



# TECHNIQUES TO IDENTIFY THEMES IN QUALITATIVE DATA

DR. LINDA BLOOMBERG  
ASSOCIATE DIRECTOR FACULTY SUPPORT & DEVELOPMENT AND PROFESSOR  
SCHOOL OF EDUCATION



- **This webinar will introduce doctoral candidates to the following:**
  - ✓ Theme identification as an approach for understanding Qualitative data analysis
  - ✓ various strategies to identify analytic themes in Qualitative data
  - ✓ analytic strategies for addressing different types of Qualitative data.
  
- **Webinar outcomes:**
  - ✓ Understand what constitutes rigorous Qualitative data analysis
  - ✓ Understand theme identification a fundamental task in Qualitative research
  - ✓ Understand specific theme identification strategies
  - ✓ Understand how and in what ways theme identification strategies are applicable to different types of Qualitative data
  - ✓ Receive guidelines and suggestions regarding rigorous Qualitative data analysis techniques.

- Theme identification is one of the most fundamental tasks in qualitative research analysis.
- It also one of the most mysterious processes.
- Explicit descriptions of theme discovery are rarely described in articles and reports and if so usually appear in appendices or footnotes.
- During the proposal-writing phase students often struggle to clearly explain and justify plans for discovering themes as they are as yet unfamiliar with qualitative traditions.

- This webinar presents techniques that are drawn from across social science disciplines.
- These techniques are useful tools to conduct analysis of data.
- These techniques range from quick word counts to in-depth, line-by-line scrutiny.
- Some methods work well for short answers to open-ended survey questions while others are more appropriate for rich, complex narratives produced through interview or focus group transcripts.
- You may find some techniques easier than others.
- No single technique is “does it all”.

- Techniques are based on the following:
  - analysis of words (word repetitions, key-words-in contexts)
  - careful reading of larger blocks of texts (compare & contrast, social science queries, searching for missing information)
  - intentional analysis of linguistic features (metaphors, transitions, connectors)
  - physical manipulation of texts (unmarked texts, pawing, and cut and sort procedures).

# WHERE DO THEMES COME FROM?

- At the heart of qualitative data analysis is the task of discovering themes.
- Themes are abstract, often fuzzy, constructs which researchers identify before, during, and after data collection.
- Themes emanate from the following sources:
  - reviewing the literature
  - characteristics of the phenomena being studied
  - already-agreed-upon professional definitions
  - common-sense constructs
  - researchers' own values, theoretical orientation, and personal and/or professional experience with the subject matter.

# A. WORD-BASED TECHNIQUES

- Word-based techniques are typically fast and efficient ways to start looking for themes.
- This is a particularly useful at early stages of theme identification.
- These techniques are relatively easy for novice researchers to apply.

# 1. WORD REPETITION

- We often begin our data analysis with word-based techniques.
- If you want to understand what people are talking about, we need to look at the words they use.
- Words that occur a lot are often seen as being salient in the minds of research participants.
- Word repetitions can be analyzed informally and formally.
- In the informal mode, researchers simply read the text and note words or synonyms that people use frequently.
- A more formal analysis of word frequencies can be done by generating a list of all the unique words in a text and counting the number of times each occurs.
- Computer programs such as *ATLAS.ti*, NVivo and others allow you to conduct this kind of exercise by way of a software program.



## 2. KEY WORDS IN CONTEXT

- Key-words-in-context (KWIC) is based on a simple observation: if you want to understand a concept, then look at *how* words or phrases are used.
- In this technique, researchers identify key words and then systematically search the text to find all instances of the word or phrase.
- Each time you find a word/phrase, make a copy of the word/phrase and its immediate context.
- Themes are identified by physically sorting the examples into piles of similar meaning.

## B. SCRUTINY-BASED TECHNIQUES

- Word-based techniques are useful.
- Nothing however beats a careful scrutiny of the texts for finding themes that may be more subtle or that are not directly obvious in the text.
- Scrutiny-based techniques are more time-intensive and require a lot of attention to details and nuances.

