the explanation.

Words and Numbers

Research data come in two broad forms. **Qualitative research** collects words, such as what people say in an interview. **Quantitative research** collects numbers, such as answers to a survey with fixed choices. Table 2 compares them.

Table 2. *Qualitative and quantitative research compared.*

	Qualitative research	Quantitative research
Typical question	Why, and how	How many, and how much
Questions asked	Open, with follow-up	Fixed, the same for everyone
Sample	Small, chosen for what each person can explain	Large, chosen to represent a population
Data	Words	Numbers
Analysis	Reading, coding, and comparison	Counting and statistics
Common use	Exploratory research	Descriptive and causal research

Neither form is better than the other. Each answers a different kind of question. A study that uses both is a **mixed methods** study.

The two forms also support different claims. Fifteen interviews can show that a reason exists and how it works in a person's life. They cannot show what share of all patients have that reason. Chapter 1 returns to this limit, because most misuse of qualitative research comes from ignoring it.

This book covers qualitative research. Its companion, *Before the Numbers*, covers quantitative research. Each book can be read without the other.

The People Involved

Three parties take part in most studies, and each has duties to the others.

- **Participants** give their time and their words. They are owed an honest account of the study, a free choice about taking part, and protection of what they share.
- **The researcher** owes the sponsor careful work and an accurate report, including results the sponsor did not hope for.
- **The sponsor** is the person or organization that asked for the study. The sponsor owes the researcher an honest statement of the decision and must not ask for a result fixed in advance.

These duties weigh more in qualitative research. A participant's own words can identify her even when her name is removed. Chapter 4 covers consent and confidentiality, and every chapter has an Ethics Moment.

What This Book Covers

This book follows the research process from the first question to the written report, for studies that collect words.

- **Asking and Listening.** Chapter 1 explains what qualitative research can and cannot show. Chapter 2 covers the interview.
- **Participants and Trust.** Chapter 3 covers whom to interview and when to stop. Chapter 4 covers consent and confidentiality.
- **From Text to Themes.** Chapter 5 covers coding, which is the labeling of text. Chapter 6 covers how labels become findings.
- **Evidence and Reporting.** Chapter 7 covers how to check an analysis. Chapter 8 covers text that already exists and the written report.

Two topics are left to other books. Surveys, sampling for surveys, and experiments belong to *Before the Numbers*. Statistical tests belong to *Margin of Error*, the statistics book in this series.

Each chapter opens with a note that names the step of the process it covers. Chapter 1 begins with a school district that is losing students and a survey that could not say why.

Key Terms

- **Applied research.** Research that serves a specific decision.
- **Basic research.** Research that aims to add to general knowledge.

- **Causal research.** Research that tests whether one thing produces a change in another.
- **Decision statement.** A statement of what the decision-maker has to choose.
- **Descriptive research.** Research that measures how many, how often, and who.
- **Exploratory research.** Research on a problem that is not yet understood.
- **Mixed methods.** A study that uses qualitative and quantitative research together.
- **Primary data.** Data collected for the study in hand.
- **Qualitative research.** Research that collects words.
- **Quantitative research.** Research that collects numbers.
- **Research.** A planned and careful way of producing information to support a decision.
- **Research design.** The overall plan for a study.
- **Research objective.** A statement of the information that would help with a decision.
- **Secondary data.** Data that someone else collected for another purpose.

Further Reading

- Churchill, G. A., and Iacobucci, D. (2015). *Marketing Research: Methodological Foundations.*
- Zikmund, W. G., Babin, B. J., Carr, J. C., and Griffin, M. (2013). *Business Research Methods* (9th ed.). Cengage.

PART I

Asking and Listening

CHAPTER 1

What Words Can Show

Where this chapter fits. Steps 1 and 3 of the research process are to formulate the problem and to choose the design. The Introduction describes all eight steps. This chapter explains which questions call for qualitative research and what such a study can claim.

1.1 The Checkbox That Said "Other"

Consider a school district that has lost 6 percent of its students in three years. The case is a composite written for this book, and the district is not a real one. Its details are typical of what district staff describe.

The district has good numbers. It knows how many students left each school, in which grade, and in which month. When a family withdraws a child, the office staff fills in a form. The form has a box labeled "reason for leaving," with five choices. Sixty percent of the forms are marked "moved out of district." Thirty percent are marked "other."

The school board asks the superintendent why families are leaving. The numbers do not answer the question. "Moved" does not say why the family moved. "Other" says nothing at all.

A staff member proposes a different approach. She will interview 15 families who left in the past year. She will ask each one to describe the decision in their own words. Three excerpts from those interviews follow. They are synthetic.

> Synthetic "We didn't want to leave. They cut the bus route. I start work at six. There was no way to get him there."

> Synthetic "The rent went up 300 dollars. We found a place we could afford, and it was across the district line. That's the whole reason. We loved that school."

> Synthetic "She has a reading plan. The teacher who ran it left, and nobody picked it up. I asked four times. The other district had a spot."

Each excerpt names a cause that the form could not record. One is transportation. One is the cost of housing. One is a service the school stopped providing. The first and third are things the district can act on. The second is outside its control.

The 15 interviews cannot show what share of all departing families left for each reason. They can show what the reasons are and how families weighed them. They can also show which reasons the district's own form was hiding. This chapter explains what kind of knowledge that is and when a researcher should seek it.

1.2 What Qualitative Research Is

Qualitative research studies how people experience and interpret some part of their lives. Its data are words, and sometimes images and observed behavior. The researcher collects those data through interviews, group discussions, observation, and documents. The researcher then analyzes them by reading closely, labeling what is there, and looking for patterns.

Quantitative research records characteristics as numbers and analyzes them with statistics. Each approach fits some questions better than others. Table 1.1 sets them side by side.

Table 1.1. *Quantitative and qualitative research compared.*

	Quantitative research	Qualitative research
Typical question	How many? How much? Is there a difference?	How? Why? What is it like?
Data	Numbers from surveys, records, and experiments	Words from interviews, discussions, observation, and documents
Participants	Many, often chosen at random	Few, chosen on purpose for what they can explain
Instrument	A fixed questionnaire	A flexible guide, with the researcher asking follow-up questions
Analysis	Statistics	Coding and the search for themes
Typical finding	An estimate with a margin of error	A pattern, with quotations as evidence
Main risk	Measuring the wrong thing with precision	Seeing a pattern the data do not support

Neither approach is the more rigorous one. Each has standards, and each can be done badly. A survey with a biased sample is poor research. So is a set of interviews in which the researcher heard only what she expected to hear.

1.3 Questions That Call for Words

Qualitative research fits four kinds of questions.

Questions about process. How does a family decide to change schools? What steps does a tenant go through after receiving a notice? A survey can record the outcome. An interview can trace the sequence.

Questions about meaning. What does "stable housing" mean to a person who has moved five times? Two people can give the same answer on a survey and mean different things by it.

Questions where the categories are unknown. The district's form offered five reasons for leaving. The interviews produced reasons that were not among the five. When a researcher does not yet know the possible answers, a fixed list of choices will miss some of them.

Questions about people who are hard to count. Some groups do not appear in records or do not respond to surveys. A researcher can sometimes reach them through trusted contacts and a conversation.

Qualitative research is a poor fit for questions of how many and how much. Fifteen interviews cannot show that 40 percent of families left because of bus routes. A report that makes such a claim from 15 interviews has gone beyond its data.

1.4 A Well-Known Example

In 2008 and 2009, the sociologist Matthew Desmond lived in two poor neighborhoods of Milwaukee. He lived first in a trailer park and then in a rooming house. He followed eight families and two landlords through evictions, moves, and court hearings. His book *Evicted* (2016) reports what he observed.

The fieldwork showed him something that court records did not. Many families were forced out with no court case at all. A landlord would tell a tenant to leave, or pay the tenant a small sum to go, or remove the front door. None of these events appeared in any record.

Desmond then designed a survey of about 1,100 Milwaukee renters to count what he had seen. The survey found that informal evictions were about twice as common as formal ones. Court records had been missing most forced moves.

The order of the two studies matters. The fieldwork found a category that no dataset contained. The survey then measured how large it was. Without the fieldwork, the survey would not have included the question.

1.5 What a Qualitative Study Can Claim

A probability sample supports a claim about a population. A qualitative study does not, and it should not try to. Its claims are of other kinds.

A claim that something exists. The district's interviews establish that some families left because a bus route was cut. One clear case is enough to show that a reason exists.

A claim about how something works. Interviews can show the steps that connect a cause to an outcome. A rent increase leads to a search. The search finds a cheaper home across the district line. The child then changes schools.

A claim about range. A well-chosen set of participants can show the variety of experiences in a group, including ones the researcher did not expect.

A claim that may transfer. A reader in another district can judge whether the findings are likely to apply there. The researcher supports that judgment by describing the setting and the participants in detail. Lincoln and Guba (1985) call this property **transferability**. It is the reader, and not the researcher, who decides whether the findings transfer.

A qualitative report should use words that match these claims. "Several families described" and "one pattern in the interviews was" are accurate. "Most families feel" is a claim about a population, and the study cannot support it.

1.6 Five Traditions

Qualitative researchers work in several traditions. Creswell and Poth (2018) describe five. Table 1.2 summarizes them.

Table 1.2. *Five traditions of qualitative research, after Creswell and Poth (2018).*

Tradition	What it studies	Typical data
Narrative research	The life story of one person or a few people	Long interviews, letters, diaries
Phenomenology	What an experience is like for those who have lived it	In-depth interviews with people who share the experience

Tradition	What it studies	Typical data
Grounded theory	A process, with the aim of building a theory from the data	Interviews collected and analyzed in rounds
Ethnography	The shared practices of a group or a place	Long-term observation, with interviews
Case study	One bounded case, such as a program or an organization	Interviews, documents, and observation together

Much applied research in business, health, and public policy does not follow one tradition closely. It uses interviews or focus groups to answer a practical question and analyzes them for themes. This book teaches that general approach. The methods in the chapters ahead are common to all five traditions.

1.7 The Four Sources of Data

Qualitative data come from four main sources.

An **interview** is a guided conversation between a researcher and one participant. Chapter 2 covers it in detail. A **focus group** is a guided discussion among six to ten participants. It shows how people respond to one another's views. **Observation** means watching what people do in a setting and recording it in **field notes**. **Documents** include anything already written, such as letters, case files, public comments, and posts online. Chapter 8 covers documents.

Each source has a blind spot. Interviews record what people say they do, which can differ from what they do. Observation records behavior and not the reasons behind it. Documents were written for a purpose other than research. Researchers often combine sources so that each one can check the others. This practice is called **triangulation**, and Chapter 7 returns to it.

1.8 Work With the Data: Read Two Interviews

This book uses one set of 12 interviews about housing in every chapter. The interviews are synthetic. They were written for this book, and no real person said these words. Read two of them now, before Chapter 2.

The interviews are plain text files, named H01 through H12, at qual.marginoferrorme

1. Read each interview once from start to finish.
2. In each one, mark a statement that a survey question would have been unlikely to record.

3. Write one question you would want to ask the participant next.

Keep your notes. Chapter 5 returns to these interviews for coding.

1.9 Using Words and Numbers Together

A study that combines qualitative and quantitative methods is a **mixed methods** study. Creswell and Plano Clark (2018) describe three basic designs.

1. **Words first, then numbers.** Interviews identify the categories, and a survey measures how common each one is. Desmond's work followed this order. So would the school district, if it rewrote its withdrawal form after the interviews.
2. **Numbers first, then words.** A survey finds a pattern, and interviews explain it. A district that sees withdrawals rise sharply in one school could interview families from that school.
3. **Both at once.** The two kinds of data are collected in the same period and compared. If they agree, confidence in the finding grows. If they disagree, the researcher has a new question.

The choice depends on what is already known. If the categories are unknown, begin with words. If a pattern needs to be explained, begin with numbers.

1.10 The Researcher Is Part of the Study

In a survey, every respondent receives the same questions in the same order. In an interview, the researcher decides what to ask next. In the analysis, the researcher decides what a passage means. The researcher's judgment is part of the method.

This has a cost. Two researchers can read the same transcript and notice different things. A researcher's background, beliefs, and hopes for the study all affect what she hears.

Qualitative research manages this cost in the open. **Reflexivity** is the practice of examining how the researcher's own position affects the study, and of reporting it. A former teacher who interviews parents about schools should say so in the report. She should also consider which answers she may have been ready to hear.

The other safeguards are procedures. They include written rules for labeling the data, a second person who labels the same material, and a record of

each decision. They also include a search for cases that contradict the pattern. Chapters 5 through 7 cover each of these.

Ethics Moment: Whose Account Is It

An interview gives a researcher someone else's words. The researcher then selects, cuts, and arranges them. A participant who said ten things may find one of them in the report, placed next to a claim the participant would not accept. The researcher has a duty to represent each person's account fairly. Keep quotations in context. Do not join parts of two answers into one quotation. When a participant's view does not fit the pattern, report it.

AI Reality Check: Invented Quotations Are Not Data

A language model can write an interview transcript on any topic in seconds. The transcripts in this book were drafted that way, and each one is marked as synthetic. They are teaching material. They are not evidence about housing, and no finding about real people can rest on them. The same limit applies to any text a model produces. If you ask a model what parents think about school bus routes, it will write fluent and plausible answers. No parent said them. Qualitative evidence begins with a real person who said or wrote something. A model cannot supply that step.

1.11 Putting It All Together

Qualitative research studies how people experience and interpret their lives. Its data are words from interviews, focus groups, observation, and documents. It fits questions about process and meaning. It also fits questions where the possible answers are not yet known.

A qualitative study does not estimate how common something is. It can show that something exists, how it works, and what range of experiences a group contains. A reader judges whether the findings transfer to another setting.

Words and numbers can be combined. Interviews can supply the categories that a survey then counts, and interviews can explain a pattern that a survey found. The researcher's judgment is part of the method, and reflexivity and written procedures keep that judgment open to inspection.

1.12 Looking Ahead

The staff member in the school district now has to conduct 15 interviews. What she asks, and how she asks it, will shape everything that follows. Chapter 2 covers the interview. The case is a set of conversations with parents who used a crisis nursery.

1.13 Further Reading and References

- Creswell, J. W., and Poth, C. N. (2018). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (4th ed.). Sage. The source of the five traditions.
- Creswell, J. W., and Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). Sage.
- Desmond, M. (2016). *Evicted: Poverty and Profit in the American City.* Crown. The appendix, "About This Project," describes the fieldwork and the survey.
- Desmond, M., and Shollenberger, T. (2015). Forced displacement from rental housing: Prevalence and neighborhood consequences. *Demography, 52*(5), 1751 to 1772. The survey results on formal and informal evictions.
- Lincoln, Y. S., and Guba, E. G. (1985). *Naturalistic Inquiry.* Sage. The source of the term transferability.
- Tracy, S. J. (2020). *Qualitative Research Methods* (2nd ed.). Wiley-Blackwell. A readable general textbook.

1.14 Key Terms

- **Documents.** Text that already exists, such as letters, case files, and public comments.
- **Field notes.** A researcher's written record of what was observed.
- **Focus group.** A guided discussion among six to ten participants.
- **Interview.** A guided conversation between a researcher and one participant.
- **Mixed methods.** A study that combines qualitative and quantitative methods.
- **Observation.** Watching what people do in a setting and recording it.
- **Qualitative research.** The study of how people experience and interpret some part of their lives, using words as data.

- **Quantitative research.** Research that records characteristics as numbers and analyzes them with statistics.
- **Reflexivity.** Examining and reporting how the researcher's own position affects the study.
- **Transferability.** The degree to which findings may apply in another setting, as judged by the reader.
- **Triangulation.** The use of more than one source or method to check a finding.

1.15 Exercises

Check Your Understanding

1. Give two kinds of questions that qualitative research fits well and one kind that it fits poorly.
2. The district interviewed 15 families. Which of these claims can the interviews support? (a) Some families left because a bus route was cut. (b) Bus routes were the most common reason for leaving. (c) Families weighed the commute against the quality of a program.
3. What did Desmond's fieldwork find that court records had missed? Why did the survey come after the fieldwork?
4. Define reflexivity. Give one example of a researcher's background that should be reported.

