**09 Analyzing Open-ended Response Data with Qualitative Data Analysis**

**Note – add discussion of credibility of data – having participants check results**

**1. Open-ended Questionnaire Items**

Open-ended questionnaire items require data analysis that can handle text well. Basic qualitative data analysis (QDA) is well suited to analyzing responses to open-ended items.

Example questionnaire with open-ended items:

<http://www.bwgriffin.com/gsu/courses/edur9131/activities/sample_survey_open_ended_authorship.pdf>

**2. Generic Steps for Qualitative Data Analysis (QDA)**

LeCompte (2000) likens QDA to assembling pieces of a jigsaw puzzle.

* Many pieces to the puzzle – the raw text of responses to open-ended items
* Sort pieces into common piles – read responses and identify common responses
* Form themes of puzzle (e.g., sky, barn, water, flowers) – do the same for responses (e.g., anxiety, confidence, frustration)
* Find linking pieces of puzzle to connect themes – determine how response themes relate (e.g., when I experience frustration and I also experience anxiety)

**2a. Data Preparation**

Since most qualitative data are in the form of words, it is important that interviews, field notes, documents, etc. be transcribed and recorded in such a way that can be easily accessed and read.

First note that data analysis in qualitative research is often cyclical and may, perhaps should, begin once data collection commences. The cycle of collecting data and analyzing data during the data collection phase is known as interim analysis (analyzing data during the interim while data collection continues). Beginning data analysis early can help identify important themes or areas that should be explored.

At this initial stage researchers should read all their data carefully, and then re-read, then repeat again (and again). Why? The more familiar researchers are with their data, the more easily they can begin spotting or identifying important concepts in those data and see connections between concepts. With each reading researchers should record their impressions of the data, record their thoughts and interpretation of the data. These recordings will help build one’s memory and provide insight when sorting/collecting data into broad categories and concepts.

LeCompte (2000, p. 148) suggests one use the following in preparation for QDA (if not using computer analysis systems):

* Make copies of all data so none is lost or ruined when memo-ing (adding researcher comments/notes to data)
* Put all notes and interviews in files by date of creation
* Create other files based on
  + types of data (e.g., interviews, questionnaires, field notes, artifacts),
  + participants (e.g., students, teachers, staff),
  + organizations (e.g., health agencies, foundations, schools)
  + subject or topic (e.g., recruitment of students, parent involvement);
  + do the above based upon needs and what seems reasonable.
* Catalog and store all documents and artifacts
* Label all files and boxes according to their contents.
* Create index or table for all contents for all data.
* Review research questions comparing them against data collected to ensure each question is addressed.
* Identify holes in data collection and address missing data so research questions can be answered.
* Collect additional data if needed.

**2b. Develop Initial Codes and Code Data**

At this stage the researcher will begin coding data; this means labeling relevant or important data points with unique labels to help separate data into unique and meaningful components. The researcher, when coding, is attempting to identify key ideas, behaviors, interactions, incidents, and terminology/phrases available in the data. In short, coding is labeling or naming things found in one’s data.

Codes used for labeling data may be derived in several ways:

**Deductive/A priori/Preset Codes** – Researcher develops a classification scheme of codes prior to collecting data. This approach may not allow important new information to be identified; probably few qualitative researchers employ this approach although can be a useful approach is one is interested in theory testing.

**Inductive/Post hoc/Emergent Codes** – Codes for classifying data are developed while reading and coding the data. This approach allows data to speak and potentially enables the richness of the data to be revealed.

**Mixed Preset and Emergent Codes** – This approach represents a combination of the two in which researchers develop an initial classification scheme with codes, but adds to these codes as new information is learned. Likely a common approach for many researchers.

Coding data and developing codes is an iterative process and requires much time and effort. When data from multiple interviews or long interviews are used, one can expect this coding process to last many hours or even days.

LeCompte (2000, p. 148) writes that researchers usually use three approaches to identifying things to code or name:

* Frequency – items are coded because they appear often (e.g., how many students expressed some form of anxiety, or how many students indicated the instructor is disorganized)
* Declaration – items are important because participants tell us they are important (e.g. students tell us the instructor’s videos were very helpful)
* Omission – something expected did not occur, why and what does this mean (e.g., students never mention being assessed or tested); this approach probably only works when using some frame of reference to set expectations

Code Sheet Example:

<http://www.bwgriffin.com/gsu/courses/edur9131/activities/sample_code_sheet_open_ended_authorship.pdf>

This sheet was printed twice for each returned for questionnaire, and used by two coders separately and independently to code responses. Once completed, both were attached to the questionnaire and then responses were compared to assess inter-coder agreement levels.

