



World Leaders in Research-Based User Experience

How to Analyze Qualitative Data from UX Research: Thematic Analysis

Summary: Identifying the main themes in data from user studies — such as: interviews, focus groups, diary studies, and field studies — is often done through thematic analysis.

By [Maria Rosala](#) on September 29, 2019

Topics: [Research Methods](#)

Uncovering themes in qualitative data can be daunting and difficult. Summarizing a quantitative study is relatively clear: you scored 25% better than the competition, let's say. But how do you summarize a collection of qualitative observations?

In the early stages of a project, exploratory research is often carried out. This research often produces a lot of qualitative data, which can include:

Qualitative attitudinal data, such as people's thoughts, beliefs and self-reported needs obtained from user interviews, focus groups and even diary studies

Qualitative behavioral data, such as observations about people's behavior collected through contextual inquiry and other ethnographic approaches

Thematic analysis, which anyone can do, renders important aspects of qualitative data visible and makes uncovering themes easier.

What Is a Thematic Analysis?

Definition: Thematic analysis is a systematic method of breaking down and organizing rich data from qualitative research by tagging individual observations and quotations with appropriate codes, to facilitate the discovery of significant themes.

As the name implies, a thematic analysis involves finding *themes*.

Definition: A theme:

- is a description of a belief, practice, need, or another phenomenon that is discovered from the data
- emerges when related findings appear multiple times across participants or data sources

Challenges with Analyzing Qualitative Data

Many researchers feel overwhelmed by qualitative data from exploratory research conducted in the early stages of a project. The table below highlights some common challenges and resulting issues.

CHALLENGES	RESULTING ISSUES
Large quantity of data: Qualitative research results in long transcripts and extensive field notes that can be time-consuming to read; you may have a hard time seeing patterns and remembering what’s important.	Superficial analysis: Analysis is often done very superficially, just skimming topics, focusing on only memorable events and quotes, and missing large sections of notes.
Rich data: There are lots of detail within every sentence or paragraph. It can be hard to see which details are useful and which are superfluous.	Analysis becomes a description of many details: The analysis simply becomes a regurgitation of what participants’ may have said or done, without any analytical thinking applied to

	any strategy that thinking applied to it.
Contradicting data: Sometimes the data from different participants or even from the same participant contains contradictions that researchers have to make sense of.	Findings are not definitive: Analysis is not definitive because participant feedback is conflicting, or, worse, viewpoints that don't fit with the researcher's belief are ignored.
No goals set for the analysis: The aims of the initial data collection are lost because researchers can easily become too absorbed in the detail.	Wasted time and misdirected analysis: The analysis lacks focus and the research reports on the wrong thing.

Without some form of systematic process, the problems outlined easily arise when analyzing qualitative data. Thematic analysis keeps researchers organized and focused and gives them a general process to follow when analyzing qualitative data.

Tools and Methods for Conducting Thematic Analysis

A thematic analysis can be done in many different ways. The best tool or method for this process is determined based on the:

- data
- context and constraints of the data-analysis phase
- the researcher's personal style of work

3 common methods include:

- Using software
- Journaling
- Using affinity diagramming techniques

Using Software

To analyze large amounts of qualitative data, qualitative researchers often use software, known as CAQDAS (Computer-Aided Qualitative-Data–Analysis software) — pronounced “cak·das”. Researchers upload transcripts and field notes into a software program and then analyze the text systematically through formal coding. The software helps with the discovery of themes by offering various visualization tools, such as word trees or word clouds, that allow the coded data to be manipulated in many different ways.

Benefits

- The analysis is very thorough.
- A physical project file (which contains the raw data and the analysis) can be shared with others. (This method is popular in student projects at academic institutions.)

Drawbacks

- Time-consuming, as it results in many codes which need to be condensed into a small, manageable list
- Expensive
- Hard to analyze with others synchronously
- Requires some learning of the software
- Can feel restrictive

Journaling

Writing thought processes and ideas you have about a text is common among researchers practicing grounded-theory methodology. Journaling as a form of thematic analysis is based on this methodology and involves manual annotation and highlighting of the data, followed by writing down the researchers’ ideas and thought processes. The notes are known as memos (not to be confused with the office memo delivering news to employees).

Benefits

The process encourages reflection through the writing of detailed notes

- The process encourages reflection through the writing of detailed notes.
- Researchers have a record of how they arrived at their themes.
- The analysis is cheap and flexible.

Drawbacks

- Hard to do collaboratively

Affinity-Diagramming Techniques

The data is highlighted, cut out physically or digitally, and reassembled into meaningful groups until themes emerge on a physical or digital board. (See a [video demonstrating affinity-diagramming](#).)

