



# What Do I Do With It Now? Beginning and Advanced Theme Analyses

## Techniques for Open-Ended Survey Data

Carol VanZile-Tamsen

University at Buffalo, The State University of New York

REACHING OTHERS

### ABSTRACT

Open-ended survey items are often never analyzed. This poster outlines a simple process of thematic analysis for such items and demonstrates how to derive the "big picture" to effectively inform campus decision-making. This process involves the use of Microsoft productivity software that most researchers already own. Areas of emphasis include theme refinement, identification of relationships among themes, group comparisons, and presentation of results.

### OVERVIEW OF THEME ANALYSIS

Many surveys contain open-ended items that allow free-form text entry. These items are meant to allow respondents to freely communicate in their own words what is important about an issue or topic. Researchers are often interested in summarizing this information in some way in order to learn more about the issue or topic they are surveying.

In truth, however, this data is often ignored. Qualitative analysis is seen as a time-consuming and labor-intensive endeavor with little to add to results of quantitative analysis. This presentation outlines a simple method for conducting theme analysis with open-ended data.

Themes represent the most important issues that are mentioned many times over by the respondents. Identification of themes can help support the typical quantitative analysis done on most survey items in several ways. First, the resulting themes can help the researcher gain a deeper understanding of complex issues. Second, the themes can provide a context within which to frame statistical analyses of survey responses. The combination of the results of theme analysis and with the results of statistical analysis falls under the Mixed-Methods Approach to research.

### CONDUCTING THEME ANALYSIS

The responses to an open-ended item will determine if qualitative analysis of responses is necessary or even possible. In general, theme analysis of open-ended items is necessary when:

- Most responses are complex and cannot be easily distilled into a single idea (or numeric code); and
- Forcing the responses into a single thought results in the loss of context and meaningful clarification.
- Generally, "other, specify" items do not qualify for qualitative analysis.

When the open-ended data is sufficiently complex to warrant qualitative analysis, two types of theme analysis can be employed. In **Emergent Theme Analysis**, the analyst has no pre-conceived notions about the categorization of responses. Instead, these categorizations are developed based on the most common responses. In **Convergent Theme Analysis**, the analyst has a pre-existing set of categories into which the open-ended data is meant to be sorted. This method may be used when the topic area is already well-researched, and the current study is viewed to be confirmatory in nature.

Once the researcher has determined that theme analysis is possible and has chosen to proceed with either Emergent or Convergent Theme Analysis, the following steps are used to identify salient themes from the responses:

### CONDUCTING THEME ANALYSIS (continued)

#### Step 1: Compiling and Organizing Data

- It is easiest to work with data in an electronic form, preferably in Excel or some other spreadsheet program.
- If data has been collected via paper and pencil surveys, it should be typed into Excel with one row per respondent, maintaining an identifier, and one worksheet tab per open-ended question.
- Sort the responses alphabetically and add a column for theme name(s).
- The identifier helps link responses back to other key data elements.
- Although tempting to ask a student or an administrative assistant to engage in this largely mechanical step, it can provide an important orientation to the data that makes subsequent steps easier.

#### Step 2: First Read-Through

- Beginning qualitative analysis is often the hardest part since the quantity of information can seem overwhelming.
- Additional anxiety is created if one is reading comments about one's own program.
- The first read-through allows the researcher to begin analysis without really beginning: themes are beginning to be identified mentally but nothing is yet set in stone.
- If anxiety sets in, it is easy to say, "I only have to read these now."
- Underlining and note-taking is okay if it doesn't get in the way of the reading.

#### Step 3: Categorizing into Themes

- Begin anywhere in the data and "name" the ideas that the responses seem to represent.
- Type these names in the theme column and remember that one response may represent several themes.
- Consistency in theme name is not essential at this point since it will be checked in a later step.
- Having the responses sorted alphabetically may result in the grouping of responses that start with similar words to reflect similar ideas.
- If the first read-through suggests that some words and phrases are used many times, the Find button can be used to quickly locate occurrences.

#### Step 4: Organizing Data by Themes

- In this step, responses are reorganized by themes.
- It is best to copy and past the coded data into a new tab so that raw data and initial themes are saved intact.
- A new sorting scheme will be used to group all responses representing a single them and allows a consistency check.
- Any responses representing more than one theme should be included with the comments for each theme.
- A hand-written version of a theme table might look like this:

Theme Name	Original	Thematic Comments	Comments
General	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		

#### Step 5a: Sub-Theme Analysis

- If there are complexities of the data that are not captured by the primary organizing theme, sub-themes are likely required.
- For the theme table presented above, a sub-theme analysis might look like this:

Theme Name	Original	Thematic Comments	Comments
General	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		
Specific	1) I was surprised!		