# 1. COMPARE AND CONTRAST

- The compare and contrast approach is based on the idea that themes represent the ways in which texts are either similar or different from each other.
- Glazer and Strauss (1967) refer to this as the "constant comparison method."
- Typically, Grounded Theorists begin by conducting a careful line-by-line analysis.
- Researchers read each line or sentence and ask themselves, "What is this about?" and "How does it differ from the preceding or following statements?"
- This kind of detailed work keeps the researcher focused on the data themselves.
- This approach is like interviewing the text and is similar to the ethnographic interviewing style where the researcher compares pairs of texts by asking "How is this text different from the preceding text?"; "What kinds of things are mentioned in both?" and "What does this remind me of?"
- Like a good journalist, researchers compare answers to questions across/between research participants.

## 2. SOCIAL SCIENCE QUESTIONS

- Researchers are often interested in understanding how data illuminates questions of importance to social science.
- This includes searching for evidence of social conflict, cultural contradictions, informal methods of social control, things that people do to manage social relationships, methods by which people acquire and maintain achieved and ascribed status, and information about how people solve problems.
- Examine the setting and context, as well as the perspectives of research participants' ways of thinking about people, objects, processes, activities, events, and relationships.
- To do this you will need to be sensitive to conditions, actions/interactions, and consequences of a phenomenon.
- Querying the text is a powerful technique because you are searching for specific kinds of topics that are likely to generate major social and cultural themes.
- By examining the data in this way you must be careful not to *overfit the data* – that is, find only what you are looking for.
- There is a trade-off between bringing a lot of prior theorizing to the theme-identification effort and going at it with an open mind.
- Prior theorizing (including your theoretical framework) can inhibit the forming of fresh ideas and the making of surprising connections.

### 3. SEARCHING FOR MISSING INFORMATION

- The final scrutiny-based approach works in reverse from typical theme identification techniques.
- Of all the scrutiny-based techniques, searching for missing information is the most difficult.
- Instead of identifying themes that emerge from the text, researchers will be searching for themes that are *missing in the text*.
- Much can be learned from a text by what is NOT mentioned.
- Sometimes silences indicate areas that people are unwilling or afraid to discuss sensitive or painful issues.
- Other times, absences may indicate participants' assumptions. When people tell their stories, they sometimes assume that the researcher share their assumptions and so they leave out information that "everyone knows." This is called *abbreviating*.
- There are other reasons why people do not mention topics: People may not fully trust the interviewer, may not wish to speak when others are present, or may not fully understand the researcher's questions.
- Distinguishing between participants' unwillingness to discuss topics and when they assume the researcher already knows about the topic requires familiarity with the subject matter.

- Linguistic features such as metaphors, topical transitions, and keyword connectors to help identify themes. By making these features explicit, you can sharpen your ability to discover themes:
  - Look at patterns of speech and at the repetition of key words and phrases, paying particular attention to participants' use of metaphors and the commonalities in their reasoning.
  - Another linguistic approach is to look for naturally occurring shifts in thematic content. This includes subtle or abrupt shifts in topics, pauses, change in tone, or particular phrases that may indicate themes.
  - A third linguistic approach is to look carefully at words and phrases that indicate relationships among things. For example, causal relationships are often indicated by such words and phrases as *because*, *since*, and *as a result*. Words such as *if* or *then*, *rather than*, and *instead of* often signify conditional relationships. You can discover themes by searching for groups of words and looking to see what kinds of things the words connect.

## D. TACTILE APPROACHES

Each of the next three techniques requires some physical manipulation of the text itself:

# 1. UNMARKED TEXTS

- The first technique requires multiple readings of a text.
- On the first reading, salient themes are clearly visible and can be quickly and readily marked with different colored pencils or highlighters.
- In the next stage, the search is for themes that remain unmarked.
- This tactic—marking obvious themes early and quickly—forces the search for new, and less obvious themes.



## 2. PAWING

- Begin by proofreading the material and underlining key phrases "because they make sense."
- In this method, you get a feel for the text by handling your data multiple times.
- Researchers have been known to spread their texts out on the floor, tack bunches of them to a bulletin board, and sort them into different file folders or boxes.
- By living with the data you wait for patterns to emerge.
- This may not seem like a scientific way to do things, but it is one of the best ways to begin hunting for patterns in qualitative data.
- This method gives you a strong and intuitive feel for themes and the relations among the themes.