Apply It

5. Rewrite the district's withdrawal form. Use the three synthetic excerpts to add at least three new choices for "reason for leaving."
6. For each question, name the mixed methods design that fits best and say why. (a) A clinic's survey shows that patients over 70 miss more appointments. (b) A city wants to survey residents about a new park and does not know which features they care about.
7. Read transcript H05 or H09 in the corpus. Write two sentences that report what the participant said, using wording that matches what one interview can support.

Think Deeper

8. A report based on 12 interviews states, "Most renters feel that landlords do not listen." Rewrite the sentence so that it claims only what the study can support. What was lost, and what was gained?

9. A researcher who once lost her own home plans to interview people facing eviction. How could her experience help the study? How could it distort it? What should she do about both?
10. Choose a number from your own work or studies that you do not fully understand. Describe a small qualitative study that could help explain it. Whom would you talk to, and what would you ask?

CHAPTER 2

The Interview

Where this chapter fits. Step 4 of the research process is to design the method and the forms, and step 6 is to collect the data. This chapter covers the interview guide and the interview.

2.1 Two Ways to Ask the Same Thing

A crisis nursery gives parents a safe place to leave young children for a few hours or days during an emergency. Suppose a nursery wants to learn how parents come to use it. The case is a composite written for this book, and the exchanges below are synthetic.

An evaluator interviews a parent. Here is one way the conversation could go.

> Synthetic
> **Interviewer:** Were you satisfied with the nursery?
> **Parent:** Yes.
> **Interviewer:** Did you feel stressed before you came?
> **Parent:** Yeah, pretty stressed.
> **Interviewer:** Would you use it again?
> **Parent:** Probably.

The evaluator asked three questions and received six words. Each question could be answered with yes or no, and each one suggested its own answer. The transcript contains nothing that a three-item survey would not have recorded.

Here is a second version with the same parent.

> Synthetic
> **Interviewer:** Tell me about the day you first brought your daughter here.
> **Parent:** It was a Tuesday. I had a court date for the custody thing, and my sitter canceled at seven that morning. I called my mom, she was at

> work. I called my neighbor, no answer. I was sitting in the car crying, and I remembered a flyer from the clinic.
> **Interviewer:** You remembered a flyer.
> **Parent:** It was on the wall by the scale. I'd seen it maybe ten times and thought, that's for people in real trouble. And then I was the person in real trouble.
> **Interviewer:** What was it like to make that call?
> **Parent:** Hard. I thought they'd ask a lot of questions. Like, why can't you handle your own kid. They didn't. The woman just said, can you be here by nine.

The second version begins with an invitation to tell a story. The interviewer then repeats a phrase the parent used and waits. The parent explains how she learned of the nursery and why she had not used it before. She also describes what she feared and what happened. None of this was in the first version.

An interview is a method, and it takes skill. This chapter covers how to plan one, what to ask, how to follow up, and how to record what was said.

2.2 Three Kinds of Interview

Interviews differ in how much the questions are fixed in advance.

A **structured interview** uses the same questions, in the same words and order, for every participant. It is a survey that is read aloud. It allows answers to be compared directly and leaves no room to follow an unexpected answer.

An **unstructured interview** has a topic and no fixed questions. The participant leads, and the interviewer follows. It suits early exploration and long-term fieldwork. The interviews that result can be hard to compare with one another.

A **semi-structured interview** uses a prepared list of questions and allows the interviewer to change the order and to ask follow-up questions. It is the most common form in applied research, and it is the form this chapter teaches. Every participant is asked about the same topics, and each one can take the conversation where the fixed questions did not reach.

2.3 The Interview Guide

The prepared list of questions is the **interview guide**. A guide for a one-hour interview usually has six to ten main questions. More than that leaves no time for follow-up.

A guide has four parts.

1. **The opening.** The interviewer explains the purpose, confirms consent, and asks permission to record. Chapter 4 covers consent.
2. **A warm-up question.** The first question is easy and invites a long answer. It shows the participant what kind of conversation this will be.
3. **The main questions.** These follow the research question. They move from general to specific and from less sensitive to more sensitive.
4. **The close.** The interviewer asks whether anything was missed and thanks the participant. "Is there anything I should have asked you and didn't?" often produces the most useful answer of the hour.

Table 2.1 shows a guide for the nursery study.

Table 2.1. *An interview guide for a study of parents who used a crisis nursery.*

Part	Question	Purpose
Warm-up	Tell me a little about your family.	Easy to answer. Establishes who is in the household.
Main 1	Tell me about the day you first used the nursery.	Invites a story. Shows the path to the service.
Main 2	How did you first hear about the nursery?	Identifies sources of information.
Main 3	What did you expect it would be like?	Brings out fears and assumptions.
Main 4	Walk me through what happened when you arrived.	Describes the service as the parent experienced it.
Main 5	What would you have done that day if the nursery had not been there?	Shows the alternatives the parent had.
Main 6	What has been different for you since then, if anything?	Explores effects without assuming any.
Close	Is there anything I should have asked you and didn't?	Catches what the guide missed.

2.4 Writing Good Questions

Four rules cover most of what separates a useful interview question from a weak one.

Ask open questions. An **open question** cannot be answered with yes or no. It begins with "how," "what," or "tell me about." A **closed question** can be answered in a word. Closed questions have a place when the interviewer needs a fact, such as a date. They should be the exception.

Ask for stories and examples. People describe a specific event more accurately than they describe their general habits. "Tell me about the last time that happened" produces more detail than "What usually happens?"

Ask one thing at a time. "How did you hear about us, and what made you decide to come?" is two questions. The participant will answer one and forget the other.

Do not suggest the answer. "That must have been frightening" tells the participant what to feel. "What was that like?" does not.

Table 2.2 shows four weak questions and a revision of each.

TABLE 2.2. *Four weak interview questions and a revision of each.*

Weak question	Problem	Revision
Were the staff helpful?	Closed, and suggests the answer	Tell me about the staff you met.
Don't you think the hours are too short?	Leading	How do the hours work for you?
Why didn't you ask your family for help?	Sounds like blame	Who else did you think about asking that day?
How did you find us and what did you think?	Two questions	How did you first hear about the nursery?

A question that begins with "why" needs care. "Why did you wait so long?" can sound like an accusation, and the participant may defend the choice. "What was going on for you during that time?" asks for the same information with no judgment attached.

2.5 Work With the Data: Interview Guide Checker

Table 2.2 showed four weak questions and a revision of each. The tool checks a whole guide for the same problems. It marks closed questions, leading questions, double questions, and questions that begin with "why."

The tool runs in a web browser at qual.marginoferrormedia.com/tools/guide-checker.html. It needs no account and no software.

1. Press "Check the guide" for the sample guide of seven questions. The tool flags four of them.
2. Rewrite each flagged question in the box. Use the revisions in Table 2.2 as models. Check the guide again.
3. Replace the sample with a guide of five to seven questions for a topic you know. Put one question on each line.
4. Revise until most questions are open and none is flagged for leading.

The tool uses fixed rules. It cannot tell whether your questions are in a sensible order or whether they cover the research question. Read the guide aloud to another person before the first interview.

2.6 Probes

A **probe** is a follow-up that asks the participant to say more. Probes are not written in advance. The interviewer chooses them while listening. Most of the detail in a good interview comes from probes.

Five probes cover most situations.

- **Silence.** Wait three or four seconds after the participant stops. People often continue with something more considered. New interviewers find this the hardest probe to use.
- **The echo.** Repeat the participant's last few words. "You remembered a flyer." It shows that you are listening and invites more.
- **The request.** "Tell me more about that."
- **The example.** "Can you give me an example?"
- **The clarification.** "When you say 'real trouble,' what do you mean?"

The last probe matters most for analysis. Participants use words that carry meaning for them. The interviewer should ask what those words mean and should not assume. "Real trouble" in the exchange above shows how the parent had sorted herself into a group that did not need the service.

2.7 Listening

An interviewer has three tasks at once. The first is to hear what the participant is saying. The second is to decide what to ask next. The third is to keep track of which topics on the guide have been covered. Beginners often attend to the guide and miss what was said.

Three habits help. Bring the guide on one page and learn it well enough that you need only glance at it. Take brief notes on phrases to return to. Let the participant finish, and do not fill a pause with the next question.

The interviewer should talk far less than the participant. In a transcript of a good interview, the interviewer's turns are short, and the participant's turns are long.

The interviewer also has to manage the relationship. **Rapport** is the trust that lets a participant speak freely. It comes from interest and respect. It does not come from agreeing with everything the participant says. An interviewer

who reacts with approval to some answers teaches the participant which answers are wanted.

2.8 Recording and Transcribing

Most interviews are recorded on audio, with the participant's permission. A recording lets the interviewer listen fully and preserves the exact words.

A **transcript** is a written record of the recording. A **verbatim transcript** records every word, including repeated words and unfinished sentences. For most applied research, a verbatim transcript that leaves out "um" and "uh" is detailed enough.

A transcript needs a few conventions, applied the same way throughout. Mark who is speaking. Mark pauses and laughter in brackets where they affect the meaning. Mark words that could not be heard as [inaudible] and do not guess. Replace names of people and places with labels such as [sister] or [employer]. Chapter 4 explains why.

One hour of audio takes four to six hours to transcribe by hand. Automatic transcription is now common and much faster. It makes errors, and the errors are more frequent with accents, crosstalk, and poor audio. Someone has to check every automatic transcript against the recording.

Right after each interview, write a **memo**. A memo is a note from the researcher to herself. Record what stood out, what surprised you, how the participant seemed, and what to ask differently next time. These notes are data. They also show a later reader how the researcher's thinking developed.

2.9 Focus Groups

A **focus group** brings six to ten people together for a guided discussion of 60 to 90 minutes. A **moderator** asks the questions and manages the conversation.

A focus group is not a way to interview eight people at once. Its value is the interaction. Participants respond to one another, disagree, and build on what others say. A group can show which views are shared and which are disputed.

A focus group suits topics that people discuss with others in ordinary life, such as a neighborhood, a product, or a public service. It is a poor choice for sensitive personal topics. A parent may not describe a custody dispute in front of seven strangers. A group can also be dominated by one or two voices, and some participants will agree with the room and keep their own view private.

A moderator needs a shorter guide than an interviewer does. Four to six main questions are enough. The moderator also has two extra tasks. One is to draw out the quiet participants. The other is to keep any one person from taking over.

2.10 Before the First Interview

Test the guide before using it. Conduct one or two **pilot interviews** with people similar to the participants. A pilot shows which questions are unclear, which produce short answers, and how long the interview runs. Revise the guide and test it again if the changes are large.

The guide may also change during the study. If the third participant raises a topic that no question covered, the interviewer can add a question for the remaining interviews. This is a normal part of qualitative research. Record each change and the date it was made.

In Person, by Phone, or by Video

An interview in person gives the most information, because the interviewer can see the participant's reactions. A phone interview costs less and can feel more private for some topics. A video interview falls between the two. Choose the mode that is easiest for the participant. A parent with no car and two young children may prefer a phone call during a nap.

Ethics Moment: When a Participant Becomes Upset

An interview about a hard time can bring that time back. A parent who describes the day she called the nursery may begin to cry. The interviewer's first duty is to the person and not to the data. Stop the questions. Offer a pause, and offer to end the interview. Do not press for more because the account is useful. Before the study begins, prepare a short list of local services that the interviewer can give to any participant who wants one. The consent process should tell participants that they may skip any question and stop at any time.

AI Reality Check: The Transcript Needs a Human Check

Automatic transcription saves many hours. It also changes words without notice. It can turn "I couldn't pay" into "I could pay," and it often

> drops short words such as “not.” It may assign a sentence to the wrong speaker. An analysis built on an unchecked transcript can quote a participant as saying the opposite of what was said. Listen to the recording while you read the transcript and correct it. There is also a privacy question. Before you upload a recording to any service, find out where the audio is stored and who can use it. The consent form should say which service will be used.

2.11 Putting It All Together

A semi-structured interview uses a prepared guide and allows follow-up questions. A guide has an opening, a warm-up, six to ten main questions, and a close. Good questions are open, ask for stories, ask one thing at a time, and do not suggest an answer.

Probes produce most of the detail. The main ones are silence, the echo, the request for more, the request for an example, and the clarification of a term. The interviewer listens more than she talks and writes a memo after each interview.

Interviews are recorded and transcribed, and an automatic transcript has to be checked against the recording. A focus group shows how people respond to one another and suits topics that are not private.

2.12 Looking Ahead

The guide is ready. The next decision is whom to interview. A survey draws a random sample. A qualitative study chooses its participants on purpose, and the reasons for each choice matter. Chapter 3 covers how to choose and how to know when enough people have been interviewed.

2.13 Further Reading and References

- Rubin, H. J., and Rubin, I. S. (2012). *Qualitative Interviewing: The Art of Hearing Data* (3rd ed.). Sage. A practical guide to main questions, probes, and follow-up questions.
- Spradley, J. P. (1979). *The Ethnographic Interview.* Holt, Rinehart and Winston. The source of the “grand tour” question that asks a participant to describe a typical event.

- Weiss, R. S. (1994). *Learning From Strangers: The Art and Method of Qualitative Interview Studies.* Free Press.
- Krueger, R. A., and Casey, M. A. (2015). *Focus Groups: A Practical Guide for Applied Research* (5th ed.). Sage.
- Brinkmann, S., and Kvale, S. (2015). *InterViews: Learning the Craft of Qualitative Research Interviewing* (3rd ed.). Sage.

2.14 Key Terms

- **Closed question.** A question that can be answered with yes, no, or a single word.
- **Focus group.** A guided discussion among six to ten participants.
- **Interview guide.** The prepared list of questions for an interview.
- **Memo.** A note from the researcher to herself about what was observed or thought.
- **Moderator.** The person who leads a focus group.
- **Open question.** A question that cannot be answered with yes or no.
- **Pilot interview.** A trial interview used to test the guide.
- **Probe.** A follow-up that asks the participant to say more.
- **Rapport.** The trust that lets a participant speak freely.
- **Semi-structured interview.** An interview with prepared questions and freedom to change the order and follow up.
- **Structured interview.** An interview with the same questions in the same words and order for everyone.
- **Transcript.** A written record of a recorded interview.
- **Unstructured interview.** An interview with a topic and no fixed questions.
- **Verbatim transcript.** A transcript that records every word.

2.15 Exercises

Check Your Understanding

1. Name the three kinds of interview and state how they differ.
2. List the five probes in this chapter. Which one did the interviewer use in the second exchange with the parent?
3. Why is a focus group a poor choice for a study of custody disputes?
4. Why does an automatic transcript have to be checked against the recording?

Apply It

5. Rewrite each question as an open question that does not suggest an answer. (a) "Was the application process confusing?" (b) "Did you feel judged by the staff?" (c) "Why didn't you call sooner?"
6. Write an interview guide of six main questions for this study. "How do first-year college students decide where to live?" Include a warm-up question and a closing question.
7. Read transcript H03 in the corpus. Find one place where the interviewer used a probe and one place where a probe would have helped. Write the probe you would have used.

Think Deeper

8. An interviewer nods and says "right, exactly" whenever a participant criticizes the landlord. What effect could this have on the interview? What should the interviewer do in its place?
9. During the fifth interview of a study, a participant raises a topic that none of the first four mentioned. Should the interviewer add a question about it for the remaining interviews? What is gained, and what problem does it create?
10. Conduct a 15-minute practice interview with a friend about their last move. Record it with permission. Listen to it and count how many of your questions were closed. Report what you would change.

PART II

Participants and Trust

CHAPTER 3

Whom to Talk To

Where this chapter fits. Step 5 of the research process is to plan the sample. This chapter covers whom to interview, how to reach them, and when to stop.

3.1 Twelve People at a Meeting

A city wants to understand how residents of one neighborhood think about safety. The case is a composite written for this book.

A staff member attends the monthly neighborhood meeting and invites people to be interviewed. Twelve agree. All twelve own their homes, and ten have lived in the neighborhood for more than twenty years. Ten are over age 60. They describe speeding cars, a vacant lot, and teenagers in the park after dark.