Benefits

- Can be done collaboratively
- Quick arriving at themes
- Cheap and flexible
- Visual, and supports an iterative-analysis process

Drawbacks

- Not as thorough as other methods as often segments of text aren't coded multiple times
- Hard to do when data is very varied, or there is a lot of data

Codes and Coding

All methods of thematic analysis assume some amount of coding (not to be confused with writing a program in a programming language).

Definition: A code is a word or phrase that acts as a label for a segment of text.

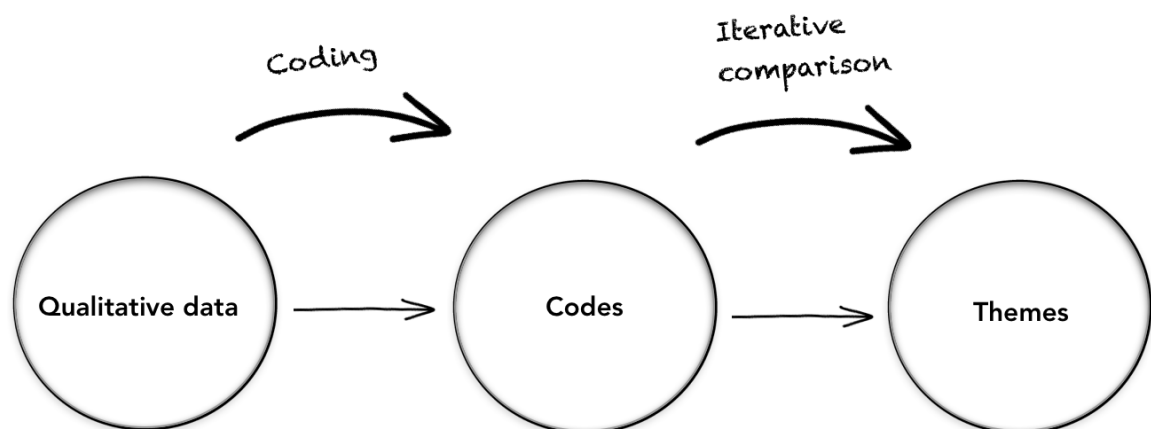
A code describes what the text is about and is a shorthand for more complicated information. (A good analogy is that a code describes data like a keyword describes an article or like a hashtag describes a tweet.) Often, qualitative

researchers will not only have a name for each code but will also have a description of what the code means and examples of text that fit or don't fit the code. These descriptions and examples are especially useful if more than one person is responsible for coding the data or if coding is done over a longer period of time.

Definition: Coding refers to the process of labeling segments of text with the appropriate codes.

Once codes are assigned, it's easy to identify and compare segments of text that are about the same thing. The codes allow us to sort information easily and to analyze data to uncover similarities, differences, and relationships among segments. We can then arrive at an understanding of the essential themes.

THEMATIC ANALYSIS



NNGROUP.COM **NN/g**

A thematic analysis starts with coding qualitative data. Through a systematic process of comparing segments of text within and between codes, the researcher arrives at themes.

Code Types: Descriptive and Interpretive

Codes can be:

- **Descriptive:** They describe what the data is about
- **Interpretive:** They are an analytical reading of the data, adding the researcher's interpretive lens to it.

To see examples of descriptive and interpretive codes, let's look at a quote from an interview I performed with a UX practitioner earlier this year (as part of our UX Careers research, to be published in our [UX Careers report](#)).

"I was petrified about facilitating a meeting and my company offered a day-and-a-half– long course. So, I went in there and the instructor did something that I felt was horrible at the time, but I've since really come to appreciate it. The first thing that we did was we filled out a sheet of paper with our name and wrote down our worst fear of moderating or facilitating and we turned it in and then he said, okay, tomorrow you're going to act out this situation (...) the next day we came back and I would leave the room while the rest of the team read, they read my worst fear, figured out how they'd act it out, and then I'd walk in and facilitate for 10 minutes with that. And that really helped me realize that there isn't anything to be afraid of, that our fears are really in our head most of the time and facing that made me realize I can handle these situations."

Here are possible descriptive and interpretive codes for the text above:

Descriptive code: *how skills are acquired*

Rationale behind the code label: Participants were asked to describe how they came to possess certain skills.

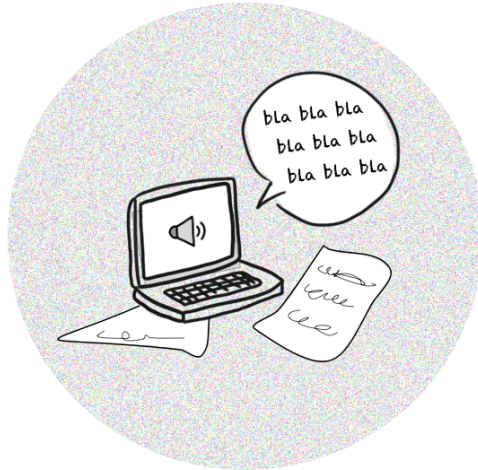
Interpretive code: *self-reflection*

Rationale behind the code label: The participant describes how this experience changed her beliefs about facilitation and how she reflected on her fear.

