**Qualitative Research II: Data Analysis**

***C. Stanton***



**From** Beddoes, K. & Borrego, M. (2011). Feminist theory in three engineering education journals: 1995-2008. *Journal of Engineering Education, 100*(2), 285.

**See** Saldana, J. (2016). *The coding manual for qualitative researchers* (3rd Ed.)*.* Thousand Oaks, CA: SAGE. Chapter 1 (1st Ed.) available at: <https://www.sagepub.com/sites/default/files/upm-binaries/24614_01_Saldana_Ch_01.pdf>. Summary available at: <http://salmapatel.co.uk/academia/coding-qualitative-research>

**Code**: A “word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data” (p. 4)

**Category**: A synthesis of similar codes to refine and organize data for further analysis

**Theme:** An interrelated set of codes and/or categories expressed in a succinct word or phrase and aligned with and/or generative of theory



(Saldana, 2016, p. 14)

A **coding pattern** can be characterized by:

* similarity (things happen the same way)
* difference (they happen in different ways)
* frequency (they happen often or seldom)
* sequence (they happen in a certain order)
* correspondence (they happen in relation to other activities or events)
* causation (one appears to cause another)

**Questions to guide coding:**

* Who is represented? What are they doing? What are they trying to accomplish? Where? When? Why?
* How, exactly, do they do this? What specific means and/or strategies do they use?
* How do members talk about, characterize, and understand what is going on?
* What assumptions are they making? What evidence demonstrates these assumptions?

**Questions to guide researcher memos:**

* What do I see going on here?
* What did I learn from these notes?
* What surprised/intrigued me?
* What disturbed me?

**Step One: First Cycle of Coding**

The first cycle of coding typically breaks down qualitative data into discrete parts so they can be further examined and compared. Researchers can use open/descriptive coding, holistic coding,

**Open/Descriptive Coding** (the most popular approach) summarizes the basic topic(s) of each segment of data (e.g. words, phrases, sentences, paragraphs, single images, moments in observations, etc.). Some specific approaches to open coding include:

* **In Vivo Coding** uses a word or short phrase from the actual language found in the data to create a code (e.g. “Not Girlie”).
* **Process Coding** uses gerunds (“-ing” words) exclusively to connote action in the data to create a code (e.g. “Doing vs. Being”).
* **Affective Coding** uses emotions, values, or conflicts to create a code (e.g. “Skepticism”, “Resisting Social Norms”).
* **Attribute Coding** identifies basic information (e.g. demographic data).
* **Magnitude Coding** describes characteristics such as intensity or frequency (e.g. Strongly Supportive of X, Moderately Supportive of X).
* **Discursive Coding** analyzes grammatical choices (e.g. use of active vs. passive voice).

**Holistic Coding** applies a single code to large units of data (e.g. whole pages or interviews) to capture a sense of the overall contents and categories that may develop. Usually, this approach is followed by a round of more detailed open/descriptive coding.

**Protocol Coding** uses a list of previously generated codes (a “codebook”) based on existing research or theory to guide coding. Like open/descriptive coding, the researcher goes line-by-line through the data, but instead of allowing codes to “emerge” from the data segments, the researcher uses the pre-established protocol to guide the process of assigning codes.

**Step Two: Organizing and Categorizing Codes**

Next, the researcher organizes the codes (by frequency, similarities/differences, patterns, magnitude, etc.) in order to generate larger categories. Depending on your methodology, these categories might span participants, data sources (e.g. interviews, observations, content analysis), and/or time.

Strategies include:

* Highlighting/Color Coding
* Using Post-It Notes
* Visual Mapping (e.g. Wordle, concept mapping, flowcharts, models)
* Word Searches and Cut/Paste (in Word document)
* Creating Matrices for Comparisons
* Using Qualitative Data Software (e.g. nVivo)

**Step Three: Second Cycle of Coding**

The second round of coding allows the researcher to refine codes and categories in order to develop themes and align data with research and theory (note: these themes may be pre-established through research or theory in the form of typologies, domains, tenets, etc.). Following this phase, researchers may complete another round of organization and categorization. Approaches to second cycle coding include:

* **Pattern coding,** which groups similar codes/categories into a smaller number of sets, themes, or constructs.
* **Focused coding,** which searches for the most frequent or significant codes/categories to create or align with conceptual themes.
* **Axial coding,** which explores how codes, categories, and subcategories relate to each other.

**Step Four: Adding to Theory**

Qualitative educational research is not typically viewed as generalizable (i.e. it is not meant to be methodologically or pedagogically prescriptive across contexts). As qualitative educational researchers, our main goal is to build upon theory, including that which shapes research methodology and teaching pedagogy. The expanded/refined theory can then be modified and applied across contexts.

After creating codes, categories, and themes, qualitative educational researchers identify implications for theory, research, and practice. In addition, referring to researcher memos can be important for this phase, as they help provide a check on researcher bias and assumptions.

Finally, checks for trustworthiness should include member checks (i.e. verifying codes, categories, and themes with participants) and/or expert verification (i.e. having a colleague or more experienced researcher check the codes, categories, and themes).

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Analysis** | **Example Data** | **Example Analysis** |
| *Interviews* (individual, focus groups) | * Open coding (attributes, structures, descriptions)
* Axial coding (values, connections to other coded data)
* Focused coding (categorization & connections to theory)
 | “Growing up, I always wanted to make things, but all of my friends wanted to play with Barbies. It wasn’t a ‘girlie’ thing to do to actually *do* something.” (Ann) | Open Codes: childhood connection, motivation, social norms, doing = masculineCategories: “Doing” vs. BeingFocused Codes: Normalized PassivityTheory: Interactional Feminism |
| *Observations* (field notes from observations in classrooms, labs, communities; videos/photos; researcher memos) | Coding (see above) | See Powerpoint imageMemos: “girls’ team had max distance”; “boys’ confounded looks” | Open Codes: masculine body language (crossed arms, closer look, strong stance), women leadingCategories: Leading, Observing, “Doing” vs. BeingFocused Codes: Challenging “Leadership” Theory: Standpoint Feminism |
| *Content, text, discourse, artifacts*(classroom talk, images, textbooks, policies) | Coding (see above) | See Powerpoint images | Open Codes: male (7), female (2), crossed arms (2), low angle shot (4), high angle shot (1), doing (6), collaboration (1)Categories: Representations of Women, Representations of Men, “Doing” vs. BeingFocused Codes: Masculinity and PowerTheory: Masculinity Studies  |