The city now has twelve interviews, and they agree with one another. It would be easy to report that "residents are concerned about youth in the park."

The twelve were the people who attend evening meetings. Renters make up half the neighborhood, and none were interviewed. No one under 30 was interviewed, and no parent of a teenager. The teenagers in the park were not asked what they think about safety either. The interviews show how one group of residents sees the neighborhood. They do not show how the neighborhood sees itself.

The problem is not the number twelve. Twelve can be enough. The problem is which twelve. This chapter covers how to choose participants for a qualitative study and how to decide when to stop.

3.2 A Different Logic of Sampling

A survey draws a random sample so that the results can be generalized to a population. Chance decides who is included, and the researcher's judgment is kept out.

A qualitative study works the other way. The researcher chooses each participant for what that person can explain. This is **purposive sampling**. The aim is to learn about the range of experiences and the reasons behind them. Patton (2015) calls the people chosen "information-rich cases."

Purposive sampling is not the same as convenience. A **convenience sample** takes whoever is easiest to reach, as the city did at the meeting. A purposive sample follows a plan, and the researcher can state the reason for each choice.

3.3 Strategies for Choosing

Several strategies are in common use. A study often combines two or three.

Maximum variation. The researcher seeks participants who differ as much as possible on the characteristics that matter. For the safety study, those could be tenure, age, length of residence, and the block a person lives on. A pattern that appears across very different people is stronger evidence than one found in a uniform group.

Criterion sampling. Every participant meets a stated condition. A study of eviction could include only people who received a notice in the past year.

Typical case. The researcher chooses cases that are ordinary for the group, to describe the common experience.

Extreme or deviant case. The researcher chooses unusual cases, such as the one block with no reported incidents. Unusual cases can show what makes the ordinary ones ordinary.

Snowball sampling. Each participant is asked to suggest others. It reaches people who are on no list, and it depends on trust passed from one person to the next. People tend to refer others like themselves, and that is the weakness of the method. A snowball that starts at the neighborhood meeting will stay among the people who attend it. Start from several unconnected people to reduce this.

Theoretical sampling. The researcher analyzes the first interviews and then chooses the next participants to test or extend the ideas that are forming. If early interviews suggest that renters and owners see safety differently, the researcher seeks more renters. This strategy comes from grounded theory (Glaser and Strauss, 1967).

A **key informant** is a person with special knowledge of the setting, such as a school principal, a store owner, or a block captain. Key informants can describe the setting and introduce the researcher to others. They speak from their own position, and their account should not stand in for the residents'.

3.4 A Sampling Plan

A simple table turns a strategy into a plan. List the characteristics that matter down the side and the groups across the top. Then set a target for each cell. Table 3.1 shows a plan for the safety study.

TABLE 3.1. *A sampling plan for a study of neighborhood safety, with a target number of interviews in each cell.*

	Lived here under 5 years	Lived here 5 to 20 years	Lived here over 20 years
Renters	3	2	1
Homeowners	1	2	3
Youth ages 15 to 19	2	2	Not applicable
People who work in the neighborhood	2, of any length of time		

The plan has 18 interviews. It does not match the neighborhood in proportion, and it does not need to. Its purpose is to make sure that each kind of experience is heard. The plan also shows the reader of the final report who was included and why.

The plan can change. If the first interviews show that parents of young children have distinct concerns, the researcher can add a row. Record each change with its reason.

3.5 Reaching People

A plan names the people the study needs. Recruitment is the work of reaching them.

A **gatekeeper** is a person or organization that controls access to a group. A tenant association, a youth center, and a faith leader are examples. A gatekeeper's support can make a study possible. It also shapes who is reached. A youth center can introduce the researcher to the teenagers who attend the center, and not to those who avoid it. Use more than one gatekeeper where possible.

Three practical points apply to most studies.

1. **Go where people are.** Residents who do not attend meetings may be at the laundromat, the bus stop, or the school pickup line. Recruit at different places and at different times of day.
2. **Make it easy to say yes.** Offer times outside working hours. Offer to meet by phone. Provide child care or a ride if the budget allows.
3. **Pay people for their time.** A payment of 25 to 50 dollars for an hour is common. It shows respect, and it lets people take part who cannot afford to give an hour away. Chapter 4 covers the point at which payment becomes pressure.

Keep a record of everyone who was invited, who agreed, and who declined. People who decline are information too. If every renter the researcher approached said no, the report should say so.

3.6 When to Stop

A survey sets its sample size in advance with a formula. A qualitative study has no formula. The usual guide is **saturation**. Saturation is the point at which new interviews add little that is new to the analysis.

The idea can be made concrete. Figure 3.1 uses the 12 synthetic housing interviews from this book. Chapter 5 will code them with 14 codes. The figure shows how many of the 14 codes had appeared by the end of each interview.

The line rises quickly and then flattens. After interview 3, 13 of the 14 codes have appeared. Interviews 4 and 5 add none. Interview 6 adds the last one, and interviews 7 through 12 add none.

The step at interview 6 carries a warning. That participant moved by choice and described the move as a good one. The first five participants had all been forced to move or feared that they would be. Had the study interviewed only renters in difficulty, the line would have stayed flat after interview 3. The researcher could have stopped there and never heard the other kind of account. Saturation is measured within the kinds of people a study has reached. A flat line among similar participants shows that the study has heard that group. It does not show that the study has heard every group.

Research on real studies gives rough guidance on numbers. Guest, Bunce, and Johnson (2006) analyzed 60 interviews and tracked when each code first appeared. Most of the frequent codes had appeared within the first six interviews, and nearly all had appeared by the twelfth. Hennink and Kaiser (2022) reviewed 23 studies that measured saturation. Those studies reached it after 9 to 17 interviews or 4 to 8 focus groups. In most of them, the participants were fairly similar and the aims were narrow.

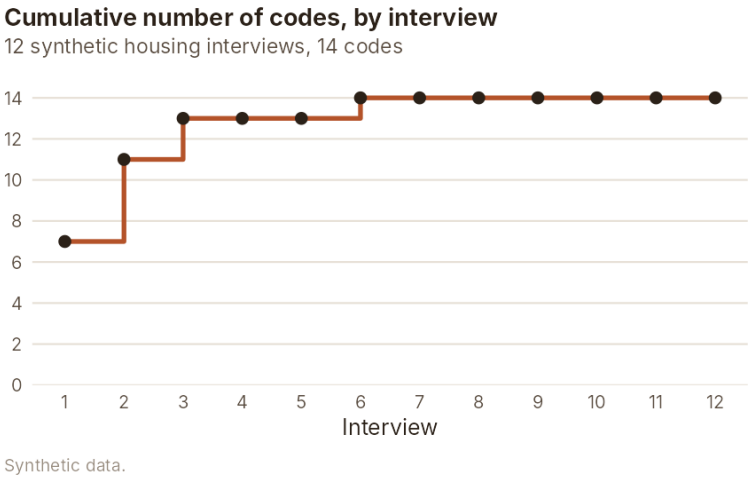


FIGURE 3.1. *Number of distinct codes that had appeared after each of the 12 synthetic housing interviews, in order. Thirteen of the 14 codes appear in the first three interviews. The last one appears in interview 6, which is the first with a participant who moved by choice. Synthetic data.*

These figures are a starting point. A study with a broad question or a varied group needs more interviews. Malterud, Siersma, and Guassora (2016) suggest that researchers ask how much **information power** the sample holds. A sample holds more when the aim is narrow, the participants are well chosen for the aim, and the interviews are rich. Such a sample can be smaller.

3.7 Work With the Data: Saturation Tracker

Saturation is easier to judge with a chart. The 12 synthetic housing interviews were coded with 14 codes, and Chapter 5 shows how. The tool plots how many distinct codes have appeared after each interview.

The tool runs in a web browser at qual.marginoferrormedia.com/tools/saturation-tracker.html. It needs no account and no software.

1. Start with "Original order." Find the interview after which the line stops rising.
2. Choose "Renters in difficulty first." Note where the line goes flat, and which later interviews add the last codes.
3. Choose "Most different people first." Compare how quickly the line rises.

4. Press "Shuffle" several times. The same 12 interviews produce a different curve each time.

The point at which the line goes flat depends on the order of the interviews. A flat line among similar people shows that those people have little more to add. It does not show that every kind of person has been heard. This is why a stopping rule should be tied to the sampling plan.

The file `housing-interviews-index.csv` at qual.marginoferrormedia.com/data lists the 12 participants.

State the Stopping Rule in Advance

Decide before the study how you will judge saturation. One workable rule comes from Francis and colleagues (2010). Conduct a set number of interviews, such as ten. Then continue until three interviews in a row add no new codes. Report the rule and what happened when you applied it.

3.8 Describing the Sample

A qualitative report describes its participants in enough detail for a reader to judge the findings. The description usually takes the form of a table. Table 3.2 shows the participants in the synthetic housing corpus.

TABLE 3.2. *Participants in the 12 synthetic housing interviews. Names are invented. Synthetic data.*

Transcript	Pseudonym	Description
H01	Avery	Renter, age 34, two children, works in a warehouse
H02	Sasha	Renter, age 61, lives alone, retired on a fixed income
H03	Riley	Renter, age 27, restaurant server, lives with a partner
H04	Jordan	Renter, age 45, home health aide, three children
H05	Casey	Renter, age 52, lost housing after a divorce, slept in a car for four months
H06	Quinn	Renter, age 38, moved by choice to a smaller city, works remotely
H07	Robin	Landlord, age 58, owns six units in two older buildings
H08	Akira	Housing caseworker at a nonprofit, age 41, nine years in the job
H09	Devon	Renter, age 23, community college student, shares a house with four others
H10	Noor	Renter, age 36, arrived in the country six years ago, two children, works in food processing
H11	Imani	Renter, age 48, bus driver, cares for a parent with dementia
H12	Rene	Renter, age 31, single parent, moved into subsidized housing eight months ago

The table shows a reader several things at once. Ten of the twelve are renters. One is a landlord and one is a caseworker. The renters range in age from 23 to 61. No participant is a property manager for a large company, and none is a current shelter resident. A reader can see what the corpus covers and what it leaves out.

The table also raises a privacy question. In a small community, a description such as "bus driver, age 48, cares for a parent with dementia" could identify a real person. With real participants, a researcher has to decide how much detail is safe to print. Chapter 4 covers that decision.

Ethics Moment: The Gatekeeper's Influence

A caseworker offers to recruit her own clients for a study of the housing program she works for. The offer is generous, and it creates two problems. Clients may feel that they cannot say no to a person who controls their benefits. They may also feel that they cannot speak freely about the program. The researcher should ask the caseworker only to pass along a flyer. Clients then contact the researcher directly, and the caseworker never learns who took part. The consent form should state that the decision to take part has no effect on services.

AI Reality Check: Participants Who Are Not Who They Say

Studies that recruit online and pay participants now attract people who invent an identity to collect the payment. Some use a language model to write plausible answers to screening questions. Ridge and colleagues (2023) describe such "imposter participants" in health research. A few checks help. Ask a screening question that only a real member of the group could answer in detail. Hold a short call before the interview. Be cautious when many replies arrive within minutes of posting. Recruitment through a trusted local organization is slower and much harder to fake. A related product, the AI "persona" that answers as a renter or a parent, is not a participant at all.

3.9 Putting It All Together

A qualitative study chooses participants on purpose for what they can explain. Maximum variation, criterion, typical case, extreme case, snowball, and theoretical sampling are the main strategies. A sampling plan sets a target for each kind of participant and shows a reader who was included and why.

Recruitment shapes the sample as much as the plan does. Gatekeepers open access to some people and not to others, and a record of who declined is part of the data.

Saturation is the point at which new interviews add little. Published studies often reach it between 9 and 17 interviews when the group is fairly similar. Saturation within one kind of participant does not show that other kinds have been heard.

3.10 Looking Ahead

The researcher now knows whom to interview. Those people will be asked to describe private and sometimes painful parts of their lives. Chapter 4 covers what the researcher owes them. The case is a study with immigrant families who have reasons to be careful about what they say and to whom.

3.11 Further Reading and References

- Patton, M. Q. (2015). *Qualitative Research and Evaluation Methods* (4th ed.). Sage. The standard reference on purposive sampling strategies.
- Guest, G., Bunce, A., and Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59 to 82.
- Hennink, M., and Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science and Medicine, 292*, 114523.
- Malterud, K., Siersma, V. D., and Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research, 26*(13), 1753 to 1760.
- Francis, J. J., Johnston, M., Robertson, C., and others (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology and Health, 25*(10), 1229 to 1245.
- Glaser, B. G., and Strauss, A. L. (1967). *The Discovery of Grounded Theory.* Aldine.
- Ridge, D., Bullock, L., Causer, H., and others (2023). "Imposter participants" in online qualitative research, a new and increasing threat to data integrity? *Health Expectations, 26*(3), 941 to 944.

3.12 Key Terms

- **Convenience sample.** A sample of whoever is easiest to reach.
- **Criterion sampling.** Choosing participants who all meet a stated condition.
- **Gatekeeper.** A person or organization that controls access to a group.
- **Information power.** How much a sample can show about the research aim.
- **Key informant.** A person with special knowledge of a setting.
- **Maximum variation.** Choosing participants who differ as much as possible on the characteristics that matter.
- **Purposive sampling.** Choosing each participant for what that person can explain.
- **Saturation.** The point at which new interviews add little that is new to the analysis.
- **Snowball sampling.** Asking each participant to suggest others.
- **Theoretical sampling.** Choosing the next participants to test or extend ideas from the analysis so far.

3.13 Exercises

Check Your Understanding

1. Explain the difference between purposive sampling and convenience sampling. Which one did the city use at the neighborhood meeting?
2. What is the main weakness of snowball sampling, and how can a researcher reduce it?
3. Define saturation. Why can a study reach saturation and still have missed an important group?
4. Why should a researcher keep a record of people who declined to take part?

Apply It

5. Build a sampling plan for this study. "How do employees at a mid-sized company experience the return to the office?" Name two characteristics that matter, draw the table, and set a target for each cell.
6. Use the file `housing-coded-segments.csv`. Count the distinct codes in transcripts H01 through H04. Then count how many new codes transcripts H05 through H08 add.

7. Write a one-paragraph recruitment flyer for the neighborhood safety study that is aimed at renters under age 30. State where you would post it.

Think Deeper

8. A researcher completes ten interviews with nurses at one hospital and reports saturation. A reviewer asks whether the findings would hold at a rural clinic. How should the researcher respond?
9. A study pays 100 dollars for a one-hour interview in a low-income neighborhood. Argue that the payment is fair. Then argue that it is a problem. Where would you set it?
10. Look at Table 3.2. Suppose these were real participants in a town of 5,000 people. Which descriptions could identify someone? Rewrite two of them to lower the risk and note what information is lost.

CHAPTER 4

Consent and Trust

Where this chapter fits. Steps 4 and 6 of the research process both involve the people who take part. This chapter covers the consent text and the protection of participants during and after the interview.

4.1 A Signature Some People Cannot Give

A community health center wants to learn why some immigrant families in its area do not use a free dental program for children. The case is a composite written for this book.

A researcher plans 20 interviews with parents. She prepares the consent form that her organization uses for every study. It runs four pages. It asks for the participant's printed name, signature, address, and date.

The first three parents she approaches decline. A staff member who knows the families explains why. Some of the parents do not have legal status, and others have relatives who do not. They have learned to be careful about any paper that carries their name and address. The form that was written to protect them is the reason they will not take part. The signed form would also be the only record that links their names to the study.

The researcher returns to her review board with a different plan. She will explain the study aloud, in the parent's language, from a one-page script. Parents will agree by voice and sign nothing. She will record no names, and each interview will carry a number. The board approves the change, and parents begin to agree.

Nothing about the ethics of the study was relaxed. The second plan protects participants better than the first. This chapter covers what a researcher owes the people who take part. It also covers how to meet those duties in a way that fits the people and the setting.

4.2 Three Principles

Modern research ethics in the United States grew out of a public scandal. From 1932 to 1972, the US Public Health Service ran a study in Tuskegee, Alabama. It followed 399 Black men who had syphilis, to observe the disease without treatment. The men were not told their diagnosis. They were not offered penicillin when it became the standard treatment in the 1940s. The study ended only after a news report in 1972.