## 3. CUTTING AND SORTING

- Cutting and sorting is a more formal way of pawing, and is particularly useful for identifying sub-themes.
- First read through the text and identify quotes that may seem important.
- Cut out each quote (making sure to maintain some of the context in which it occurred) and paste the material on small index cards.
- On the back of each card write down the quote's reference—who said it and where it appeared in the text.
- Sort the quotes into piles of similar quotes, naming each pile.
- These piles become your themes.
- An advantage of this technique is that the data can be used to systematically describe how themes are distributed across participants.
- If you do not opt for the manual approach, cutting and sorting can be done on your computer (cut, copy and paste)
- This approach can be done with the use qualitative software such as AtlasTI, INVivo, and others.



# WHICH TECHNIQUES ARE MOST APPROPRIATE FOR MY STUDY?

- The choice of techniques depends on one or more of the following:
  - kind and amount of text
  - experience of the researcher
  - goals of the project
  
- Techniques range from quick word counts to laborious, in-depth, line-by-line scrutiny. Some work well for short answers to open-ended questions while others are more appropriate for rich, complex narratives.
  
- No single technique does it all!!

- Word-based techniques (word repetitions & KWIC) are the least labor intensive.
- A careful look at the frequency and some quick pile sorts can help to start identifying themes.
- Word-based techniques are versatile. They can also be used relatively easily by novice and expert researchers alike.
- Given their very nature these techniques are best used in combination with other approaches.

- Scrutiny-based techniques (compare & contrast, querying the text, examining absences) are most appropriate for rich narrative data derived from interviews and focus groups
- This approach would be overkill for analyzing short answer responses.
- Researchers who are just beginning to explore a new topical area might want to start with compare-and-contrast techniques before moving on to the more difficult tasks of querying the text or searching for missing information.
- Do not use the latter two techniques unless you are fluent in the language in which the data are collected.
- If the primary goal is to discover as many themes as possible, then nothing beats scrutiny-based techniques on a line-by-line basis.

- Like scrutiny-based techniques, linguist-based approaches are better used on rich narrative data rather than short answer responses.
- Looking for transitions is the easiest technique to use, especially if the texts are actually written by research participants themselves (rather than transcribed from tape recordings of verbal interviews).
- Searching for metaphors is also relatively easy once novices have been trained on what kind of things to look for in the texts.
- Looking for connecting words and phrases is best used as a secondary wave of finding themes, once the researcher has a more definite idea of what kinds of themes he or she finds most interesting.

- In the early stages of exploration, nothing beats a thorough reading and pawing through of the data.
- This approach is the easiest for novice researchers to master and is good for identifying major themes. As the exploration progresses, researchers will often look for subthemes within the major themes.
- The cutting and sorting techniques help researchers to identify all text passages that are related to a major theme, cut them out, and sort them into sub-thematic categories.
- If you are marking texts for each newly discovered theme, you can apply the unmarked text technique as you go.
- These three techniques can be applied successfully to both rich narrative data and more simple responses to open-ended questions.



- A powerful strategy is to combine multiple techniques in a sequential manner.
- For example, you might begin by pawing through the data to see what kinds of themes stand out. As part of this process, you might want to make comparisons between paragraphs and across participants. A quick analysis of word repetitions would also be appropriate for identifying themes at an early stage of the analysis. If key words are present, researchers might follow-up by conducting more focused KWIC analyses. Texts representing major themes can be marked either on paper or by computer. Investigators can then search areas that are not already marked for additional themes or cut and sort marked texts into subthemes. You could also consider identifying metaphors, marking them, cutting them out, and sorting them into thematic categories.
- There is no single way to discover themes. In theme discovery, please assume that more is always better than less!

# HOW DO I KNOW WHEN I HAVE FOUND ALL THE THEMES?

- There is no magic formula to answer this question.
- This is similar to asking members of a population to list all the illnesses they know. One can never be sure of the full range of illnesses without interviewing the entire population. This is true because there is always the possibility that the last person interviewed will mention a new disease. We can simplify the process considerably, however, if we are willing to miss rarely-mentioned illness. One strategy would be to interview people until some number of respondents in a row fail to mention any new illnesses.
- In text analysis, grounded theorists refer to the point at which no new themes are being identified as *theoretical saturation* (Strauss and Corbin 1990). When and how theoretical saturation is reached, however, depends on the number of texts and their complexity, as well as on researcher experience and fatigue, and the number of investigators examining the texts.
- Be aware of *premature closure* where the researcher fails to move beyond the face value of the content in the narrative.