Steps to Conduct a Thematic Analysis

Regardless of which tool you use (software, journaling, or affinity diagramming), the act of conducting a thematic analysis can be broken down into 6 steps.

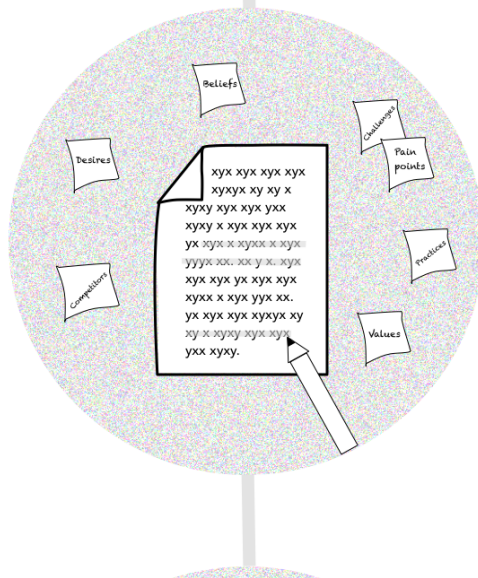
6 STEPS TO DOING A THEMATIC ANALYSIS



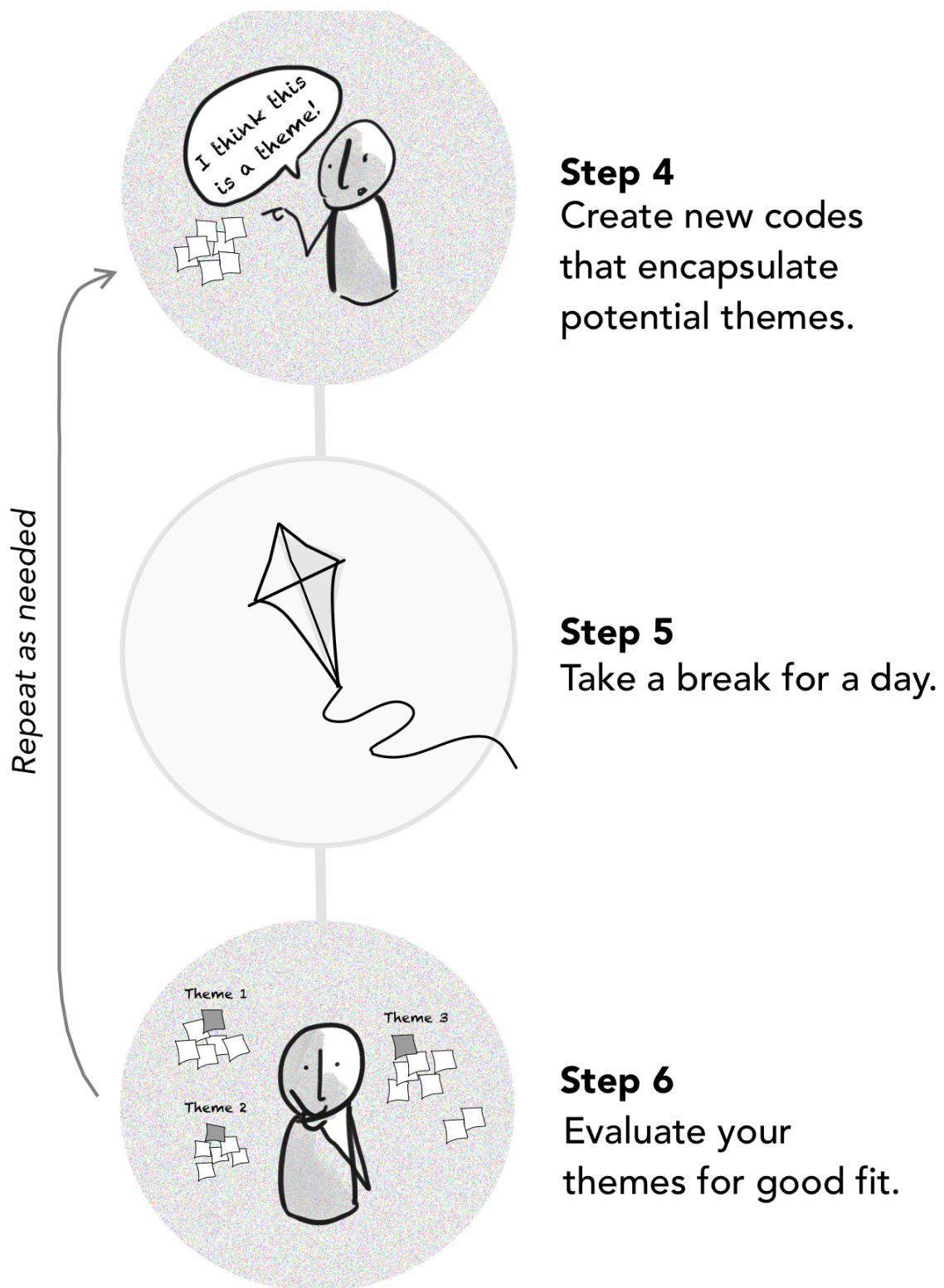
Step 1
Gather your data.