Congress responded by creating a national commission. In 1979, the commission published the *Belmont Report*, which states three principles.

1. **Respect for persons.** People decide for themselves whether to take part, with full information. People who are less able to protect their own interests receive extra protection.
2. **Beneficence.** The researcher works to reduce possible harm and to increase possible benefit.
3. **Justice.** The burdens and benefits of research are shared fairly. A group should not carry the risks of a study because it is easy to recruit or unable to refuse.

An **institutional review board** applies these principles. It is a committee that reviews a study with human participants before the study begins. Universities, hospitals, and many agencies have one. A researcher who works outside such an institution is still bound by the principles.

4.3 Informed Consent

Informed consent means that a person agrees to take part after understanding what the study involves. It has three requirements.

Information. The person is told the purpose of the study, what they will be asked to do, and how long it will take. The person is also told the risks and benefits and how their information will be protected.

Understanding. The information is given in a form the person can follow. This means plain words, in the person's own language, at a pace that allows questions.

Voluntary choice. The person is free to say no and free to stop. No one with power over the person applies pressure.

Many consent forms meet the first requirement and fail the second. A four-page form in legal language informs no one. A researcher should write the consent text for the reader with the least schooling among the participants.

Table 4.1 shows three sentences from a standard form and a plain version of each.

TABLE 4.1. *Three sentences from a standard consent form and a plain-language version of each.*

Standard wording	Plain wording
``Participation in this research is entirely voluntary, and you may withdraw at any time without penalty or loss of benefits to which you are otherwise entitled."	``You do not have to take part. You can stop at any time. Saying no will not change any service you receive."
``Audio recordings will be retained in a secure repository and destroyed upon completion of the transcription process."	``I will keep the recording in a locked file. I will delete it after I type up what you said."
``There are no anticipated risks beyond those encountered in daily life."	``Some questions may bring up hard memories. You decide what to share."

The third row shows a second problem. The standard sentence says the study has no special risks. For a parent without legal status, an interview about family and health carries a risk. A consent text should name the risks that apply to these participants.

A full plain-language template is in the book's files as `consent-template.txt`.

Consent is also a process. A participant agreed at the start of the interview without knowing what would come up. The researcher can check again during the conversation. "Is it all right if I ask about that?" takes five seconds. At the end, the researcher can ask whether there is anything the participant would like left out.

4.4 Written and Spoken Consent

Consent is usually recorded with a signature. Review boards can approve spoken consent, with no signature, in two situations. In the first, the signed form would be the only record that links a person to the study. The main risk to the person is that the link becomes known. In the second, the study involves little risk, and the same activity outside research would not normally require a signature.

The dental study fits the first situation. With spoken consent, the researcher reads a script, answers questions, and notes the date and that the person agreed. She keeps no name.

When participants speak another language, the consent text should be translated by a person fluent in both languages and checked by a second person. If an interpreter takes part in the interview, the interpreter is bound by

the same promise of confidentiality. A family member should not interpret, and a child should never do so. A parent will not speak freely about some topics through a relative.

4.5 Confidentiality in Practice

Chapter 7 of *Before the Numbers* separates two terms. A study is **anonymous** if no one, including the researcher, can link the data to a person. A study is **confidential** if the researcher can make the link and promises not to reveal it. Most interview studies are confidential. The researcher met the participant and knows who said what.

A promise of confidentiality has to be kept through the whole life of the data. Five practices cover most of it.

1. **Separate names from data.** Give each participant a code number. Keep the list that links codes to names in a different place from the transcripts, and delete it when it is no longer needed.
2. **Remove identifying details from transcripts.** Replace names of people, employers, schools, and streets with labels in brackets.
3. **Protect the files.** Store recordings and transcripts on an encrypted drive. Do not send them by ordinary email.
4. **Limit who has access.** State who will hear the recordings and read the transcripts, and hold to it.
5. **Set a date to delete the recordings.** A voice identifies a person. Once the transcript has been checked, the recording is usually no longer needed.

A **pseudonym** is an invented name that replaces a participant's real one in a report. Choose pseudonyms that do not point back to the person. Do not use the same initials. In a small group, do not keep a detail that only one person fits.

4.6 When a Name Is Not the Only Clue

Removing a name does not always protect a person. **Deductive disclosure** occurs when a reader can work out who a participant is from the details in a report (Kaiser, 2009).

Consider this synthetic excerpt.

> Synthetic "I work nights at the poultry plant. My husband was deported in March, so now it's me and the four kids. The dentist is only open when I'm asleep."

The excerpt has no name in it. In a town with one poultry plant, a neighbor could still identify a night-shift worker with four children whose husband was deported in March. The details that make the quotation vivid are the same details that identify her.

A researcher has three options, and each has a cost.

- **Generalize the details.** "The poultry plant" becomes "a factory." "Four kids" becomes "several children." "March" becomes "this year." The quotation is safer and less specific.
- **Remove the risky clause.** The quotation can keep the point about the dentist's hours and drop the deportation.
- **Paraphrase.** The researcher reports the point in her own words and does not quote.

A fourth step helps with all three. Ask the participant. At the end of the interview, the researcher can say which parts she may want to quote. She can then ask whether any of it should be changed or left out.

Any change to a quotation should be made by rule and reported. The methods section should state that identifying details were altered and should describe the kinds of changes. It should never present an altered quotation as exact.

4.7 Work With the Data: Identifying Details Checker

A quotation can identify a speaker with no name in it. The tool marks details in a quotation that could help a reader work out who said it. Do not paste real names or real quotations from a study that is under way.

The tool runs in a web browser at qual.marginoferrormedia.com/tools/detail-checker.html. It needs no account and no software.

1. Press "Check the quotation" for the sample. The tool marks five details.
2. For each marked detail, choose one of the three changes described above. Make it more general, cut the clause, or paraphrase.
3. Type your revised quotation and check it again.
4. Read the revision as a neighbor of the speaker would. Decide whether that reader could still tell who it is.

The tool uses fixed rules and knows nothing about the setting. In a small town, a detail that the tool leaves unmarked can identify a person. The last step is the one that protects the participant.

A one-page consent text in plain language is in the file `consent-template.txt` at qual.marginoferrormedia.com/corpus.

4.8 The Limits of a Promise

A researcher should promise only what she can deliver. Two limits apply in many studies.

Mandatory reporting. In most US states, members of certain professions are required to report suspected child abuse. If the researcher is a mandatory reporter, the consent text has to say so before the interview begins.

Legal demands. Research records can be subpoenaed. For studies funded by the National Institutes of Health, a **Certificate of Confidentiality** protects identifiable data from most legal demands. Researchers with other funding can apply for one when a study collects sensitive information. A study that collects no names, like the dental study, has a simpler protection where the design allows it.

In a focus group, the researcher cannot promise confidentiality at all. She can promise her own silence and ask the group to keep what is said in the room. She cannot guarantee that eight other people will do so. The consent text should say this.

4.9 Power, Payment, and Position

An interview takes place between two people who are not equal. The researcher chose the topic, will write the report, and often comes from an institution with standing. The participant may depend on a related institution for services. Three practices reduce the effects of that imbalance.

Keep the service provider out of recruitment. Chapter 3 described the problem of a caseworker who recruits her own clients. The health center in this case should not ask its dentists or front-desk staff to recruit.

Pay fairly and without conditions. A payment thanks the participant for their time. It becomes a problem when it is large enough to lead a person to accept a risk they would otherwise refuse. Review boards call this **undue inducement**. Pay an amount in line with an hour or two of local wages. Pay it whether or not the participant finishes the interview.

Examine your own position. **Positionality** is the researcher's social position in relation to the participants, including language, class, race, citizenship, and role. An **insider** shares the participants' background. An **outsider** does not. Each has advantages. Participants may speak more openly with an insider. They may also leave unsaid what "everyone knows." They may explain more to an outsider and trust the outsider less. A researcher should state her position in the report and consider how it affected what she was told.

4.10 Giving Something Back

Participants give time and trust. A study that takes both and returns nothing has used them. Three forms of return are practical in most studies.

Send participants a short summary of the findings in plain language. Report the findings to the organization in a form it can act on. The dental study could recommend evening hours and a statement that the clinic asks no questions about legal status. Where participants are willing, invite a few of them to react to the draft findings before the report is final. Chapter 7 covers this last practice as a check on the analysis.

A Short Test for a Consent Text

Read the consent text aloud to someone outside your field. Then ask three questions. What will happen if you say yes? What could go wrong? What happens if you change your mind? If the person cannot answer all three, rewrite the text.

Ethics Moment: A Quotation the Participant Would Recognize

A participant reads the final report and finds her own words. Her name is not there, and no outsider could identify her. She recognizes herself at once, and so would her sister. The quotation describes a private matter, and it sits under a heading she finds unflattering. No rule was broken, and she still feels exposed. Confidentiality protects a participant from being identified by others. It does not protect her from how it feels to see her own account in print. A researcher should choose quotations and headings as if each participant will read the report, because some of them will.

AI Reality Check: Uploading a Transcript Is a Disclosure

Language models can transcribe recordings, summarize interviews, and suggest codes. Each of these uses sends participants' words to a company's servers. That is a disclosure to a third party, and it falls under the promise the researcher made. Before using any such tool, check three things. Does the provider store the text, and for how long? Can the provider use it to train models? Did the consent text tell participants that such a service would be used? If the answer to the third question is no, remove identifying details first, or do not upload the text. Some institutions provide approved tools that keep the data inside the institution.

4.11 Putting It All Together

The *Belmont Report* states three principles. They are respect for persons, beneficence, and justice. Informed consent puts the first into practice. It requires information, understanding, and a voluntary choice. A consent text should be written in plain words and should name the risks that apply to the participants in the study.

Most interview studies are confidential and not anonymous. The researcher separates names from data, removes identifying details, protects the files, and deletes recordings on a stated date. Removing a name is not always enough, because the details in a quotation can identify a person. The researcher can generalize the details, cut a clause, or paraphrase, and she reports that she did so.

A researcher promises only what she can deliver, keeps service providers out of recruitment, pays fairly, and states her own position. She returns the findings to the people who made them possible.

4.12 Looking Ahead

The interviews are complete, and the transcripts are checked and stripped of identifying details. The researcher now has a few hundred pages of text. Chapter 5 covers the first step of analysis, which is coding. The material is the set of 12 synthetic housing interviews.

4.13 Further Reading and References

- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979). *The Belmont Report.* US Department of Health, Education, and Welfare.
- Jones, J. H. (1993). *Bad Blood: The Tuskegee Syphilis Experiment* (new ed.). Free Press. The standard history of the study.
- Kaiser, K. (2009). Protecting respondent confidentiality in qualitative research. *Qualitative Health Research, 19*(11), 1632 to 1641. The source of the discussion of deductive disclosure.
- Saunders, B., Kitzinger, J., and Kitzinger, C. (2015). Anonymising interview data: Challenges and compromise in practice. *Qualitative Research, 15*(5), 616 to 632.
- Dwyer, S. C., and Buckle, J. L. (2009). The space between: On being an insider-outsider in qualitative research. *International Journal of Qualitative Methods, 8*(1), 54 to 63.
- US Department of Health and Human Services. *Federal Policy for the Protection of Human Subjects* (45 CFR 46). The regulation that governs review boards and consent.

4.14 Key Terms

- **Anonymous.** No one, including the researcher, can link the data to a person.
- **Beneficence.** The duty to reduce possible harm and increase possible benefit.
- **Certificate of Confidentiality.** A federal protection that shields identifiable research data from most legal demands.
- **Confidential.** The researcher can link the data to a person and promises not to reveal the link.
- **Deductive disclosure.** Identifying a participant from the details in a report.
- **Informed consent.** Agreement to take part after understanding what a study involves.
- **Insider.** A researcher who shares the participants' background.
- **Institutional review board.** A committee that reviews research with human participants before it begins.
- **Justice.** The fair sharing of the burdens and benefits of research.

- **Outsider.** A researcher who does not share the participants' background.
- **Positionality.** The researcher's social position in relation to the participants.
- **Pseudonym.** An invented name that replaces a participant's real name.
- **Respect for persons.** The principle that people decide for themselves whether to take part.
- **Undue inducement.** A payment large enough to lead a person to accept a risk they would otherwise refuse.

4.15 Exercises

Check Your Understanding

1. Name the three principles of the *Belmont Report* and give one practice that follows from each.
2. Why did the review board approve spoken consent for the dental study?
3. Explain deductive disclosure. Why is removing a name not always enough?
4. Why can a researcher not promise confidentiality in a focus group?

Apply It

5. Rewrite this sentence in plain language. "Data will be reported in aggregate, and no individually identifiable information will be disseminated."
6. Revise the synthetic excerpt about the poultry plant so that it could be published. Then state what you changed and what was lost.
7. Fill in `consent-template.txt` for a study of how college students manage the cost of food. Name one risk that applies to those participants.

Think Deeper

8. A participant says, during the interview, "Please don't write this down," and then describes something central to your research question. What do you do with what you heard?
9. You are a former client of the program you are now studying. List two ways this could help the interviews and two ways it could distort them. How would you describe your position in the report?
10. A colleague wants to paste full transcripts into a public AI tool to get a quick summary. The consent form did not mention any such tool. Write a short reply that explains the problem and suggests two alternatives.

PART III

From Text to Themes

CHAPTER 5

Coding

Where this chapter fits. Step 7 of the research process is to prepare, analyze, and interpret the data. This chapter covers the first part of the analysis, which is coding.

5.1 Twelve Transcripts on a Table

A researcher has finished 12 interviews about housing. Each one has been transcribed and checked. Names and places have been replaced with labels. The transcripts are printed and stacked on a table.

She has read each one at least once. She remembers Avery's son, who stopped doing homework in a crowded living room. She remembers Casey, who showered at a gym before work and told no one he was living in his car. She has a strong sense that the interviews are "about" something. She cannot yet say what, in a way that another person could check.

Memory is a poor tool for this step. It keeps the vivid stories and drops the plain ones. It also favors whatever the researcher expected to find. A reader of her report would have to trust her impressions.

Coding replaces impression with a procedure. The researcher goes through every transcript and attaches a short label to each passage that bears on the research question. When she finishes, she can pull every passage with a given label and read them together. This chapter covers how to do that.

The 12 interviews in this chapter are synthetic. They were written for this book, and no real person said these words. The full transcripts are in the book's files as H01 through H12.

5.2 What a Code Is

A **code** is a short label that names what a passage of text is about. A **segment** is the passage the label is attached to. A segment can be a phrase, a sentence, or several sentences. It should be long enough to make sense when read alone.

Here is one answer from Avery, the first participant.

> Synthetic "I called 211. They gave me a list. Every place on the list either had no money left or you had to have an eviction notice. I didn't have a notice. I left before it got to that. So I didn't count."

A researcher could attach the code *Help requires a crisis* to this segment. The label is short. It names what the passage is about in a way that could apply to other passages. It stays close to what Avery said and does not yet explain it.

A code is not a summary of one passage. "Avery called 211 and was turned away" describes this passage only. A code has to be general enough to be used again. When the caseworker in interview 8 says "No lease, no help," the same code applies.

5.3 Read Before You Code

Coding begins with reading. Read every transcript from start to finish before attaching any label. Braun and Clarke (2006) call this step familiarization.

Read with a pen. Mark passages that seem important, and write a few words in the margin about why. Do not try to be consistent yet. The purpose is to learn what is in the material as a whole.

After the first reading, write a memo of a page or less. Note what came up often, what surprised you, and what you expected and did not find. Chapter 2 introduced the memo. From this point on, the researcher writes one whenever her thinking changes.

5.4 Where Codes Come From

Codes come from two directions.

In **inductive coding**, the researcher builds the codes from the data. She reads a passage, asks what it is about, and writes a label. The codes are not decided in advance. This approach suits a study in which the researcher does not yet know what matters to participants.

In **deductive coding**, the researcher begins with a list of codes taken from the research question, from earlier research, or from a framework. She then

looks for passages that fit each one. This approach suits an evaluation that has set questions to answer.