Coded Examples:

<http://www.bwgriffin.com/gsu/courses/edur9131/activities/open_ended_coding_example_1140.pdf>

<http://www.bwgriffin.com/gsu/courses/edur9131/activities/open_ended_coding_example_420.pdf>

**2c. Organize Data into Categories**

At this stage most data will be identify via codes (although the process is iterative so new codes may be identified still or data may be labeled or relabeled with existing codes), so now the process combining like codes into categories begins. Here one attempts to identify redundancies in codes and create subsets of codes to form broader categories of data. This reduction process helps to bring meaning to data; it allows one to more succinctly grasp key ideas found in the data.

One approach to determine unique codes is to compare and contrast data, and to sort items (units of data) into similar and dissimilar groupings.

**2d. Further Refinement: Categories to Themes/Concepts/Taxonomies**

In many cases one will be able to organize categories into still boarder themes/concepts. Sometimes this may not be possible, or categories may be themes/concepts (the two overlap). The notion, however, is that if there are many categories of data, it may be possible to further combine these into more general concepts that better reveal important information or meaning in the data. At this point some categories may be discarded as unimportant or because these categories provide little relevant, helpful information for telling the story of this research.

**2e. Find Relations among Concepts and Categories/Themes/Taxonomies**

Often one may be able to identify how various themes interrelate for study participants and researchers. This can lead to significant meaning and reveal important findings.

**2f. Displaying Results**

**Textual Display**

Most qualitative researchers present results in textual format; they describe the study setting, their perspectives and biases, summary of what they found, and often supplement this with quotations. I illustrate some of this below in “**3. Illustrated Example of Data Analysis**” especially in **3d** and **3e**.

**Graphical Display**

Sometimes textual results are also coupled with graphical displays. Kodish and Gittelsohn (2011) present a graphical display of data results from QDA showing linkages found in diabetes study.

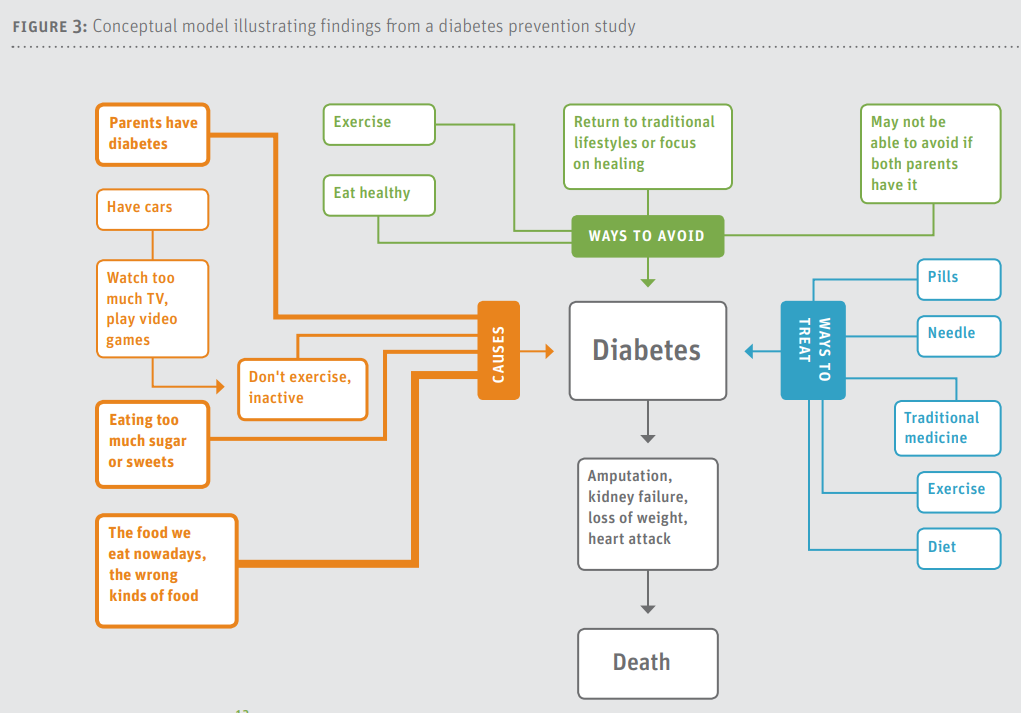
In this example there are Categories which form Themes:

Categories –

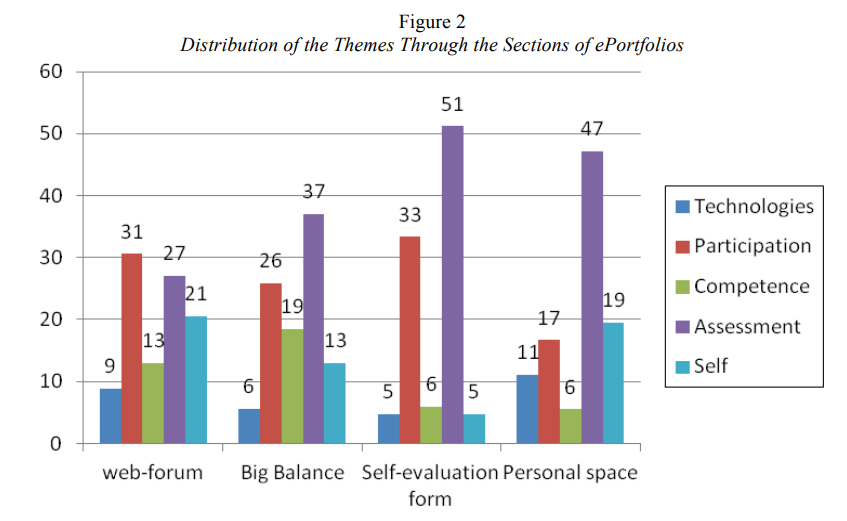
* Items displayed on the outside, examples
  + Parents have diabetes
  + Exercise
  + Pills

Themes –

* Causes
* Don’t Exercise/Inactive
* Ways to Avoid
* Ways to Treat

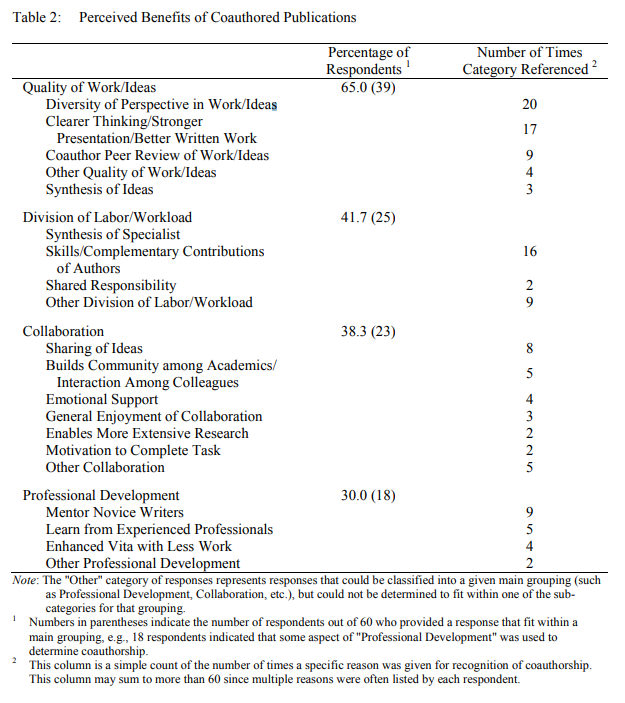


Impedovo, Ritella, and Ligorio (2013) provide the following bar chart showing frequency of themes for different sections of e-portfolios examined. The X-axis contains four sections of the e-portfolio and the labels to the right are the data themes.



**Tabular Display**

Another approach to displaying data is in tables. Below is an example from Moore and Griffin (2006) who asked participants to identify the benefits of co-authoring research.



**3. Illustrated Example of Data Analysis**

Below is an example showing how data from an interview may be coded to identify important concepts.

**3a. Data Preparation**

The interview between a researcher and teacher is transcribed and presented below.

Interviewer:

"Please tell me what it was like to work under your previous principal, and how is it different with your current principal."

Teacher:

"Wow, the difference is like night and day. Mr. Sykes was so controlling. He had to approve . . . like everything we did. If I wanted to try something new, maybe a new computer program with my students, I had to get his permission. Uh, I remember once . . . I wanted to try a new workbook that I downloaded from the internet. One day he was observing me teach, and he asked about it in kind of a disapproving way. I don’t know, but it just seems he was so oppressive and didn’t want us to try new things. I kind of lost interest in teaching when he was here."

"Now with Mr. Rosen things are so different. When he first arrived, he said he wanted us to explore new ways of teaching, you know, to try different things in the classroom. I wasn’t sure I believed him. So, I asked him about using some stuff I found on the web and he said ‘Sure, go for it and let me know if it works.’ So, I did and now I am constantly trying new things to help my students. It is much more exciting now to be in the classroom because I can teach the way I want."