### CONDUCTING THEME ANALYSIS (continued)

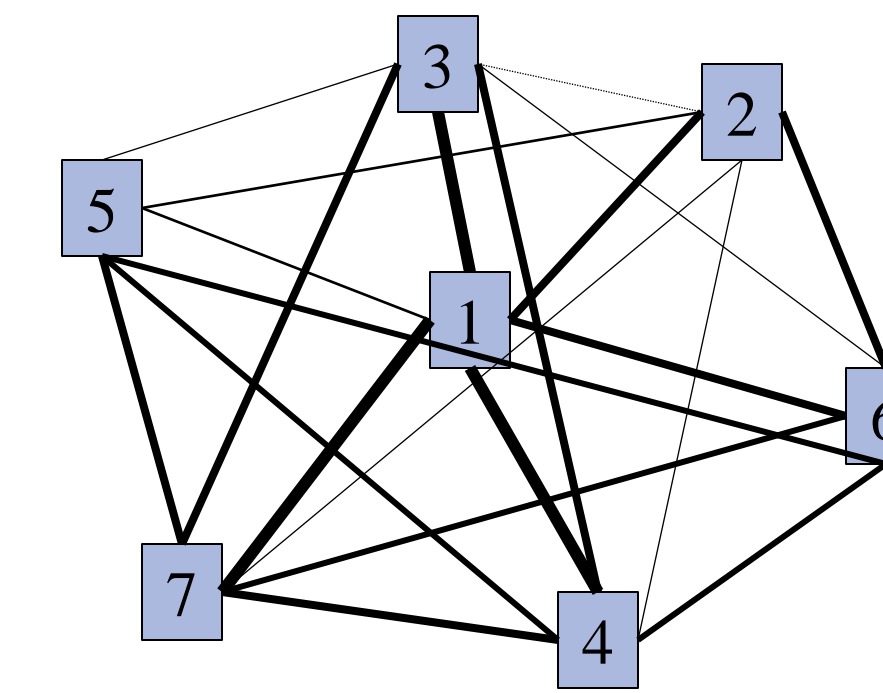
#### Step 5b: Higher-Order Theme Analysis

- In general, the more themes identified in the data, the more difficult it is to develop a "big picture."
- The theme analysis can be simplified by identifying Higher-Order Themes based on theory, if relevant, or the data itself.
- When two or more themes generally appear together in respondent comments, there is likely a higher-order them involved.
- Graphical displays can help in the identification process:
  - Qualitative Correlation Matrix

Theme	1	2	3	4	5	6	7
1 Complete Text Quizzes							
2 Relate Material to Other Classes							
3 Memorize Notes							
4 Mnemonic Devices							
5 Read/Highlight Text							
6 Ask Self "How" and "Why" Questions							
7 Compile Own Notes with Lecture Outline							

Sample data based on N=30. "Describe your study techniques for exams in this class."

#### Web Diagram



#### Step 6a: Identifying Relationships Among Themes

- Once all themes (higher-order, sub-themes, etc.) have been identified, relationships among themes may be noted.
- Types of Relationships:
  - Time-Ordered or Process-Oriented;
  - Role-Ordered;
  - Conceptual Networks; and
  - Cause and Effect.
- Different organizational schemes can be used to determine if there is some relationship among themes (see Miles & Huberman, 1994).
- In general, start with the highest level of themes identified to map relationships.

#### Step 6b: Group Differences Among Themes

- Cross-case analysis may be used to identify group differences.
- To clarify, role-ordered analyses are inherently cross-case analyses, but there may be additional analyses within roles (e.g., gender groups, age groups, ethnic groups, etc.)
- In essence, the resulting theme analysis is compared among different groups.