Step 2
Read all your data from beginning to end.



Step 3
Code the text based on what it's about.

NNGROUP.COM **NN/g**

A thematic analysis involves 6 different phases: gathering the data, reading all the data from beginning to end, coding the text based on what it's about, creating new codes that encapsulate candidate themes, taking a break and coming back to the analysis later, and evaluating your themes for good fit.

Step 1: Gather All Your Data

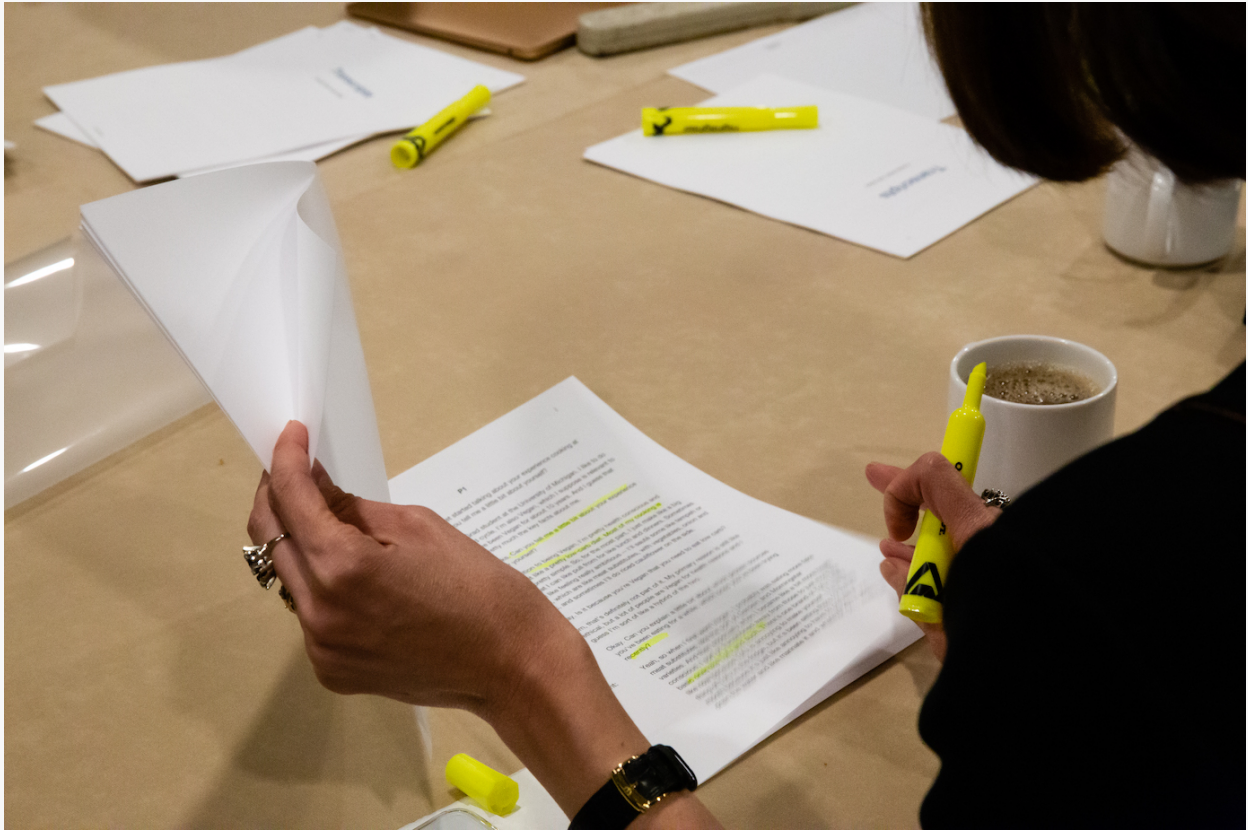
Start with the raw data, such as interview or focus-group transcripts, field notes, or [diary study](#) entries. I recommended [transcribing audio recordings from interviews](#) and using the transcriptions for analysis instead of [relying on patchy memory](#).