Most applied studies combine the two. The researcher starts with a few codes that the research question requires and adds others as she reads. In the housing study, a question about help-seeking was in the interview guide, so a code for it was planned. No one planned a code for concealment. It came from the data, after several participants described hiding their situation.

5.5 Kinds of Codes

Saldaña (2021) describes many kinds of codes. Three are enough to begin.

A **descriptive code** names the topic of a passage in a word or short phrase. *Rent increase* and *Children and school* are descriptive codes.

An **in vivo code** uses the participant's own words as the label. Sasha, in interview 2, says, "It's the not knowing." *Not knowing* became the label for a code about uncertainty. In vivo codes keep the analysis close to how participants see their own situation.

A **process code** names an action, usually with a word that ends in "-ing." *Doubling up* and *Cutting back* are process codes. They suit passages that describe what people do.

Table 5.1 shows how the first part of Avery's interview was coded.

Table 5.1. *Six coded segments from interview H01. Synthetic data.*

Segment	Code	Kind
``New owner sent a letter, rent goes from 1,150 to 1,600."	Rent increase	Descriptive
``they all want three times the rent in income"	Screening barriers	Descriptive
``Five people. We did that for five months."	Doubling up and informal housing	Process
``My son stopped doing homework. There was no place to do it."	Children and school	Descriptive
``Because then you're that family."	Concealment	Descriptive
``You have to be all the way at the bottom before they can help you."	Help requires a crisis	Descriptive

A segment can carry more than one code. When Jordan's 12-year-old daughter finds the notice and texts, "Mom, are we getting kicked out?", the

passage concerns both a child and a dealing with the landlord. It can be coded for both.

5.6 The Codebook

As codes are created, the researcher records them in a **codebook**. A codebook is a list of every code with the rules for using it. It is what makes coding consistent from the first transcript to the last and from one coder to another.

MacQueen and colleagues (1998) recommend that each entry contain five things.

1. **A name.** A short label.
2. **A definition.** One or two sentences that state what the code covers.
3. **When to use it.** The features that a passage has to have.
4. **When not to use it.** Passages that look similar and belong under a different code.
5. **An example.** A segment from the data that fits well.

Here is one full entry from the housing codebook.

TABLE 5.2. *One entry from the codebook for the synthetic housing interviews.*

Name	Help requires a crisis
Definition	Help is tied to a notice, a document, or a waitlist.
Use when	The participant describes a rule that made help unavailable until the problem got worse, or a wait too long to be of use.
Do not use when	The participant was refused housing by a landlord. Use *Screening barriers*.
Example	``You have to be all the way at the bottom before they can help you.'' (H01)

The fourth line does the most to keep coders consistent. Two codes often cover neighboring ideas. A rule that states where one ends and the other begins prevents the same kind of passage from being coded two ways.

Table 5.3 lists all 14 codes with their definitions.

TABLE 5.3. *The 14 codes used for the synthetic housing interviews, with definitions.*

Code	Definition
Rent increase	The participant describes a rise in rent or a notice of one.
Cutting back and debt	The participant cuts other spending, pays fees, or takes on debt to cover housing.
Landlord's costs	A landlord describes taxes, insurance, repairs, or pressure to sell.

Table 5.3. *The 14 codes used for the synthetic housing interviews, with definitions.*

Code	Definition
Doubling up and informal housing	The participant stays with others, in a car, or in a room with no lease.
Concealment	The participant hides the situation or rejects a label such as homeless.
Not knowing	The participant describes fear of a notice or an inability to plan.
Help requires a crisis	Help is tied to a notice, a document, or a waitlist.
Screening barriers	Credit checks, income rules, or refusals keep the participant out of housing.
A person who knows	One person with knowledge or authority changes the outcome.
Dealing with the landlord	The participant describes contact, or a lack of contact, with a landlord or company.
Children and school	Housing problems affect children or schooling.
Work	Housing problems affect work, pay, or the commute.
Health and caregiving	Illness or caregiving is tied to the housing problem.
Moving by choice	The participant describes a move as chosen and positive.

5.7 Work With the Data: Coding Workbench

A codebook is easier to understand after you have used one. The tool shows the 12 synthetic interviews with the 14 codes from Table 5.3. Code one interview before you read further.

The tool runs in a web browser at qual.marginoferrormedia.com/tools/coding-workbench.html. It needs no account and no software.

1. Choose interview H01. Read all of it once before you code anything.
2. Select a passage. Choose a code and read its definition under the list. If the definition fits, press "Add code."
3. Code the rest of the interview in the same way. A passage can receive more than one code.
4. Press "Compare." The tool lists the segments coded for this chapter and marks the codes you did not use.
5. For each difference, read the definition again. Decide whether you would change your coding or the definition.

A difference between two coders is information about the codebook. It usually points to a definition that is too broad or to a boundary between two codes that is unclear. The next section shows how a codebook is revised for this reason. You can download your coded segments as a CSV file and keep them for Chapter 6.

5.8 The Codebook Changes

A codebook is not finished after the first transcript. It changes as the researcher reads more, in three ways.

New codes are added. A passage does not fit any existing code. In the housing study, interview 6 was with Quinn, who moved by choice and was glad of it. No existing code fit. A new code, *Moving by choice*, was added.

Codes are split. A code that covers too much is divided. An early code named *Money* covered rent increases, late fees, credit card debt, and landlords' tax bills. It was split into *Rent increase*, *Cutting back and debt*, and *Landlord's costs*.

Codes are merged. Two codes that cannot be told apart in practice are combined. *Fear of a notice* and *Cannot plan* were merged into *Not knowing*.

Researchers sometimes describe themselves as lumpers or splitters. A lumper uses a few broad codes, and a splitter uses many narrow ones. Too few codes hide differences that matter. Too many make the next step unmanageable. For a study of 10 to 20 interviews, 15 to 40 codes is a common range.

Every change to the codebook creates work. If a code is split after six transcripts have been coded, the researcher goes back and recodes those six under the new rules. Record each change in a memo with the date and the reason.

5.9 What the Coding Shows

When all 12 transcripts are coded, the researcher can count. Figure 5.1 shows how many of the 12 interviews contain each code.

These counts need careful reading. They describe the 12 transcripts. They do not estimate how common each experience is among renters. Chapter 1 explained why a qualitative sample does not support that kind of claim.

Counts are still useful inside the study. They show which codes have enough material to examine closely. They also guard against a common error. A researcher who remembers two vivid accounts of concealment may believe that "everyone" described it. The count shows the number of interviews in which it appeared.

A code that appears once is not unimportant. *Moving by choice* appears in one interview. Chapter 6 shows that this one case changes how the other eleven should be read.

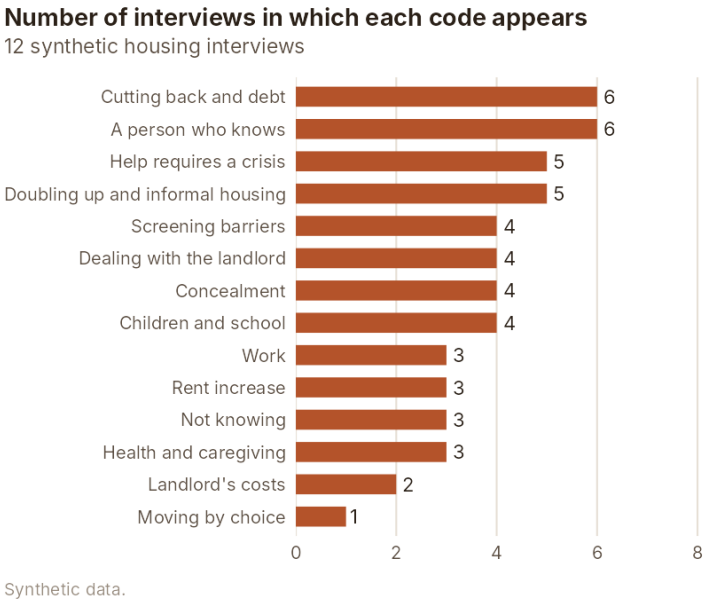


FIGURE 5.1. *Number of the 12 synthetic housing interviews in which each code appears. Eight codes appear in at least four interviews. One code, Moving by choice, appears in a single interview. Synthetic data.*

5.10 Tools for Coding

Coding can be done on paper with colored pens, in a spreadsheet, or in software made for the purpose. Common programs include NVivo, ATLAS.ti, MAXQDA, and Dedoose. All of them let a researcher select a passage, attach codes, and pull every passage with a given code.

The software stores and retrieves. It does not decide what a passage means. For a study of a dozen interviews, a spreadsheet works. Use one row for each segment, with columns for the transcript, the segment, and the code. The file `housing-coded-segments.csv` in this book is laid out that way.

A Test for a Code Definition

Give the definition and three passages to someone who has not seen the study. Ask which passages fit the code. If the person's answers differ from yours, the definition needs another sentence. Chapter 7 turns this test into a formal check.

Ethics Moment: The Label Is a Judgment

A code is the researcher's word for what a participant said. A careless label can pass judgment on the person. A passage in which a parent describes leaving before an eviction could be coded *Giving up* or *Leaving early to avoid a record*. The first label blames her. The second describes a decision she made for a reason. Read each code name and ask whether the participant would accept it as a fair description. Where one fits, use the participant's own words as the label.

AI Reality Check: Fast Codes That Need Checking

A language model can read a transcript and return a list of codes with matching quotations in under a minute. The output looks finished, and three problems are common in it. A model may "quote" words that do not appear in the transcript. Search for every quotation before you accept it. A model may also give different codes to the same passage when asked twice. Give it a written codebook and ask it to apply those codes only. Finally, a model tends to code the surface topic and miss what a passage implies. It will label Casey's account of the gym and the library as *Daily routine*. A careful reader sees that it is about concealment. Treat a model's codes as a second coder's first attempt. Chapter 7 shows how to measure how far that attempt agrees with your own.

5.11 Putting It All Together

Coding attaches a short label to each passage of text that bears on the research question. It replaces a general impression with a record that another person can inspect. Codes can be built from the data, taken from the research question, or both. Descriptive, in vivo, and process codes are three common kinds.

A codebook lists each code with a definition, rules for when to use it and when not to, and an example. The codebook changes as the researcher reads, and each change means recoding earlier transcripts. Counts of codes describe the study's own material and help the researcher check her impressions. They do not estimate how common an experience is in a population.

5.12 Looking Ahead

The 12 interviews now carry 65 coded segments under 14 codes. A list of codes is not yet a finding. Chapter 6 covers how codes are grouped into themes and how a theme is stated as a claim. It also covers how to test that claim against the cases that do not fit.

5.13 Further Reading and References

- Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers* (4th ed.). Sage. The standard reference on kinds of codes.
- MacQueen, K. M., McLellan, E., Kay, K., and Milstein, B. (1998). Codebook development for team-based qualitative analysis. *Cultural Anthropology Methods, 10*(2), 31 to 36. The source of the codebook structure used here.
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77 to 101.
- Miles, M. B., Huberman, A. M., and Saldaña, J. (2020). *Qualitative Data Analysis: A Methods Sourcebook* (4th ed.). Sage.
- Charmaz, K. (2014). *Constructing Grounded Theory* (2nd ed.). Sage. Chapters 5 and 6 cover initial and focused coding.

5.14 Key Terms

- **Code.** A short label that names what a passage of text is about.
- **Codebook.** A list of every code with the rules for using it.
- **Coding.** Attaching codes to passages of text by a consistent procedure.
- **Deductive coding.** Coding with a list of codes set in advance.
- **Descriptive code.** A code that names the topic of a passage.
- **In vivo code.** A code that uses the participant's own words as the label.
- **Inductive coding.** Building codes from the data.
- **Process code.** A code that names an action.
- **Segment.** The passage of text to which a code is attached.

5.15 Exercises

Check Your Understanding

1. Explain the difference between a code and a summary of a passage.

2. Name the five parts of a codebook entry. Which part does the most to keep two coders consistent, and why?
3. A researcher splits one code into three after coding half the transcripts. What does she have to do next?
4. *Moving by choice* appears in one of 12 interviews. Give one reason it could still matter to the analysis.

Apply It

5. Read transcript H09 in the corpus. Code at least five segments with the codes in Table 5.3. List each segment with its code.
6. Write a full codebook entry, with all five parts, for the code *Concealment*.
7. Use the file `housing-coded-segments.csv`. Which code has the most segments? In how many different interviews does it appear?

Think Deeper

8. Two researchers code the same passage. One uses *Giving up* and the other uses *Leaving early to avoid a record*. What does each label assume about the participant? How would you decide between them?
9. A study of 15 interviews ends with 110 codes. Another ends with 6. What problem does each researcher face at the next step?
10. Ask a language model to code transcript H05 without giving it a codebook. Compare its codes with the ones in this chapter. Check each quotation it gives against the transcript. Report what you found.

CHAPTER 6

From Codes to Themes

Where this chapter fits. Step 7 of the research process continues. This chapter covers how codes are grouped into themes, which are the findings of the study.

6.1 A List Is Not a Finding

The researcher from Chapter 5 presents her work to a colleague. She shows the 14 codes and the number of interviews in which each one appears. The colleague studies the list and asks, "So what did you find?"

She starts to answer and stops. *Rent increase* is a code. *Concealment* is a code. Each one names something that participants talked about. None of them is a statement that could be true or false. A list of topics does not tell a housing agency what to do differently.

A finding says something about the topics. It connects them, and it can be checked against the transcripts. The step from codes to findings is the search for themes. It is the part of the analysis that depends most on the researcher's judgment, and so it needs the most discipline.

This chapter uses the same 12 synthetic interviews as Chapter 5. They were written for this book, and no real person said these words.

6.2 What a Theme Is

A **theme** is a pattern across the data that says something about the research question. Braun and Clarke (2006) describe a theme as capturing something important in the data in relation to that question.

A theme differs from a code in two ways. It is broader, and it usually draws on several codes. It is also a claim. A code names what a passage is about. A theme states what the passages, taken together, show.

A simple test separates a theme from a topic. A topic can be written as a noun, such as "help." A theme needs a full sentence, such as "Help reaches people only after their situation has become a crisis." If a heading in a report could serve as a folder label, it is a topic. Table 6.1 shows the difference.

Table 6.1. *Four topics and the theme that states a claim about each.*

Topic	Theme
Rent	Costs rise faster than households can adjust to them.
Hidden housing	Much housing instability takes forms that no count or program records.
Getting help	Help reaches people only after their situation has become a crisis.
Relationships	One person with knowledge or authority often decides the outcome.

6.3 Six Phases

Thematic analysis is a method for finding and reporting themes. Braun and Clarke (2006) describe it in six phases.

1. **Become familiar with the data.** Read everything and write a first memo.
2. **Generate codes.** Code the full dataset.
3. **Search for themes.** Sort the codes into groups that may form themes.
4. **Review the themes.** Check each candidate theme against the coded segments and against the full transcripts.
5. **Define and name the themes.** State each theme in a sentence and decide what it covers.
6. **Write the report.** Present each theme with evidence.

Chapter 5 covered the first two phases. This chapter covers phases 3 through 5, and Chapter 8 covers the report. The phases are not strictly in order. A researcher who reviews a theme in phase 4 often returns to phase 2 to recode.

6.4 Sorting Codes Into Candidate Themes

Phase 3 begins with the codes laid out where they can be moved around. Some researchers write each code on a card. Others use a whiteboard or a spreadsheet.

The researcher looks for codes that belong together and asks what connects them. In the housing study, three codes concerned money. They were *Rent increase*, *Cutting back and debt*, and *Landlord's costs*. Read together, the segments under these codes described costs that kept rising while incomes stayed fixed. Both tenants and the small landlord described this.

Three other codes concerned situations that were hard to see from outside. They were *Doubling up and informal housing*, *Concealment*, and *Not knowing*. Casey slept in a car and told no one. Devon rented a dining room with no lease. Sasha was housed and afraid of every notice on her door.

Table 6.2 shows how the 14 codes were sorted.

