

# Qualitative research

Matt Bhagat-Conway



# What is qualitative research?

- Most of what we've done so far has been quantitative - i.e. using numbers
- Research that does not use numbers (extensively)
- Often more exploratory and flexible than quantitative research



# When to use a qualitative approach

- When you are trying to answer a *qualitative* question - i.e. not how much, how far, how different, etc.
- When you want to dig into causes and motivations
- When there is a poor existing understanding of a subject
- When you want topics to emerge organically



# Grounded theory



# “Exempling”



# Double-dipping

- In quantitative work, we are often concerned about *double-dipping*: using the same data to identify and test theories
- This is problematic because it can lead to a lack of replicability and generalizability
- The same is true of qualitative research
- This is why it is critical research is an ongoing endeavor



# Coding

- Coding is arguably the most common way of organizing unstructured data
- You select *excerpts* from each document, and assign *codes* to them
- The codes are shared across documents
- This helps you understand which documents talk about similar things, and how prevalent different things are

*This section heavily based on Saldaña (2021)*

# Coding: some terminology

- A *document* is a single piece of text (e.g. an interview, an article, etc.)
- A *code* is a descriptor linked to one or multiple pieces of text



# Pre-coding

- Sometimes, it is helpful to highlight/italicize/underline during transcription or an initial read-through
- Can help get an initial idea of the body of work, and
- Make later coding more streamlined



# Analytic memos

- These are notes that you write down about each document during coding
- Generally reflect your own summary of the document and anything important or surprising
- Very valuable for generalization



# Coding methodology

- Keep in mind data-to-code ratio
- Try to re-use codes as much as possible
  - A primary goal of coding is finding recurring patterns



# Coding hierarchy

- Non-motorized
- Freight
- Highway/road
- Air
- Transit/rail
- Water
- Design
- Planning
- Policy/data analysis
- Theory
- Land use
- Multidisciplinary



# Coding hierarchy

- Multimodal
  - Non-motorized
  - Freight
  - Highway/road
  - Air
  - Transit/rail
  - Water
- Skills
  - Design
  - Planning
  - Policy/data analysis
- Theory
- Land use
- Multidisciplinary



# Emergent vs. pre-defined codes



# (Emergent) coding is not a linear process

- During the coding process, you will often revisit old codes to adjust, merge, or split them
- Even after coding, there is often a second stage of going through the codes to produce analytical insights



# Codes to narrative themes



# Brainspace for coding

- Valuable codes are going to depend on on your research questions
- It's important to figure out those questions before coding, and keep those questions top of mind while coding



# Types of coding

- There are many different types of coding
- We'll cover a few here; more in Saldaña (2021)
- We'll apply different types of coding to a few example documents



# Example document 1

We like to think that we know things and we like people to give us clear numbers. So, tomorrow it's going to rain. People have trouble with the idea that there's a 40% probability of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, we face a lot of uncertainty about the climate system, and I think we have to recognize the uncertainty. And, it doesn't mean we shouldn't do anything. It just means that we have to be clear about the things we understand and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO<sup>2</sup> will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, even if we did know—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—we don't know what the impact would be. We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

[Robert Pyndick on EconTalk, abridged](#)



# Macrocoding or “lumping”

We like to think that we know things and we like people to give us clear numbers. So, tomorrow it's going to rain. People have trouble with the idea that there's a 40% probability of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, we face a lot of uncertainty about the climate system, and I think we have to recognize the uncertainty. And, it doesn't mean we shouldn't do anything. It just means that we have to be clear about the things we understand and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO<sup>2</sup> will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, even if we did know—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—we don't know what the impact would be. We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

interpreting uncertainty   climate uncertainty   climate impacts



# Microcoding or “splitting”

We like to think that we know things and we like people to give us clear numbers. So, tomorrow it's going to rain. People have trouble with the idea that there's a 40% probability of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, we face a lot of uncertainty about the climate system, and I think we have to recognize the uncertainty. And, it doesn't mean we shouldn't do anything. It just means that we have to be clear about the things we understand and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO<sup>2</sup> will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, even if we did know—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—we don't know what the impact would be. We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

Human interpretations of uncertainty    Decisionmaking under climate uncertainty    Future climate

Impacts of climate change



# Practical recommendations for coding

- If splitting (or even if not), number of codes can become overwhelming
- Regularly revisit emergent codes to combine
- Consider being selective about what you code



# Coding paradigms

- Many different frameworks for developing codes
- The right choice will depend on the research question, the document type, and researcher preferences
- Coding paradigm may be chosen as part of the research design, or chosen based on experience reading the documents