Step 2: Read All Your Data from Beginning to End

Familiarize yourself with the data before you begin the analysis, even if you were the one to perform the research. Read all your transcripts, field notes, and other data sources before analyzing them. At this step, you can involve your team in the project. [Involving your team](#) instills knowledge of users and [empathy](#) for them and their [needs](#).

Run a [workshop](#) (or a series of workshops if your team is very large or you have a lot of data). Follow these steps:

1. Before your team members engage with the data, write your research questions on a whiteboard or piece of flipchart paper in order to make the questions easy to refer to while working.
2. Give each member a transcript or one field- or diary-study entry. Tell people to highlight anything they think is important.
3. Once team members have completed reading their entries, they can pass their transcript or entry to someone else and receive a new one from another team member. This step is repeated until all team members have engaged with all the data.
4. Discuss as a group what you noticed or found surprising.



A workshop where each team member reads each diary- or field-study entry and highlights important bits is a good way of getting team members to actively engage with the text, as opposed to just reading it and letting it wash over them.

While it's best if your team observes all your research sessions, that may not be possible if you have a lot of sessions or a big team. When individual team members observe only a handful of sessions, they sometimes walk away with an incomplete understanding of the findings. The workshop can solve that problem, since everyone will read all the session transcripts.

Step 3: Code the Text Based on What It's About

In the coding step, highlighted sections need to be categorized so that the highlighted sections can be easily compared.

At this stage, remind yourself of your research objectives. Print your research questions out. Stick them up on a wall or on a whiteboard in the room where you're conducting the analysis.

If you have adequate time, you can involve your team in this initial coding step. If

time is limited and there is a lot of data to work through, then do this step by yourself and invite your team later to review your codes and help flesh out the themes.

As you are coding, review each segment of text and ask yourself “What is this about?” Give the fragment a name that describes the data (a descriptive code). You can also add interpretive codes to the text at this stage. However, these will typically become easier to assign later.

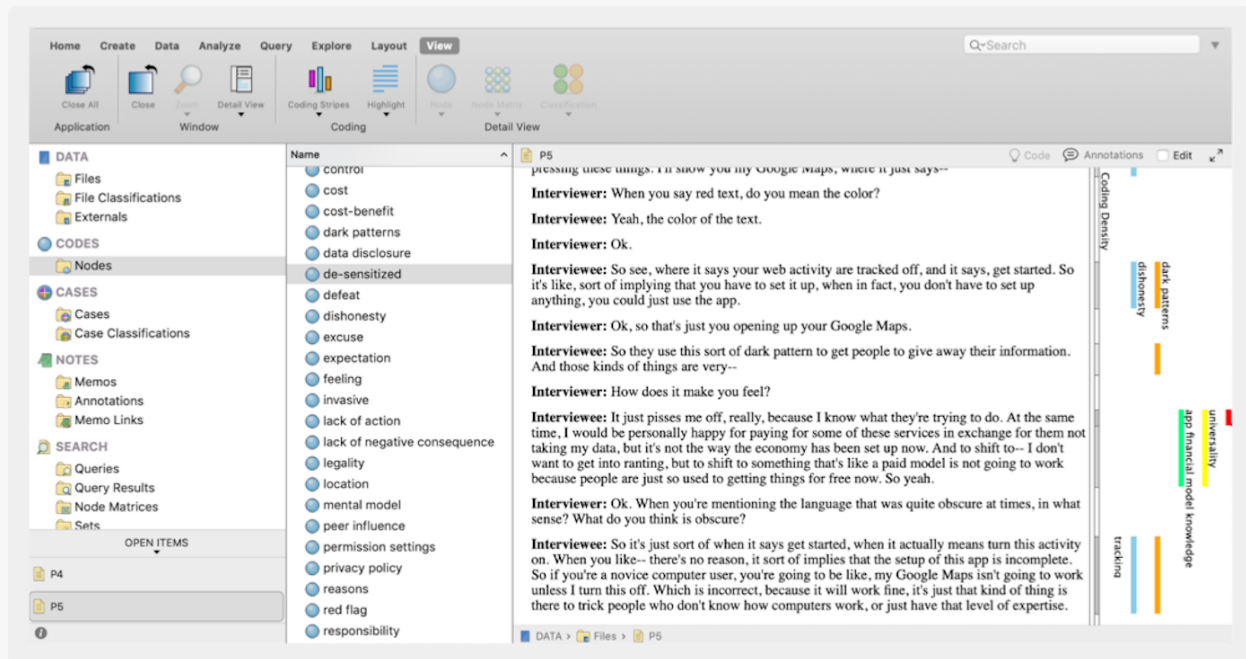
The **code can be created before or after you have grouped the data**. The next two sections of this step describe how and when you may add the codes.