Table 6.2. *Five candidate themes for the synthetic housing interviews, with the codes each one draws on. One code, Moving by choice, is held apart as a negative case.*

Theme, stated as a claim	Codes	Interviews
Costs outrun adjustment. Housing costs rise faster than households, and some landlords, can adjust.	Rent increase. Cutting back and debt. Landlord's costs.	9 of 12
Instability that is not counted. Much instability takes forms that no count or program records.	Doubling up and informal housing. Concealment. Not knowing.	8 of 12
Help arrives late. Help reaches people only after their situation has become a crisis.	Help requires a crisis. Screening barriers.	7 of 12
A person makes the difference. One person with knowledge or authority often decides the outcome.	A person who knows. Dealing with the landlord.	8 of 12
Spillovers. Housing problems spread into school, work, and health.	Children and school. Work. Health and caregiving.	6 of 12

One code was left out of the five groups. *Moving by choice* appeared only in the interview with Quinn, and it fit none of them. The researcher set it aside for the review in phase 4.

Themes do not emerge from the data unaided. The researcher builds them through the choices she makes when sorting. A different researcher could sort the same codes in another way. For that reason, the grouping has to be shown, as in Table 6.2, so that a reader can inspect it.

6.5 Work With the Data: Theme Builder

Table 6.2 shows one way to sort the 14 codes. The tool lets you build your own sorting from the same coded segments and then inspect it.

The tool runs in a web browser at qual.marginoferrormedia.com/tools/theme-builder.html. It needs no account and no software.

1. Assign each of the 14 codes to a group. Use as many groups as you need, and leave a code out if it fits nowhere.
2. Write a one-sentence claim for each group. A single noun is a topic and is not yet a theme.
3. Read the matrix. It shows which interviews contain each of your themes.
4. Read every segment listed under one theme. Decide whether all of them support your claim.
5. Press "Load the book's sorting" and compare it with yours.

Your sorting will probably differ from the one in this chapter. Both can be defensible if the segments support the claims. The reading in step 4 is the start of the review that the next section describes.

6.6 Reviewing a Theme

A candidate theme is a guess. Phase 4 tests it in two steps.

Step one. Read every segment under the theme. Pull all the coded segments that belong to the theme and read them together. Ask whether they form a coherent pattern. Segments that do not fit may need a different code, or the theme may be too broad.

For *Help arrives late*, the segments were consistent. Avery was told she needed an eviction notice. Akira, the caseworker, said that she tells people "to call back when it's worse." Imani said, "Everything is a list." Rene nearly lost a housing voucher because the letter went to an old address. The pattern holds across tenants and a service provider.

Step two. Go back to the full transcripts. Read each transcript again with the themes in mind. Ask whether the themes, taken together, give a fair account of what was said. Look for material that was never coded and that bears on a theme.

A **matrix** helps with both steps. A matrix is a table with one row for each participant and one column for each theme. Each cell records whether the theme appears for that participant, often with a short quotation. Ritchie and

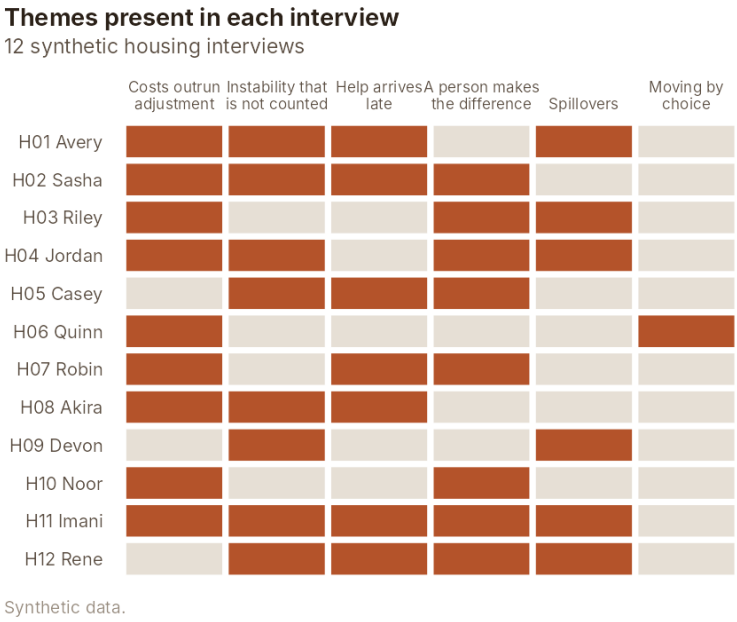


FIGURE 6.1. *Themes present in each of the 12 synthetic housing interviews. A filled square means that at least one segment in the interview was coded under the theme. Interview H06, with Quinn, has one of the five themes and the code Moving by choice. Synthetic data.*

Spencer (1994) built their framework method around such a table. Figure 6.1 shows the matrix for the housing study.

The matrix shows the analysis in one view. Reading down a column shows who supports a theme. Reading across a row shows one participant's whole account. Empty cells are informative. A researcher should ask whether a theme is absent for a participant or was only never asked about.

6.7 The Case That Does Not Fit

Row H06 in the matrix stands out. Quinn moved, as most of the others did. Quinn's interview has almost none of the themes.

A case that contradicts the pattern is a **negative case**. A researcher who is confident in her themes may want to leave it out. **Negative case analysis** is the practice of doing the reverse. The researcher searches for cases that do not fit and revises her claims until they account for those cases too.

Quinn says this about the move.

> Synthetic "Something we did. That's the difference, right? We had savings. We had a job that came with us. If I'd had to move because the rent went up and I had two weeks, that's a different story with the same truck."

This passage changes the analysis. Before reading it closely, the researcher had described her study as being about "people who had to move." Quinn moved and was not harmed by it. The harm that the other participants described did not come from the move itself. It came from a move with no time, no savings, and no say in the matter.

The researcher revises two things. She renames her subject from "moving" to "forced and unplanned moves." She also adds a condition to her first theme. Costs outrun adjustment for households that have no cushion. Quinn faced the same high rent and had the means to leave on chosen terms.

One case in twelve improved the main claim. If the researcher had dropped it as an exception, the report would have been less accurate.

A second complication comes from Robin, the landlord. An early version of the first theme read "Landlords raise rents faster than tenants can pay." Robin describes tax and insurance bills that nearly doubled. Robin also keeps one long-term tenant's rent 200 dollars below the market rate. The theme was reworded so that it describes rising costs and does not assign them to one party.

6.8 Defining and Naming

In phase 5, the researcher writes a short definition of each theme. A definition states the claim, says what the theme covers, and says where it stops. Here is the definition of one theme.

> **Help arrives late.** Participants described help that was available only after a situation had become a crisis. Programs required an eviction notice, a lease, or a place at the top of a long waitlist. People who sought help early were told that they did not yet qualify. The theme covers formal help from agencies and programs. It does not cover help from family and friends, which belongs to the theme *A person makes the difference*.

A good theme name is short and states the claim. *Help arrives late* is better than *Barriers to assistance*, which is a topic. A participant's phrase can also serve. Imani's "fine on paper, drowning" would be a strong name for the theme about instability that is not counted.

Most studies of this size report three to six themes. More than that usually means that the themes are still codes. A single theme usually means that it is too broad to be useful.

Finally, consider how the themes relate. In the housing study, they form a sequence. Costs rise. Households without a cushion move into arrangements that are not counted. Formal help is not available at that stage. What changes the outcome is often one person, such as a co-signer, a legal aid lawyer, or a caseworker who checks a list. A short account of how the themes connect is often the most useful paragraph in a report.

6.9 How Many Said It

Table 6.2 reports that 7 of 12 interviews contained the theme *Help arrives late*. Should a report include such counts?

Researchers disagree. A count helps a reader judge how far a pattern extends within the study. It can also suggest a precision that a small, purposive sample does not have. A reader may take "7 of 12" to mean 58 percent of renters.

A reasonable practice is to report counts in a table and to use careful words in the text. "Most participants" and "a few participants" are accurate when the table shows the numbers behind them. State once, plainly, that the counts describe this study's participants. Do not convert them to percentages.

Three Signs That a Theme Is Not Ready

The theme's name is one noun. The theme matches one of your interview questions word for word. You can find only two or three segments to support it, all from the same participant. In each case, return to the codes.

Ethics Moment: Leaving Out What Does Not Fit

A funder hopes the study will show that tenants need more emergency aid. Eleven interviews support that reading. One participant, a landlord, describes costs that make rent increases hard to avoid. A report without her would be cleaner. It would also misstate what the researcher was told. Participants gave their accounts on the understanding that the accounts would be reported honestly. A researcher who reports only the cases that suit a sponsor has broken that understanding with every participant, including those whose words she used.

AI Reality Check: Themes That Fit Any Study

Ask a language model for themes from a set of transcripts, and it will return a tidy list. The themes are often topics, such as "financial strain," "barriers to support," and "emotional impact." These labels would fit almost any study of hardship. A model also tends to describe agreement. It is less likely to notice the one case that contradicts the others. It will rarely report that its own theme is weak. Use a model to suggest groupings you may not have considered. Then do the review yourself. Read every segment under each theme, build the matrix, and search for the negative case. A theme is supported when you have read the evidence for it.

6.10 Putting It All Together

A theme is a pattern across the data, stated as a claim that bears on the research question. A topic can be written as a noun, and a theme needs a sentence. Thematic analysis moves through six phases, from reading to the report.

The researcher sorts codes into candidate themes and then reviews each one. She reads every segment under the theme and returns to the full transcripts. A matrix of participants by themes shows the whole analysis in one view. A negative case is examined and used to refine the claims. In the housing study, one participant who moved by choice showed that the harm came from moves that were forced and unplanned.

A report gives each theme a definition, a name that states the claim, and an account of how the themes connect. Counts describe the study's participants and are not converted to percentages.

6.11 Looking Ahead

The researcher now has five themes. They rest on her own coding and her own sorting. A skeptical reader may ask how anyone can know that she did not find what she expected to find. Chapter 7 covers the checks that answer that question. One of them is a second coder, and in that chapter the second coder is a language model.

6.12 Further Reading and References

- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77 to 101. The source of the six phases.
- Braun, V., and Clarke, V. (2022). *Thematic Analysis: A Practical Guide.* Sage. The authors' fuller and more recent account.
- Ritchie, J., and Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman and R. G. Burgess (Eds.), *Analyzing Qualitative Data* (pp. 173 to 194). Routledge. The source of the framework method and its matrix.
- Gale, N. K., Heath, G., Cameron, E., Rashid, S., and Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology, 13*, 117.
- Maxwell, J. A. (2010). Using numbers in qualitative research. *Qualitative Inquiry, 16*(6), 475 to 482. A discussion of when counts help.
- Nowell, L. S., Norris, J. M., White, D. E., and Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1 to 13.

6.13 Key Terms

- **Matrix.** A table of participants by themes that shows where each theme appears.
- **Negative case.** A case that contradicts the pattern found in the other cases.
- **Negative case analysis.** Searching for cases that do not fit and revising claims to account for them.
- **Thematic analysis.** A method for finding and reporting themes in qualitative data.
- **Theme.** A pattern across the data, stated as a claim that bears on the research question.
- **Topic.** A subject that participants talked about, with no claim attached.

6.14 Exercises

Check Your Understanding

1. State the difference between a code, a topic, and a theme. Give an example of each from the housing study.
2. List the six phases of thematic analysis.
3. What are the two steps of reviewing a candidate theme?
4. How did Quinn's interview change the researcher's first theme?

Apply It

5. Turn each topic into a theme by writing a one-sentence claim. Use the corpus for evidence. (a) "Children." (b) "Landlords." (c) "Work."
6. Use the file `housing-coded-segments.csv`. Pull every segment coded under the theme *A person makes the difference*. List the people who made the difference and their role in each case.
7. Write a definition for the theme *Instability that is not counted*, modeled on the definition of *Help arrives late* in this chapter. State what the theme covers and where it stops.

Think Deeper

8. A second researcher sorts the same 14 codes into three themes in place of five. Is one of you wrong? How should a report handle the fact that another sorting was possible?
9. A report states, "58 percent of participants described help that came too late." The study had 12 participants. What is misleading about the sentence? Rewrite it.
10. Robin, the landlord, complicated the first theme. Suppose the study had interviewed no landlords. How would the findings differ, and would the researcher have known what was missing?

PART IV

Evidence and Reporting

CHAPTER 7

Checking the Analysis

Where this chapter fits. Step 7 of the research process ends with a check on the analysis. This chapter covers how to test whether the findings can be trusted.

7.1 A Second Coder That Works in Seconds

A researcher has 40 short complaints that consumers sent to a federal agency. She has a codebook with six codes. She gives the complaints and the codebook to a language model and asks it to assign one code to each complaint. The model returns 40 codes in a few seconds.

The complaints are real. Consumers sent them to the Consumer Financial Protection Bureau and agreed to have them published. The codes are also real. A language model assigned them for this book in November 2026, in one pass, and no one corrected them.

Here is one of the 40. The model assigned it the code *Money withheld or not returned*.

> "My IRS tax return was sent to a closed account and the bank refuses to return the check back to the IRS." (Complaint 10916232)

Most readers would agree with that code. Here is another. The model assigned it *Dispute or claim denied or stalled*.

> "I was scammed on the phone into buying $1400.00 in XXXX XXXX XXXX cards using my [company] debit card in exchange for a $5000.00 loan, which the scammers agreed to provide but did not. When reported to the [company], [company] posted a $1400.00 provisional credit and thirty days later removed the credit without completing a full investigation." (Complaint 15766556)

This one is less clear. The complaint begins with a scam, which fits the code *Unauthorized activity or fraud*. It ends with a bank that withdrew a credit with-

out finishing its investigation, which fits the code the model chose. A careful human coder could choose either. The codebook needs a rule for a complaint that contains both.

The model's 40 codes look like a finished piece of work. Are they right? The question has no direct answer, because no answer key exists for what a passage means. The useful questions are narrower. Does a second coder, working alone, assign the same codes? Where the two differ, what does the difference show about the codebook? This chapter covers those questions and the other checks that make a qualitative analysis worth trusting.

7.2 Trustworthiness

Quantitative research judges a study by its validity and reliability. Lincoln and Guba (1985) proposed a parallel set of standards for qualitative research. They used **trustworthiness** as the general term and divided it into four parts. Table 7.1 lists them.

TABLE 7.1. *Four standards of trustworthiness, after Lincoln and Guba (1985).*

Standard	The question it asks	Nearest term in quantitative research	Main checks
Credibility	Do the findings give an accurate account of what participants said and meant?	Internal validity	Triangulation, member checking, negative case analysis
Transferability	Can a reader judge whether the findings apply in another setting?	External validity	Detailed description of the setting and participants
Dependability	Would the analysis hold up if someone examined how it was done?	Reliability	An audit trail, a second coder
Confirmability	Do the findings come from the data and not from the researcher's wishes?	Objectivity	Reflexive memos, an audit trail, quotations as evidence

No single check establishes any of the four. A researcher chooses several checks that suit the study and reports which ones she used.

7.3 Six Checks

Triangulation. The researcher compares findings from more than one source, method, or analyst. In the housing study of Chapters 5 and 6, tenants, a landlord, and a caseworker all described help that came late. Agreement across three kinds of participants is stronger evidence than agreement among tenants alone.

Member checking. The researcher takes the findings back to participants and asks whether the account fits their experience. She can send a one-page summary or meet with a few of them. Participants may correct a detail, object to a label, or add what was missed. They do not have a veto. A participant may dislike a finding that is accurate. The researcher records each response and what she did with it.

Negative case analysis. The researcher searches for cases that contradict her themes and revises the themes to account for them. Chapter 6 showed this with one interview.

Thick description. The researcher describes the setting, the participants, and the context in enough detail for a reader to judge whether the findings may transfer. The term comes from the anthropologist Clifford Geertz (1973).

Peer debriefing. The researcher presents the analysis to a colleague who was not part of the study. The colleague questions the themes and asks for the evidence behind each one.

Audit trail. The researcher keeps a record of every step. The record includes the interview guide and its revisions, the codebook and each change to it, the memos, and the reasons for each decision. An **audit trail** allows another person to follow the path from the transcripts to the findings.

7.4 A Second Coder

A code is a judgment. One way to test a judgment is to ask a second person to make it independently. **Intercoder agreement** is the extent to which two coders assign the same codes to the same passages.