#### Step 7: Validation of Themes

- Validation can be done when analysis is complete or it can be built into the analysis process.
- When validating during analysis, multiple categorizers are used and Inter-Rater Reliability calculated:
  - <http://med-ed-online/rating/reliability.html>, or
  - Use Crosstabs in SPSS and choose the Cohen's Kappa statistic.
- Validation after the analysis involves either:
  - Member Checking, where respondents review and "sign-off" on results, or
  - Peer Review, where colleagues "audit" analysis procedures and results

### PRESENTING RESULTS OF THEME ANALYSIS

Sometimes, just knowing the results of theme analysis may be enough, but usually a report is expected. Traditional or mixed methods reports are similar to quantitative reports but use participant quotes and major qualitative findings to emphasize key points. A qualitative narrative presents results in the words of participants and clearly describes the process used to derive results.

For administrative audiences, the report should describe *major* theme areas that provide the "big picture." In describing qualitative analysis, it is easy to get carried away with details. Authors can include contact information to provide details upon request. In addition, well-designed graphical representations can be used to focus reader attention on key points. Finally, Executive Summaries, using bullet points to present key themes, can be more effective than long, detailed reports.

### REFERENCES

Aronson, J. (1994). A pragmatic view of thematic analysis. *The Qualitative Report*, 2 (1), np.

Carney, J. H., Joiner, J. F., & Tragou, H. (1997). Categorizing, coding, and manipulating qualitative data using the WordPerfect word processor. *The Qualitative Report*, 3 (1), np.

Chenail, R. J. (1995). Presenting qualitative data. *The Qualitative Report*, 2 (3), np.

Flick, U. (2007). *Designing qualitative research*. Los Angeles, CA: Sage Publications.

Flick, U. (2007). *Managing quality in qualitative research*. Los Angeles, CA: Sage Publications.

Gibbs, G. (2007). *Analyzing qualitative data*. Los Angeles, CA: Sage Publications.

Gowan, T. (2009). New hobos or Neo-Romantic fantasy? Urban ethnography beyond the Neoliberal disconnect. *Qualitative Sociology*, 32 (3), 231-257.

LeCompte, M. D., & Preissle, J. (1993). *Ethnography and qualitative design in educational research (2nd edition)*. New York, NY: Academic Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.

Livingston, J. A., Buddie, A., Testa, M., & VanZile-Tamsen, C. (2004). The role of sexual precedence in verbal sexual coercion. *Psychology of Women Quarterly*, 28, 287-297.

Livingston, J. A., Hequebourg, A., Testa, M., & VanZile-Tamsen, C. (2007). Unique aspects of adolescent sexual victimization experiences. *Psychology of Women Quarterly*, 31, 331-343.

Matute-Bianchi, M. E. (1986). Ethnic identities and patterns of school success and failure among Mexican-descent and Japanese-American students in a California high school: An ethnographic analysis. *American Journal of Education*, 95 (1), 233-255.

Michalski, G. V. (2011, May). *Complementing the numbers: A text mining analysis of college course withdrawals*. Paper presented at the Annual Forum of the Association of Institutional Research, Toronto, CA.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage Publications.

Onwuegbuzie, A. J., Leech, N. L. (2006). Linking research questions to mixed methods data analysis procedures. *The Qualitative Report*, 11 (3), 474-498.

Plano Clark, V. C., & Creswell, J. W. (2008). *The mixed methods reader*. Los Angeles, CA: Sage Publications.

Scott, K. W. (2004). Relating categories in grounded theory analysis: Using a conditional relationship guide and reflective coding matrix. *The Qualitative Report*, 9 (1), 113-126.

Testa, M., VanZile-Tamsen, C., Livingston, J. A., & Koss, M. P. (2004). Assessing women's experiences of sexual assault aggression using the sexual experiences survey: Evidence for validity and implications for research. *Psychology of Women Quarterly*, 28, 256-265.

VanZile-Tamsen, C. (1996). Metacognitive self-regulation and the daily academic activities of college students. *Dissertation Abstracts International: Section A, The Humanities and Social Sciences*, 57 (6), 2361.

VanZile-Tamsen, C. (1999). Transforming teacher education classes: Lessons from educational psychology. *Research in the Schools*, 6 (2), 33-43.

Wolcott, H. F. (2001). *Writing up qualitative research (2nd edition)*. Thousand Oaks, CA: Sage Publications.

Yeh, C. J., & Inman, A. G. (2007). Qualitative data analysis and interpretation in counseling psychology: Strategies for best practices. *The Counseling Psychologist*, 35, 369-403.

### CONTACT

Carol VanZile-Tamsen  
Research Analyst  
Office of Institutional Analysis  
536A Capen Hall  
Buffalo, NY 14260  
(716)645-3905  
cmv3@buffalo.edu