Traditional Method: Create Codes Before Grouping

In the traditional approach, as you highlight segments of the data, like sentences, paragraphs, phrases, you code them. It’s helpful to keep a record of all the codes used and outline what they are, so you can refer to this list when coding further sections of the text (especially if multiple people are coding the text). This approach avoids creating multiple codes (that will later need to be consolidated) for the same type of issue.

Once all the text has been coded, you can group all the data that has the same code.

If you’re using CAQDAS for this process, then the software automatically logs the codes you assign while coding, so you can use them again. It then provides a way for you to view all text coded with the same code.

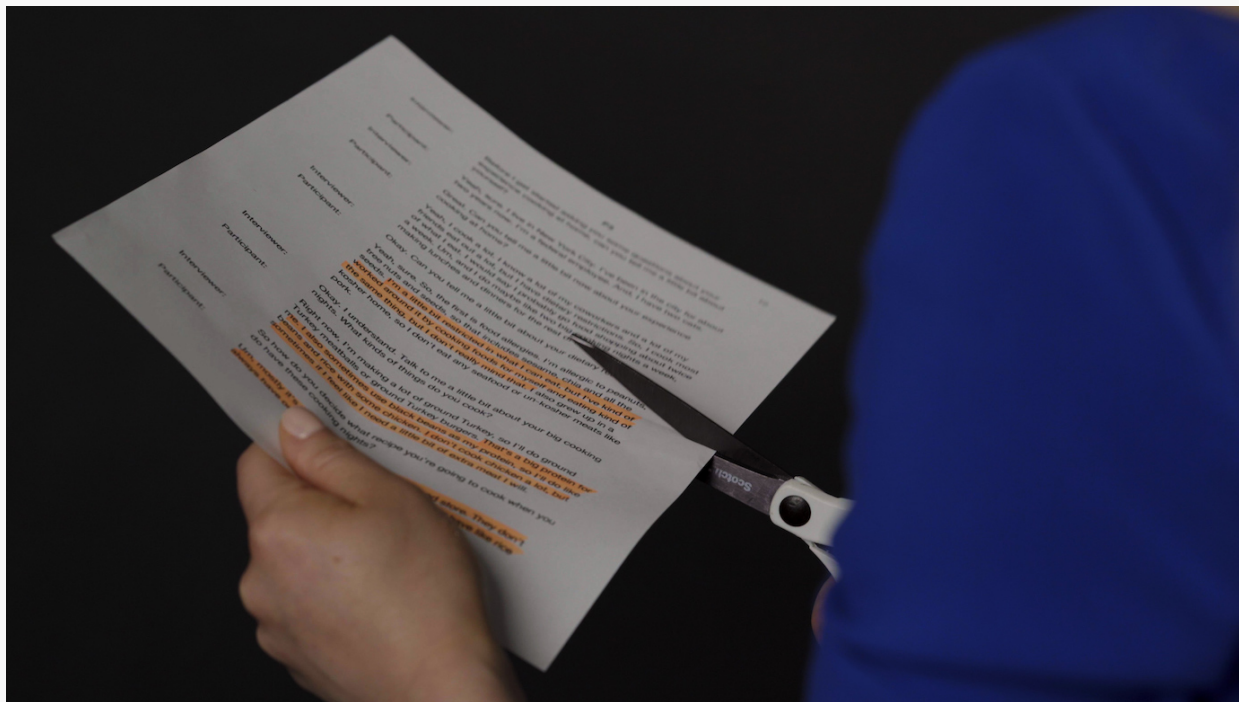


An example from Nvivo (a CAQDAS tool) is shown above. The coding stripes on the right show which parts of the text have been coded. All codes used throughout all the raw data in this project are displayed in the node panel (Nvivo refers to codes as nodes). Double-clicking on a node will display all the raw data coded with this word.