# Descriptive coding



## Descriptive coding: example

We like to think that we know things and we like people to give us clear numbers. So, tomorrow it's going to rain. People have trouble with the idea that there's a 40% probability of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, we face a lot of uncertainty about the climate system, and I think we have to recognize the uncertainty. And, it doesn't mean we shouldn't do anything. It just means that we have to be clear about the things we understand and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO<sup>2</sup> will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, even if we did know—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—we don't know what the impact would be. We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

interpreting uncertainty   climate uncertainty   climate impacts



# You don't have to code every word

- To fit on the slides, I've heavily edited this example to just a few relevant pieces
- When working with actual text, you generally won't assign a code to every word
- Some will be out of scope (e.g. interviewer questions, sometimes), or just not relevant to your research question



## *In vivo coding*

- Codes are actual quotes from the text
- Less interpretation required
- Easier to preserve voice
- Generally leads to a large number of codes to be further summarized later
  - There may be some opportunity for consolidation during coding



## In vivo coding: example

**We like to think that we know things** and we like people to give us clear numbers. So, tomorrow it's going to rain. **People have trouble with** the idea that there's a 40% **probability** of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, we face a lot of **uncertainty about the climate system**, and I think we have to recognize the uncertainty. **And, it doesn't mean we shouldn't do anything.** It just means that **we have to be clear about the things we understand** and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO<sup>2</sup> will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, even if we did know—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—**we don't know what the impact would be.** We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

we like to think that we know things    people have trouble with . . . probability

uncertainty about the climate system    it doesn't mean we shouldn't do anything

we have to be clear about the things we understand    we don't know what the impact will be



# Process/action coding

- Coding with verbs, specifically gerunds (-ing words) in each code
- Applicable in many cases
- Particularly useful for studies about human behavior



## Process coding: example

We like to think that we know things and we like people to give us clear numbers. So, tomorrow it's going to rain. People have trouble with the idea that there's a 40% probability of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, we face a lot of uncertainty about the climate system, and I think we have to recognize the uncertainty. And, it doesn't mean we shouldn't do anything. It just means that we have to be clear about the things we understand and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO<sup>2</sup> will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, even if we did know—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—we don't know what the impact would be. We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

Thinking about uncertainty   Recognizing uncertainty   Planning for uncertainty   Predicting the future



# Versus coding

- Each code compares/contrasts two viewpoints or topics

## Versus coding: example

We like to think that we know things and we like people to give us clear numbers. So, tomorrow it's going to rain. People have trouble with the idea that there's a 40% probability of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, we face a lot of uncertainty about the climate system, and I think we have to recognize the uncertainty. And, it doesn't mean we shouldn't do anything. It just means that we have to be clear about the things we understand and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO<sup>2</sup> will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, even if we did know—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—we don't know what the impact would be. We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

certainty vs uncertainty   action vs inaction   predictions vs impacts

# Emotion coding



## Example document 2

I had not expected, I think like most people, the storm to be much of anything. We figured eh, a couple inches in the basement, it'll be a little annoying, and go to sleep. But in the middle of the night, that Friday night, the electricity went out a few times and I couldn't breathe because I'm connected to a CPAP machine. I was like all, I didn't think about it, but Saturday morning when I got woke up and I was like, thank you, Jesus, that I woke up, there was nothing, there was no electricity. And I was well shook. I was like, wait a minute, hold up. We got to find someplace for me to plug in.

Long story short, I basically did find that I could go to AB Tech, the medical shelter, right across from Mission Hospital because Mission was getting slammed. They were just getting slammed. And they were like, okay, AB Tech, can you take folks less critical. You know, if they just need to plug in or they just need checking on that sort of thing, just have 'em over here.

And it was very different. I was able to eventually plug in because they had generators at first and you didn't have enough and they had to find some more and get some more coming in. And we had to wait for supplies. So the thing for me is that I had to come to an understanding for me personally that, well, I didn't control nothing yesterday. I don't control nothing today, and I don't control nothing tomorrow.

[Rev. Penny Meacham on her experience during Hurricane Helene, abbreviated](#)



## Emotion coding: example

I had not expected, I think like most people, the storm to be much of anything. We figured eh, a couple inches in the basement, it'll be a little annoying, and go to sleep. But in the middle of the night, that Friday night, the electricity went out a few times and I couldn't breathe because I'm connected to a CPAP machine. I was like all, I didn't think about it, but Saturday morning when I got woke up and I was like, thank you, Jesus, that I woke up, there was nothing, there was no electricity. And I was well shook. I was like, wait a minute, hold up. We got to find someplace for me to plug in.