The procedure has five steps.

1. The first coder writes the codebook and trains the second coder on it.
2. Both coders code the same sample of the data, working separately. A sample of 10 to 25 percent of the material is common.
3. The two sets of codes are compared, and agreement is calculated.
4. The coders meet and discuss each disagreement. Most disagreements point to a definition that is unclear.

5. The codebook is revised, and the process is repeated until agreement is acceptable. The remaining data are then coded.

The fourth step is the most valuable. A disagreement between two careful coders means that the codebook allows two readings. The scam complaint at the start of this chapter is an example. The remedy is a new rule in the codebook, such as "When a complaint describes a fraud and a failed dispute, code the failed dispute."

7.5 Measuring Agreement

The simplest measure is **percent agreement**. It is the share of passages to which both coders gave the same code.

Percent agreement has a weakness. Two coders will agree some of the time by chance, even if they assign codes at random. With only two codes, random coding produces agreement about half the time.

Cohen's kappa corrects for this (Cohen, 1960). It compares the agreement that was observed with the agreement that chance would produce.

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

In the formula, p_o is the observed agreement and p_e is the agreement expected by chance. A kappa of 1 means perfect agreement. A kappa of 0 means agreement no better than chance.

Here is a worked example with invented numbers. Two coders read 40 passages. For each one, they decide whether the code *Fraud* applies. Table 7.2 shows their decisions.

TABLE 7.2. *Decisions of two coders on whether one code applies to each of 40 passages. The numbers are invented for the example.*

	Coder B: yes	Coder B: no	Total
Coder A: yes	14	4	18
Coder A: no	2	20	22
Total	16	24	40

1. Calculate observed agreement. The coders agreed on $14 + 20 = 34$ passages. So $p_o = 34/40 = 0.85$.
2. Calculate each coder's rate of saying yes. Coder A said yes to 18 of 40, or 0.45. Coder B said yes to 16 of 40, or 0.40.

3. Calculate the chance that both say yes. Multiply the two rates. The result is $0.45 \times 0.40 = 0.18$.
4. Calculate the chance that both say no. The rates of saying no are 0.55 and 0.60. Their product is 0.33.
5. Add the two chances. So $p_e = 0.51$.
6. Apply the formula. Kappa is $(0.85 - 0.51)/(1 - 0.51) = 0.69$.

The coders agreed on 85 percent of the passages. After the correction for chance, kappa is 0.69.

Landis and Koch (1977) proposed labels for ranges of kappa that are still widely used. Values from 0.41 to 0.60 are "moderate," values from 0.61 to 0.80 are "substantial," and values above 0.80 are "almost perfect." These labels are conventions. Many researchers treat 0.70 or higher as acceptable for a codebook that is ready to use.

Agreement Is Not Accuracy

Two coders can agree and both be wrong. If the codebook defines a code badly, two well-trained coders will apply the bad definition in the same way. High agreement shows that the codebook can be applied consistently. It does not show that the codes capture what matters. That question belongs to the other checks, such as member checking and negative case analysis.

Some qualitative researchers do not calculate agreement at all. Braun and Clarke (2022) describe a reflexive form of thematic analysis in which coding is one researcher's interpretation. They argue that a second coder's agreement is the wrong test for it. They ask for careful reflexive memos in its place. A researcher should choose the checks that match her approach and say why. For applied studies with a team and a codebook, measuring agreement is standard practice.

7.6 The Model as a Second Coder

Return to the 40 complaints. The codebook that the model was given is in Table 7.3.

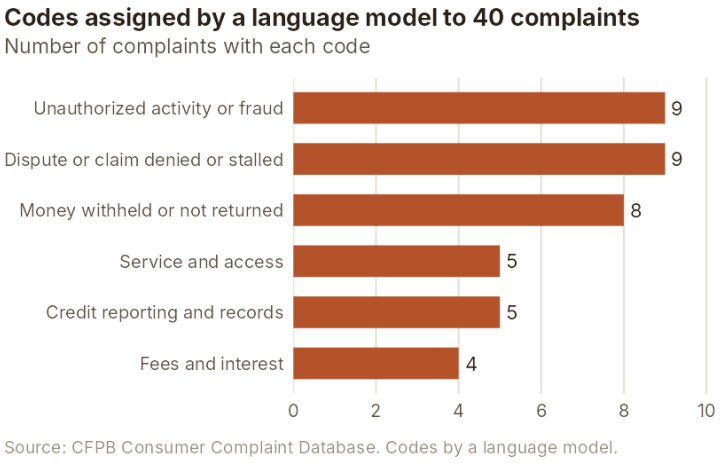


FIGURE 7.1. *Codes that a language model assigned to 40 real consumer complaints, using the six-code codebook. The codes were assigned in one pass and were not corrected. Source: complaints from the CFPB Consumer Complaint Database.*

TABLE 7.3. *The six-code codebook used for 40 real consumer complaints.*

Code	Use when the complaint is mainly about
A. Unauthorized activity or fraud	An account or a transaction the consumer did not authorize, a scam, or identity theft
B. Dispute or claim denied or stalled	A dispute or claim that the company refused, reversed, or did not finish
C. Money withheld or not returned	Funds, a refund, or a check that the consumer did not receive
D. Fees and interest	An overdraft fee, an annual fee, or interest that the consumer contests
E. Credit reporting and records	Late-payment marks, wrong information on a credit report, or the sharing of personal information
F. Service and access	Trouble reaching the company, moving money, or getting an explanation

Figure 7.1 shows how the model distributed its 40 codes.

The chart describes what the model did. It does not show whether the codes are sound. That requires a human coder who codes the same 40 complaints without seeing the model's codes. The tool for this chapter lets you be that coder.

Readers who try it should expect three things.

First, agreement will be high on complaints with one clear topic. A complaint about an overdraft fee is coded *Fees and interest* by nearly everyone.

Secondly, disagreement will be most common on complaints with two topics. The scam complaint is one. Here is another, which the model coded as *Unauthorized activity or fraud*.

> "There is a transfer of $11000.00 and I have no clue how it got on my account nor any clue on who did but I am aware I do not have access to the funds" (Complaint 3943369)

The consumer reports a transfer that they do not recognize. The consumer also cannot reach the funds, which could fit *Service and access*. A disagreement here is a finding about the codebook, which gives no rule for choosing a main topic.

Thirdly, a model's errors are not random. A human coder who is tired makes scattered mistakes. A model tends to make the same kind of mistake each time. It often follows the first topic in a passage, or the words that match a code's name. An error of that kind shifts every count in one direction.

A sound procedure for using a model as a coder has five steps.

1. Write the codebook yourself, with definitions and rules for unclear cases.
2. Code a sample yourself before you see the model's codes.
3. Give the model the same sample and the codebook. Calculate agreement.
4. Read every disagreement. Revise the codebook, and repeat until agreement is acceptable.
5. Let the model code the remaining material. Check a random sample of its work and report the agreement you found.

This is the same procedure as for a human second coder. The model is faster and never tires. It also has no knowledge of the setting and no stake in the result. Its account of its own reasoning cannot be relied on. The human coder remains responsible for every code in the final dataset.

7.7 Work With the Data: Coder Agreement

The section above compared a human coder with a language model. The tool puts you in the place of the human coder. It shows 40 real consumer complaints and the six codes used in this chapter.

The tool runs in a web browser at qual.marginoferrormedia.com/tools/coder-agreement.html. It needs no account and no software.

1. Read each complaint and choose the one code that fits best. The model's codes stay hidden while you work.
2. When you finish, read your percent agreement with the model and Cohen's kappa.
3. Compare kappa with the agreement expected by chance, which the tool also reports.
4. Read every complaint on which you and the model differ. For each one, decide whether the codebook needs a new rule.

Expect kappa to be lower than percent agreement. Kappa removes the agreement that two coders would reach by chance. Some disagreements may come from complaints with two topics, for which this codebook has no rule. Agreement shows that two coders applied the codebook in the same way. It does not show that either coder is right.

7.8 Reporting the Checks

A methods section should state which checks were used and what they showed. Here is an example for the housing study.

> One researcher coded all 12 transcripts. A second coder coded three transcripts independently with the same codebook. Agreement on those three was 82 percent, with a kappa of 0.74. The coders discussed each disagreement and revised two code definitions. The first researcher then recoded all transcripts. A summary of the themes was sent to the 12 participants, and four replied. One reply led to a change in the name of a theme. The codebook, its revisions, and the memos are available on request.

The numbers in this example are invented for illustration. The form is the point. A reader learns who coded what, how agreement was measured, what changed as a result, and where the records are.

Ethics Moment: Reporting the Checks That Failed

A researcher sends her themes to participants. Three of them say that one theme misreads their experience. She is confident in the theme and has a deadline. She could report that "member checking was conducted" and say nothing more. That sentence would be true and would mislead. A check has value only if its result is reported. State what participants said and how you responded, including the points on which you kept your own reading and why.

AI Reality Check: Asking a Model to Check Itself

It is tempting to ask a model whether its own codes are correct. The model will answer, and the answer will usually be yes. A second run of the same model is not an independent coder either. The two runs share the same training and the same blind spots, and their agreement with each other will overstate how reliable the coding is. An independent check needs a different kind of reader. The strongest one is a person who knows the setting. Report agreement between the model and a human coder, and say how many passages the human coded.

7.9 Putting It All Together

Trustworthiness has four parts. They are credibility, transferability, dependability, and confirmability. Researchers support them with triangulation, member checking, negative case analysis, thick description, peer debriefing, and an audit trail.

A second coder tests whether a codebook can be applied consistently. Percent agreement is the share of passages coded the same way, and Cohen's kappa corrects that share for chance. Disagreements show where the codebook is unclear. Agreement shows consistency and does not show that the codes capture what matters.

A language model can serve as a second coder under the same procedure as a person. The researcher writes the codebook, codes a sample first, measures agreement, reads every disagreement, and checks a sample of the model's later work. A model's agreement with itself is not an independent check.

7.10 Looking Ahead

So far, the text in this book has come from interviews. A great deal of useful text was never collected by a researcher at all. People write complaints, public comments, reviews, and answers to open-ended survey questions. Chapter 8 covers how to analyze such text, using 1,000 real consumer complaints. It also covers how to write a report in which each claim can be traced to its evidence.

7.11 Further Reading and References

- Lincoln, Y. S., and Guba, E. G. (1985). *Naturalistic Inquiry.* Sage. The source of the four standards of trustworthiness.

- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37 to 46.
- Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159 to 174.
- O'Connor, C., and Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods, 19*, 1 to 13.
- Birt, L., Scott, S., Cavers, D., Campbell, C., and Walter, F. (2016). Member checking: A tool to enhance trustworthiness or merely a nod to validation? *Qualitative Health Research, 26*(13), 1802 to 1811.
- Geertz, C. (1973). *The Interpretation of Cultures.* Basic Books. The source of the term thick description.
- Braun, V., and Clarke, V. (2022). *Thematic Analysis: A Practical Guide.* Sage.
- Consumer Financial Protection Bureau. *Consumer Complaint Database.* The source of the 40 complaints.

7.12 Key Terms

- **Audit trail.** A record of every step and decision in a study.
- **Cohen's kappa.** A measure of agreement between two coders that corrects for chance.
- **Confirmability.** The degree to which findings come from the data and not from the researcher's wishes.
- **Credibility.** The degree to which findings give an accurate account of what participants said and meant.
- **Dependability.** The degree to which the analysis would hold up under examination of how it was done.
- **Intercoder agreement.** The extent to which two coders assign the same codes to the same passages.
- **Member checking.** Taking findings back to participants and asking whether the account fits.
- **Peer debriefing.** Presenting the analysis to a colleague outside the study for questioning.
- **Percent agreement.** The share of passages to which two coders gave the same code.
- **Thick description.** A detailed account of the setting, participants, and context.

- **Trustworthiness.** The general standard for the quality of a qualitative study.

7.13 Exercises

Check Your Understanding

1. Name the four parts of trustworthiness and one check for each.
2. Why is percent agreement alone a weak measure?
3. Two coders reach a kappa of 0.90. Does this show that their codes are correct? Explain.
4. Why is a second run of the same language model not an independent check?

Apply It

5. Two coders judge 50 passages for one code. Both say yes on 12. Both say no on 30. Coder A alone says yes on 5, and Coder B alone says yes on 3. Calculate percent agreement and Cohen's kappa. Show each step.
6. Use the file `cfpb-coding-set-40.csv`. Cover the model's codes. Code the first 15 complaints yourself with Table 7.3. Calculate your percent agreement with the model.
7. Write a codebook rule that settles how to code a complaint that describes both a fraud and a failed dispute.

Think Deeper

8. After member checking, three of twelve participants reject one of your themes. List the possible reasons. How would you decide what to do, and what would you report?
9. A team plans to have a language model code 5,000 open-ended survey answers. A human will check 50 of them. Is that enough? What would you want to know before deciding?
10. Braun and Clarke argue that a second coder is the wrong test for some kinds of thematic analysis. State their argument in your own words. For which kinds of study would you still want a second coder?

CHAPTER 8

Text Beyond Interviews

Where this chapter fits. Step 8 of the research process is to write the report. This chapter also returns to step 2, because text that already exists can sometimes answer the question.

8.1 A Thousand People Who Were Not Asked

No interviewer asked the question. A person sat down, opened a form on a government website, and wrote this.

> "The said bank charged me XXXX dollars in overdraft fees – – -the account was in the positive – – [company] does not respect veterans at all – – – – – – – – —-" (Complaint 6922392)

The writer was not answering a research question. The writer wanted a bank to return some fees. The dashes and the final sentence are in the original. They are part of what was written, and a researcher should not tidy them.

The Consumer Financial Protection Bureau has received complaints from the public since 2011. When a consumer agrees, the bureau publishes the consumer's own account after removing personal details. The public database now holds several million of these accounts.

This chapter uses 1,000 of them. They concern checking and savings accounts, credit cards, and mortgages, and they were received between May 2017 and May 2026. They are real. This book has masked most company names. Each complaint keeps its number, and a reader can look it up in the bureau's database.

Text of this kind is everywhere. People write complaints, reviews, public comments on proposed rules, posts in online forums, and answers to the last question on a survey. Much of it is public, and nearly all of it was written without a researcher in the room. This chapter covers how to analyze such text and how to report findings so that each claim can be traced to its evidence.

8.2 Text That Already Exists

Researchers sometimes call this material **found text**. It differs from interview data in ways that help and in ways that hurt.

Found text has three advantages. It is plentiful, so a study can read a thousand accounts where it could afford twenty interviews. It was written without a researcher's questions, so the interviewer's wording could not have shaped it. It also often comes with useful facts attached, such as a date, a place, or a product.

It has three limits.

No follow-up. The researcher cannot ask the writer what a phrase meant. A complaint that says the bank "refuses to return the check" cannot be probed.

A purpose of its own. Every document was written to do something. A complaint is written to obtain a remedy. The writer has a reason to stress some facts and omit others. A researcher should ask who wrote the text, for whom, and to what end.

Self-selection. The people who write are not a sample of everyone affected. People who complain to a federal agency knew that the agency existed, had access to the internet, and believed that writing could help. Many people with the same problem never write.

The last limit matters most. The complaints show what kinds of problems consumers describe and how they describe them. They cannot show how common each problem is among all bank customers.

8.3 Know Where the Text Came From

Before reading a word, a researcher should establish how the documents were selected. For this chapter's file, the answer has four parts.

1. The bureau publishes a complaint's text only if the consumer agrees. Complaints without that agreement are not in the public text.
2. The bureau removes personal details and replaces them with the letters "XXXX."
3. A university research project drew these 1,000 complaints from the public database for a study of three products. The project kept only complaints that a company had closed with an explanation or with some form of relief.
4. The file is therefore not a random sample of all complaints. It contains more recent complaints than older ones, because complaint volume has grown.