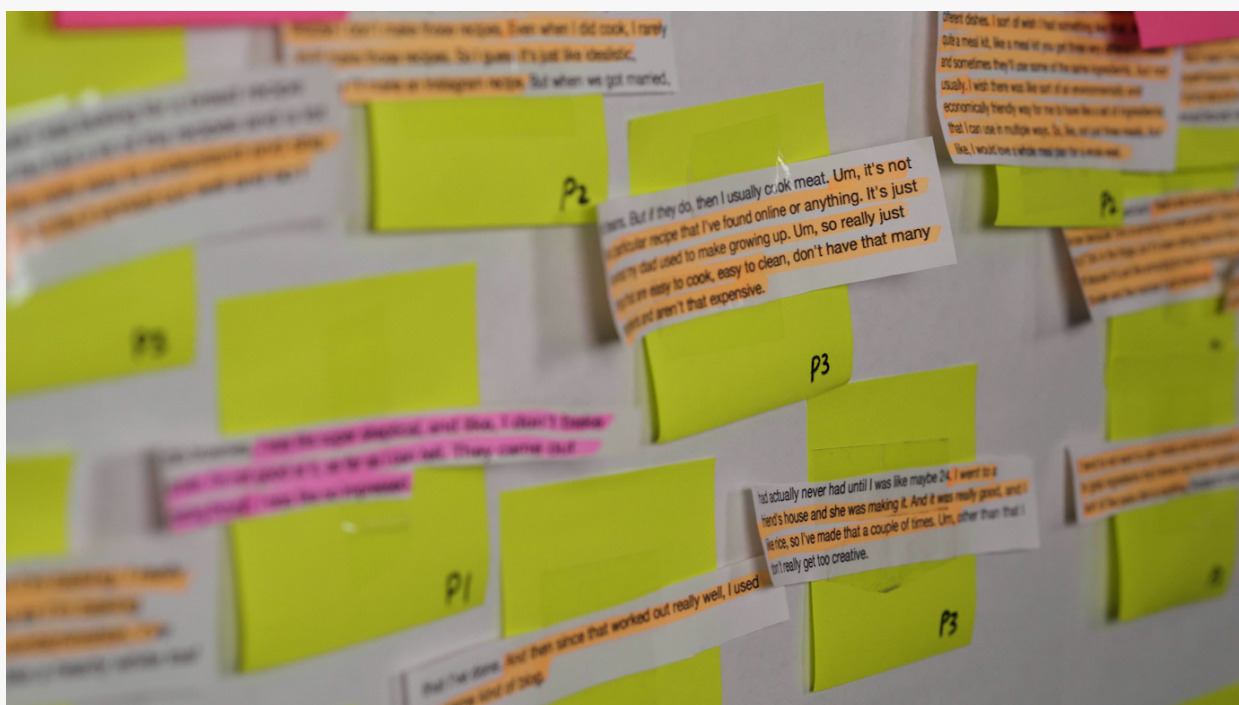
Quick Method: Group Segments of Text, Then Assign a Code

Rather than coming up with a code when you highlight text, you cut up (physically or digitally) and cluster all the similar highlighted segments (similarly to how different stickies may be grouped in an [affinity map](#)). The groupings are then given a code. If you're doing the clustering digitally, you might pull coded sections into a new document or a visual collaboration platform.

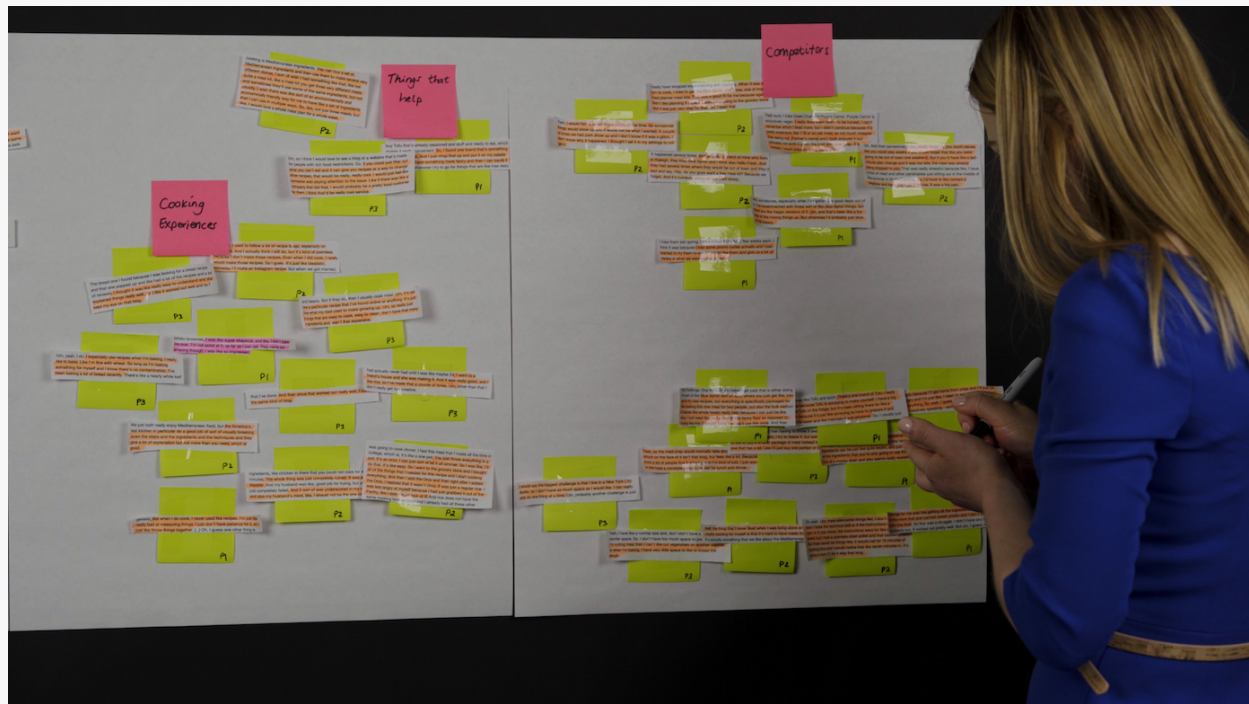
In the pictures below, the grouping was done manually. Transcripts were cut up, fixed to stickies, and moved around the board until they fell into natural topic groups. The researcher then assigned a pink sticky with a descriptive code to the grouping.