Long story short, I basically did find that I could go to AB Tech, the medical shelter, right across from Mission Hospital because Mission was getting slammed. They were just getting slammed. And they were like, okay, AB Tech, can you take folks less critical. You know, if they just need to plug in or they just need checking on that sort of thing, just have 'em over here.

And it was very different. I was able to eventually plug in because they had generators at first and you didn't have enough and they had to find some more and get some more coming in. And we had to wait for supplies. So the thing for me is that I had to come to an understanding for me personally that, well, I didn't control nothing yesterday. I don't control nothing today, and I don't control nothing tomorrow.

lack of concern   thankfulness   fear   urgency   powerlessness



# Values coding

- Assigning codes that represent values, beliefs, or worldview
- Best when intention is to extract viewpoints or understand culture



## Values coding: example

I had not expected, I think like most people, the storm to be much of anything. We figured eh, a couple inches in the basement, it'll be a little annoying, and go to sleep. But in the middle of the night, that Friday night, the electricity went out a few times and I couldn't breathe because I'm connected to a CPAP machine. I was like all, I didn't think about it, but Saturday morning when I got woke up and I was like, thank you, Jesus, that I woke up, there was nothing, there was no electricity. And I was well shook. I was like, wait a minute, hold up. We got to find someplace for me to plug in.

Long story short, I basically did find that I could go to AB Tech, the medical shelter, right across from Mission Hospital because Mission was getting slammed. They were just getting slammed. And they were like, okay, AB Tech, can you take folks less critical. You know, if they just need to plug in or they just need checking on that sort of thing, just have 'em over here.

And it was very different. I was able to eventually plug in because they had generators at first and you didn't have enough and they had to find some more and get some more coming in. And we had to wait for supplies. So the thing for me is that I had to come to an understanding for me personally that, well, I didn't control nothing yesterday. I don't control nothing today, and I don't control nothing tomorrow.

higher power   pragmatism   lack of control

# Magnitude coding

- Magnitude coding applies codes that have a value judgement (could be positive/negative or more nuanced)
- Often used when doing program evaluation, as it allows extracting sentiments towards the program

## Example document 3

I was released from prison on July 7th, 2018. I went into sober living at Biltmore Housing, and I heard about this program, and I asked them about it. I had to get a referral through the re-entry program.

I had to work hard to get get where I'm at today. I had to prove that I was serious about it otherwise I wouldn't have been eligible, and it took a couple months after I was referred for the program and paid for the car.

This car has totally changed my life. I was able to attend TransTech school and get my CDL [Commercial Driver License]. It starts at 6:00 in the morning, there's no buses that early. After I got my CDL license it enabled me to get an awesome job. I've been a CDL driver for 16 months now with the same company. I love my job. I love the people I work with. But I have to be there 5:30. There's no buses. I couldn't do what I do without the car that I got from this program.

Public transportation in Asheville, they're making strides, but it's not nearly good enough. Many people work graveyard swing shifts, there's no there's no way to get home at night.

[Kelly discussing her experience with Working Wheels, abridged](#)



## Magnitude coding: example

I was released from prison on July 7th, 2018. I went into sober living at Biltmore Housing, and I heard about this program, and I asked them about it. I had to get a referral through the re-entry program.

I had to work hard to get where I'm at today. I had to prove that I was serious about it otherwise I wouldn't have been eligible, and it took a couple months after I was referred for the program and paid for the car.

This car has totally changed my life. I was able to attend TransTech school and get my CDL [Commercial Driver License]. It starts at 6:00 in the morning, there's no buses that early. After I got my CDL license it enabled me to get an awesome job. I've been a CDL driver for 16 months now with the same company. I love my job. I love the people I work with. But I have to be there 5:30. There's no buses. I couldn't do what I do without the car that I got from this program.

Public transportation in Asheville, they're making strides, but it's not nearly good enough. Many people work graveyard swing shifts, there's no there's no way to get home at night.