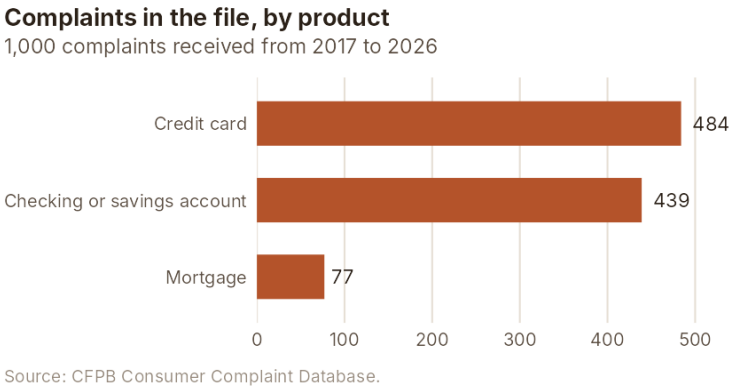


FIGURE 8.1. *Number of complaints in the file of 1,000, by product. Source: CFPB Consumer Complaint Database.*

A report has to state these facts. A reader who knows them can judge what the file can support. Figure 8.1 shows what it contains.

8.4 Counting What Is There

Content analysis is a method that sorts text into categories and counts them. It treats text as data to be measured. Krippendorff (2018) gives the standard account.

Some counts need no reading at all. The complaints range from 5 words to 3,943 words. The median is 176 words. Companies closed 771 of the 1,000 with an explanation only, 149 with monetary relief, and 80 with relief of another kind.

Other counts come from searching the text for words. Figure 8.2 shows how many of the 1,000 complaints contain each of ten terms.

The most common term in the figure concerns contact with the company and has no direct link to money. In 384 of the 1,000 complaints, the writer mentions calling or speaking with the company. That is more than the number that mention fraud (243) or a dispute (209). The count suggests something to look for in a closer reading. Many of these accounts may be about a failed attempt to get an answer as much as about the original problem.

A word count is a lead and not yet a finding. It has three weaknesses.

A word is not a meaning. "Interest" can be a charge on a credit card or a consumer's lack of interest in an offer. A search finds both.

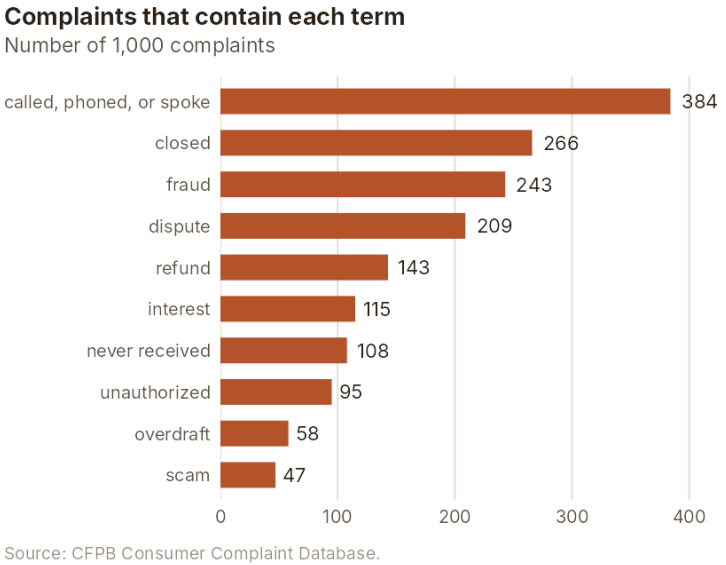


FIGURE 8.2. *Number of the 1,000 complaints that contain each term at least once. A complaint can contain more than one term. Source: CFPB Consumer Complaint Database.*

Writers use different words for one thing. One consumer writes "unauthorized." Another writes "I never opened this account." A search for the first word misses the second complaint.

A word can be denied. "This was not fraud" contains the word "fraud."

For these reasons, a keyword count should be checked by reading. Draw a random sample of the complaints that contain the term and a sample of those that do not. Read both, and estimate how often the search was right.

8.5 Counting and Reading Together

The stronger design combines the two methods.

1. **Read a sample closely.** Draw a random sample of 40 to 100 documents. Code them as Chapter 5 describes, and build a codebook.
2. **Check the codebook.** Have a second coder code part of the sample, as Chapter 7 describes.
3. **Apply the codebook to the rest.** Human coders can do this for a few thousand documents. Beyond that, researchers use keyword rules or a language model, and they check a random sample of the results by hand.

4. **Count the codes.** Report how many documents received each code.
5. **Return to the text.** For each code, read a sample of the documents and choose quotations that show what the code looks like in the writers' own words.

Chapter 7 carried out the first step for 40 of these complaints with a six-code codebook. The codebook would need more work before it could be applied to all 1,000. In particular, it has no rule for a complaint with two topics.

The counts from step 4 describe the documents in hand. With these complaints, a sentence such as "a quarter of the complaints in the file mention fraud" is accurate. "A quarter of bank customers experience fraud" is a different claim, and the file cannot support it.

8.6 Work With the Data: Text Explorer

The five steps above combine counting with reading. The tool supports both for the 1,000 real complaints. Part 1 searches and counts. Part 2 ties each claim to the complaints behind it.

The tool runs in a web browser at qual.marginoferrormedia.com/tools/text-explorer.html. It needs no account and no software.

1. In Part 1, search for "fraud." The count matches Figure 8.2.
2. Read the five complaints shown. Press "Show five other matches" and read those. Look for a complaint in which the word is denied or means something else.
3. Search for a second word for the same idea, such as "scam" or "unauthorized." Compare the counts.
4. In Part 2, write one claim that your reading supports. Press "Attach" on each complaint that supports it.
5. Download the table. It lists each claim with the complaint numbers behind it.

A claim with no complaint attached is not yet a finding. The table from step 5 has the same form as Table 8.1 later in this chapter. That table matches the wording of a report to its evidence.

8.7 Other Kinds of Found Text

Each kind of text has its own hazards.

Open-ended survey answers. Most surveys end with a box for comments. Many respondents leave the box empty, and those who write tend to

hold stronger views than those who do not. The answers are short. Code them with a small codebook and report how many respondents wrote something.

Public comments on proposed rules. Federal agencies publish the comments they receive. Many are copies of a form letter that an advocacy group wrote and its members submitted. A count of comments then measures how well a group can organize. Identify the duplicates first and report them as a separate category.

Reviews and posts online. These are plentiful and easy to collect. Some are written by people who were paid, and some by programs. A site's terms of use may also restrict collection. Check both before starting.

Organizational documents. Minutes, case notes, and reports record what an organization chose to write down. What they omit can be as informative as what they include.

8.8 Writing the Report

A qualitative report has to do two things at once. It tells the reader what was found, and it shows enough of the evidence for the reader to judge the finding. Most reports use the following order.

1. **The question and why it matters.**
2. **Methods.** Who took part or which documents were used, how they were chosen, how the data were collected and analyzed, and which checks were applied.
3. **Findings.** One section for each theme. Each section states the claim, explains it, and supports it with quotations.
4. **Discussion.** How the themes connect, what they mean for the decision at hand, and what the study cannot show.

Two published checklists list what a report should contain. The first is the Consolidated Criteria for Reporting Qualitative Research, known as COREQ. It has 32 items for studies based on interviews and focus groups (Tong, Sainsbury, and Craig, 2007). The second is the Standards for Reporting Qualitative Research, known as SRQR. It has 21 items for qualitative studies in general (O'Brien and colleagues, 2014). Many journals require one of the two.

8.9 Quotations as Evidence

A quotation in a report is evidence for a claim. Four rules govern its use.

Choose the typical quotation over the most dramatic one. The most striking quotation in a dataset is often unusual. Use a quotation that represents the segments under the theme. If you use an extreme one, say that it is extreme.

Quote exactly. Do not correct grammar or spelling. Mark any cut with an ellipsis in brackets. Do not join two passages into one quotation.

Identify the source. Give a pseudonym or a document number with each quotation, so that a reader can see that the evidence comes from more than one person. In this chapter, each complaint carries its number.

Explain the quotation. A quotation does not interpret itself. Say what it shows and how it supports the claim.

A reader should be able to trace every claim to its evidence. A simple table makes the link visible. Table 8.1 shows one for the housing study of Chapters 5 and 6.

Table 8.1. *A table that traces three claims to their evidence. The quotations are from the synthetic housing interviews.*

Claim	How far it extends	Evidence
Help reaches people only after a crisis.	7 of 12 interviews, including tenants and a caseworker	``You have to be all the way at the bottom before they can help you." (Avery, H01). ``I'm telling them to call back when it's worse." (Akira, H08)
One informed person often decides the outcome.	8 of 12 interviews	``Once I had a person next to me, the company had a person too." (Jordan, H04). ``one person who checked a list for me" (Rene, H12)
A move by choice was described as a gain.	1 of 12 interviews, a negative case	``Something we did. That's the difference, right?" (Quinn, H06)

A researcher who cannot fill in the third column for a claim does not have a finding. Build the table before writing the findings section and use it as the outline.

8.10 Matching Words to Evidence

The wording of a claim should match the strength of the evidence behind it. Table 8.2 gives a guide.

TABLE 8.2. *Wording that matches the strength of qualitative evidence.*

The evidence	Wording that fits	Wording to avoid
Present in nearly all cases, across different kinds of participants	``Participants consistently described''	``Everyone feels''
Present in more than half	``Most participants described''	``The majority of renters''
Present in several	``Several participants described''	``Many people''
Present in one or two	``One participant described''	``Some participants,'' when it was one
A relationship the researcher inferred	``These accounts suggest''	``This shows that'' or ``This proves''

The right-hand column has one fault in common. Each phrase speaks about a population or claims a certainty that a small, purposive sample cannot support.

Limitations Are Part of the Finding

State who was not included, what the documents could not show, and which checks you did not carry out. For the complaint file, the limits are plain. The writers chose to complain. The file holds three products. The bureau does not verify what consumers report. A reader who knows the limits can use the findings with more confidence.

Ethics Moment: Public Does Not Mean Free of Duty

The complaints in this chapter are public, and the writers agreed to their publication. That settles the legal question. It does not settle every ethical one. A person who posts in an online forum about a health problem has made the post public. That person did not agree to be quoted in a research report. A quotation can be traced back to the person with one search. For public text that was not written for publication, paraphrase where you can and remove identifying details. Ask whether the writer would be surprised to find the words in your report. For the complaints, this book masks most company names, because a complaint is one side's account and has not been verified.

AI Reality Check: Summaries of a Thousand Documents

A language model can read 1,000 complaints and return a one-page summary. The summary is useful as a first look. It has four known weaknesses. A model gives most of its attention to common content and tends to leave out the unusual complaint, which may be the one that matters. It may state how many complaints mention a topic, and that number is often wrong. Count with a search or a program. It may also present a "quotation" that no one wrote. Search the file for every quotation before you use it. Finally, a summary carries no record of which documents support which sentence. Ask the model to list the complaint numbers behind each point, and then read those complaints.

8.11 Putting It All Together

Found text is plentiful and was written without a researcher's influence. It also allows no follow-up, was written for a purpose of its own, and comes from people who chose to write. A researcher first establishes how the documents were selected and reports it.

Content analysis sorts text into categories and counts them. A keyword count is a lead that has to be checked by reading, because a word is not a meaning. The stronger design codes a sample closely, checks the codebook, applies it to the rest, and returns to the text for quotations.

A report states each theme as a claim and supports it with quotations that are typical, exact, identified, and explained. A table that traces each claim to its evidence keeps the report honest. The wording of each claim matches the strength of the evidence.

8.12 Looking Back

This book has followed one path. A question that numbers could not answer led to interviews. The interviews were planned, conducted with consent, and transcribed. The transcripts were coded, the codes were sorted into themes, and the themes were tested against the cases that did not fit. A second coder and the participants checked the analysis. The findings were written so that each claim could be traced to what someone said.

Before the Numbers, the companion to this book, follows the parallel path for research that produces numbers. Many questions need both. A count shows how large a problem is, and people's own accounts show what the problem is like and why it happens.

8.13 Further Reading and References

- Krippendorff, K. (2018). *Content Analysis: An Introduction to Its Methodology* (4th ed.). Sage.
- Tong, A., Sainsbury, P., and Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349 to 357.
- O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., and Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine, 89*(9), 1245 to 1251.
- Hsieh, H.-F., and Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277 to 1288.
- Consumer Financial Protection Bureau. *Consumer Complaint Database.* consumerfinance.gov/data-research/consumer-complaints. The source of the 1,000 complaints.
- Franzke, A. S., Bechmann, A., Zimmer, M., Ess, C., and the Association of Internet Researchers (2020). *Internet Research: Ethical Guidelines 3.0.* Guidance on the use of public online text.

8.14 Key Terms

- **Content analysis.** A method that sorts text into categories and counts them.
- **COREQ.** A 32-item checklist for reporting studies based on interviews and focus groups.
- **Found text.** Text that was written without a researcher's involvement.
- **Self-selection.** The process by which people choose for themselves whether to write or take part.
- **SRQR.** A 21-item standard for reporting qualitative research.

8.15 Exercises

Check Your Understanding

1. Give two advantages and two limits of found text compared with interviews.
2. Why can the complaint file not show how common a problem is among all bank customers?
3. Name three reasons a keyword count can mislead.

4. State the four rules for using a quotation as evidence.

Apply It

5. Use the file `cfpb-complaints-1000.csv`. Count the complaints that contain the word "refund." Read ten of them. In how many is the consumer asking for a refund that was not given?
6. Rewrite each sentence so that its wording matches the evidence. (a) "Renters feel ignored by landlords," from 3 of 12 interviews. (b) "This proves that early help prevents eviction," from one caseworker's account.
7. Build a table like Table 8.1 for the theme *Instability that is not counted*. Use the file `housing-coded-segments.csv` for the evidence.

Think Deeper

8. An agency receives 20,000 public comments on a proposed rule. Eighteen thousand are copies of one form letter. How should a report describe public opinion on the rule?
9. You find a public forum where people describe being evicted. The posts would make strong quotations. List the ethical questions you would settle before quoting any of them.
10. Ask a language model to summarize 50 complaints from the file and to give three quotations. Search the file for each quotation. Report how many you found word for word.

Index of Tools and Data

Tools

Each tool appears inside its chapter, in the section called Work With the Data. This page lists all seven in one place. Each one runs in a web browser and needs no account. Your work stays in your browser and is not sent anywhere.

Tool	Chapter	What it does
Interview Guide Checker	2	Marks closed, leading, and double questions in a draft interview guide.
Saturation Tracker	3	Shows how many new codes each interview adds, in any order you choose.
Identifying Details Checker	4	Marks details in a quotation that could identify the speaker.
Coding Workbench	5	Lets you code the 12 synthetic interviews and compare your codes with the book's.
Theme Builder	6	Lets you sort the 14 codes into themes and builds the matrix for your sorting.
Coder Agreement	7	Lets you code 40 real complaints and reports your agreement with a language model.
Text Explorer	8	Searches the 1,000 real complaints and traces claims to the complaints behind them.

The Two Bodies of Text

Synthetic interviews. The 12 housing interviews are plain text files at qual.marginoferrormedia.com/corpus, named H01.txt through H12.txt. They were written for this book. No real person said these words.

Real complaints. The 1,000 consumer complaints are from the public Consumer Complaint Database of the Consumer Financial Protection Bureau. Consumers agreed to their publication, and the bureau removed personal details. This book has masked most company names. Dollar amounts appear in braces in the bureau's file, and the book prints them without the braces. The bureau does not verify what consumers report.

Data Files

File	Chapters	Contents	Real or synthetic
housing-interviews-index.csv	3	The 12 participants and a description of each	Synthetic
housing-codebook.csv	5, 6	The 14 codes, their definitions, and their themes	Written for the synthetic interviews
housing-coded-segments.csv	3, 5, 6, 8	The 65 coded segments	Synthetic
cfpb-coding-set-40.csv	7	40 short complaints with the codes a language model assigned	Real text. Codes by a model
cfpb-complaints-1000.csv	8	1,000 complaints with product, issue, state, date, and outcome	Real
consent-template.txt	4	A one-page consent text in plain language	A template

The Series

Margin of Error covers statistics. *Make It Visible* covers charts and data storytelling. *Before the Numbers* covers quantitative research design. *Beyond the Numbers* covers qualitative research. All four are from Margin of Error Media.