The highlighted sections were physically cut up with scissors and taped to stickies.



The participant number or the data type (i.e., interview vs. field study) was written on the sticky (but could also be communicated through the color of the sticky). This practice facilitates an easy return to the full data, as well as comparisons across participants and data sources. Stickies allow the segments of text to be easily moved around a board or wall.



The highlighted segments were clustered by the text topic and given a descriptive code.

At the end of this step, you should have data grouped by topics and codes for each topic.

Let's look at an example. I interviewed 3 people about their experience of cooking at home. In these interviews, participants talked about how they chose to cook certain things and not others. They talked about specific challenges they faced while cooking (e.g., dietary requirements, tight budgets, lack of time and physical space) and about solutions for some of these challenges.

After grouping the highlighted clippings from my interviews by topic, I ended up with 3 broad descriptive codes and corresponding groupings:

- **Cooking experiences:** memorable positive and negative experiences related to cooking
- **Pain points:** anything that stops someone from cooking or makes cooking difficult (including navigating dietary restrictions, limited budgets, etc.)
- **Things that help:** what helps (or is believed to possibly help) someone overcome specific challenges or pain points

Step 4: Create New Codes that Encapsulate Potential Themes

Look across all the codes and explore any causal relationships, similarities, differences, or contradictions to see if you can uncover underlying themes. While doing so, some of the codes will be set aside (either archived or deleted) and new interpretive codes will be created. If you're using a physical-mapping approach like that discussed in step 3, then some of these initial groupings may collapse or expand as you look for themes.

Ask yourself the following questions:

- *What's going on in each group?*
- *How are these codes related?*
- *How do these relate to my research questions?*

Returning to our cooking topic, when analyzing the text within each grouping and looking for relationships between the data, I noticed that two participants said that they liked ingredients that can be prepared in different ways and go well with other different ingredients. A third participant talked about wishing she could have a set of ingredients that can be used for many different meals throughout the week, rather than having to buy separate ingredients for each meal plan. Thus, a new theme about the flexibility of ingredients emerged. For this theme, I came up with the code *one ingredient fits all*, for which I then wrote a detailed description.



In this research example, a new grouping was formed; the grouping included quotes mentioning a need for ingredients that can be flexibly used — either because they can be prepared in several ways or because they can be used in several different meals throughout a week. The grouping was labeled with the interpretive code one ingredient fits all. The researcher then fleshed out the description of this code.

Step 5: Take a Break for a Day, then Return to the Data

It almost always is a good idea to take a break and come back and look at the data with a fresh pair of eyes. Doing so sometimes helps you to see significant patterns in the data clearly and derive breakthrough insights.

Step 6: Evaluate Your Themes for Good Fit

In this step, it can be useful to have others involved to help you review your codes and emerging themes. Not only are new insights drawn out, but your conclusions can be challenged and critiqued by fresh eyes and brains. This practice reduces the potential for your interpretation to be colored by personal biases.

Put your themes under scrutiny. Ask yourself these questions:

- Is the theme well supported by the data? Or could you find data that don't

support your theme?

- Is the theme saturated with lots of instances?
- Do others agree with the themes you have found in the data after analyzing the data separately?

If the answer to these questions is *no*, it might mean that you need to return to the analysis board. Assuming you collected sound data, there is almost always something to be learned, so spending more time with your team repeating steps 4–6 will be worthwhile.

Conclusion

Use thematic analysis as a helpful guide for efficiently wading through lots of qualitative data. There's no one way to do a thematic analysis. Choose a method of analysis that suits the kind and volume of data you've collected. When possible, invite others into the analysis process to both increase the accuracy of the analysis and your team's knowledge of your users' behaviors, motivations, and needs. Analysis can be a lengthy process, so a good rule of thumb is to budget as much time as you had for the data collection to complete the analysis.

Learn more: [User Interviews](#), [Advanced techniques to uncover values, motivations, and desires](#), a full-day course at the UX Conference.

Share this article: [Twitter](#) | [LinkedIn](#) | [Email](#)

Copyright © 1998-2022 Nielsen Norman Group, All Rights Reserved.