± application process" + life outcomes + economic outcomes - public transit

# Multiple coders

- It is common to have multiple coders on a project
- Often, coders will double-code all or a subset of documents, to ensure consistency
- Even with a single coder, double-coding may be undertaken to ensure stability



# Frontiers in coding: large language models

- Historically, coding has been done by humans
- Large language models are a promising tool for coding, *with appropriate prompting*
- Still very much an active frontier of research; I am working with a PhD student experimenting with this, and Dr. Palm has experimented with it as well



## In vivo coding: my codes

**We like to think that we know things** and we like people to give us clear numbers. So, tomorrow it's going to rain. **People have trouble with** the idea that there's a 40% **probability** of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, we face a lot of **uncertainty about the climate system**, and I think we have to recognize the uncertainty. **And, it doesn't mean we shouldn't do anything.** It just means that **we have to be clear about the things we understand** and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO<sup>2</sup> will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, even if we did know—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—**we don't know what the impact would be.** We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

we like to think that we know things    people have trouble with . . . probability

uncertainty about the climate system    it doesn't mean we shouldn't do anything

we have to be clear about the things we understand    we don't know what the impact will be

## In vivo coding: ChatGPT's codes

We like to think that we know things and **we like people to give us clear numbers**. So, tomorrow it's going to rain. People have trouble with the idea that there's a 40% probability of rain. People want to hear things like, 'Well, 30 years from now, the temperature is going to go up by one degree Celsius or two degree Celsius or whatever, if we don't do X, Y, and Z.'

But, **we face a lot of uncertainty about the climate system**, and I think we have to recognize the uncertainty. **And, it doesn't mean we shouldn't do anything**. It just means that we have to be clear about the things we understand and the things we don't understand.

The simple fact is, we don't know what will happen. Even if we knew how much CO2 will be emitted over the coming 50 years, we wouldn't know how much the temperature would increase as a result and what will happen to sea levels. We might have estimates. But we don't know.

And, **even if we did know**—even if we knew precisely what the temperature will do and sea levels will do over the next 50 years—we don't know what the impact would be. We don't know whether that would be devastating to world GDP [Gross Domestic Product], to the world economy; whether it would be a little bit harmful, whether it would be moderately harmful.

We like people to give us clear numbers | We face a lot of uncertainty about the climate system

it doesn't mean we shouldn't do anything | Even if we did know...



# ChatGPT prompt



# Audio

- Historically, coding almost always done with text data
- Some software now can apply codes directly to audio or video
- Tradeoff is getting into the coding process faster, but listening is slower than reading



# Software for qualitative coding

- For very simple projects, coding with Word, Excel, or on paper may be sufficient
- For anything with more than a handful of short documents, purpose-built software should be used
- There are many software packages for qualitative data analysis
- Common ones include NVivo, Dedoose, Atlas.ti

# Qualitative research and causality



# Hypothesis coding

- Hypothesis coding is a type of pre-determined or deductive coding
- Codes are specified before coding, with the specific goal of measuring some hypothesis
  - Based on theory, or previous study
  - Ideally before data collection, or by someone minimally involved in data collection
- Hypotheses are made about how often the codes will occur, or how they will relate to one another



# Causality and emergent coding: train/test coding



# Causality with *outcomes*



# But wait! Causality with *processes*



# Sources of qualitative data

- Interviews
- Published documents
- Public meeting comments
- Social media
- Ethnography



# Ethnography

- Ethnography is studying people in their communities, rather than through interviews or written documents
- The researcher may be an observer in the community,
- Or become embedded into the community as a participant
- Dr. Hernandez's research is based in ethnography



# Mixed methods research

- “Mixed methods” is research that uses multiple methods, generally a mixture of quantitative and qualitative research
- The methods complement in each other in different ways



# Qualitative research before quantitative

- Qualitative research can generate hypotheses, ideas for survey questions, or motivate a research question



# Roundabouts are for rich people



© Strong Towns Langley

# Qualitative research after quantitative

- Qualitative research can help understand why observed patterns exist in quantitative data
- It can provide additional insights for subpopulations too small to support quantitative research



# Residential location choice process

- My PhD advisor was really interested in how people choose where to live
- In grad school, we did a number of quantitative studies about this
- She then followed these up with interviews to get a better understanding of the processes we were documenting with our quantitative models



# Quantitative methods for text

- There are a number of quantitative tools that can be used with textual data, blurring the lines between qualitative and quantitative work



# Quantitative analysis of coded data



# Other techniques

- Word clouds
- Sentiment analysis
- Topic modeling



# References

- Egami, Naoki, Christian J. Fong, Justin Grimmer, Margaret E. Roberts, and Brandon M. Stewart. 2022. "How to Make Causal Inferences Using Texts." *Science Advances* 8 (42): eabg2652. <https://doi.org/10.1126/sciadv.abg2652>.
- Glaser, Barney G., and Anselm L. Strauss. 2017. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. London New York: Routledge.
- Saldaña, Johnny. 2021. *The Coding Manual for Qualitative Researchers*. Fourth edition. Washington, D.C.: SAGE Publications Ltd.



This work by [Matthew Bhagat-Conway](#) is licensed under a [Creative Commons Attribution 4.0 International License](#).