**3b. Develop Initial Codes and Code Data**

Using an emergent design, note following codes within brackets [ ] and highlighted in yellow were added to the teacher’s transcribed response. This is known as open-coding (reading through data and applying/developing codes).

"Wow, the difference is like night and day. Mr. Sykes was so controlling. [authority control] He had to approve . . . like everything we did. [authority approval] If I wanted to try something new, maybe a new computer program with my students, I had to get his permission. [authority permission] Uh, I remember once . . . I wanted to try a new workbook that I downloaded from the internet. One day he was observing me teach, and he asked about it in kind of a disapproving way. [authority questioning] I don’t know, but it just seems he was so oppressive and didn’t want us to try new things. [authority oppressive] I kind of lost interest in teaching when he was here. [teacher interest lost]"

"Now with Mr. Rosen things are so different. When he first arrived, he said he wanted us to explore new ways of teaching, [authority explore] you know, to try different things in the classroom. [classroom experiment] I wasn’t sure I believed him. [teacher disbelief] So, I asked him about using some stuff I found on the web and he said ‘Sure, go for it and let me know if it works.’ [authority approval/flexibility] So, I did and now I am constantly trying new things to help my students. [teacher explore] It is much more exciting now to be in the classroom because I can teach the way I want. [teacher excited/freedom]"

**3c. Organize Data into Categories**

Below I attempt to organize the codes identified above into categories. Note there is some overlap and also one code is not used, [teacher disbelief], because it does not seem relevant to the categories that are emerging.

Control

* authority control
* authority approval
* authority permission
* authority questioning
* authority oppressive
* authority approval/flexibility

Autonomy

* teacher explore
* authority approval/flexibility
* authority control

Motivation

* teacher interest lost
* teacher excited/freedom

**3d. Further Refinement: Categories to Themes/Concepts**

As I think about this teacher’s responses and descriptions, and review those considering the codes and categories presented above, it seems there are two general concepts presented in these data:

**Teacher Autonomy Support** – In what ways does this teacher believe she is supported to be autonomous in her classroom? In what ways does she believe her level of classroom autonomy is being suppressed or controlled?

**Teacher Motivation** – While this is a category identified in the previous step, I think it also represents an important concept for this teacher. Does this teacher demonstrate motivation or de-motivation to teach in her classroom?

**3e. Find Relations among Concepts and Categories**

Given this teacher’s responses it appears that the level of autonomy support offered by the principals directly influenced her motivation to teach. As evidence, consider her statements:

"Mr. Sykes was so controlling. He had to approve . . . like everything we did."

"I kind of lost interest in teaching when he was here."

"Now with Mr. Rosen things are so different. When he first arrived, he said he wanted us to explore new ways of teaching, you know, to try different things in the classroom."

"It is much more exciting now to be in the classroom because I can teach the way I want."

**4. Codebook**

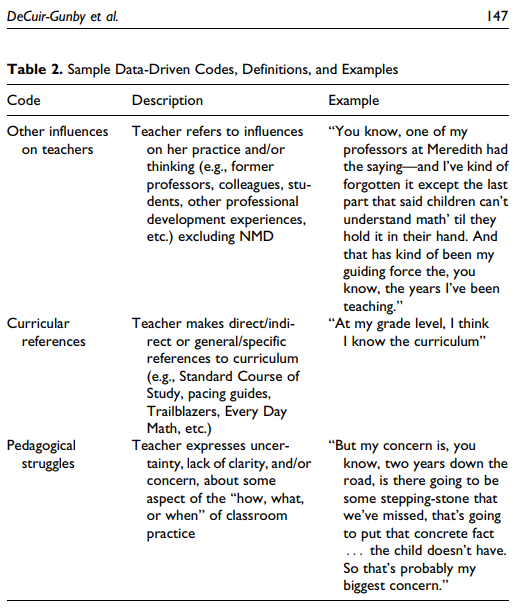
When developing codes, it is important to create a codebook to help clarify what each code represents. The codebook should contain at least three things:

* Code (e.g., Authority Control)
* Description of Code (e.g., Authority Control – individual in authority exerts control by …)
* Example of Code (e.g., Provide quotation from data that illustrates Authority Control)
* Parameters for Use (e.g., this explains when it should and should not be used; Authority Control is not applicable if someone who is not an authority attempts to control a teacher’s behavior such as a peer teacher.)
* Number or Abbreviated Letters (shorthand coding, e.g., Authority Control = AC or 3.10, etc.) – this simply makes coding large chunks of data faster as compared with writing the code label repeatedly.

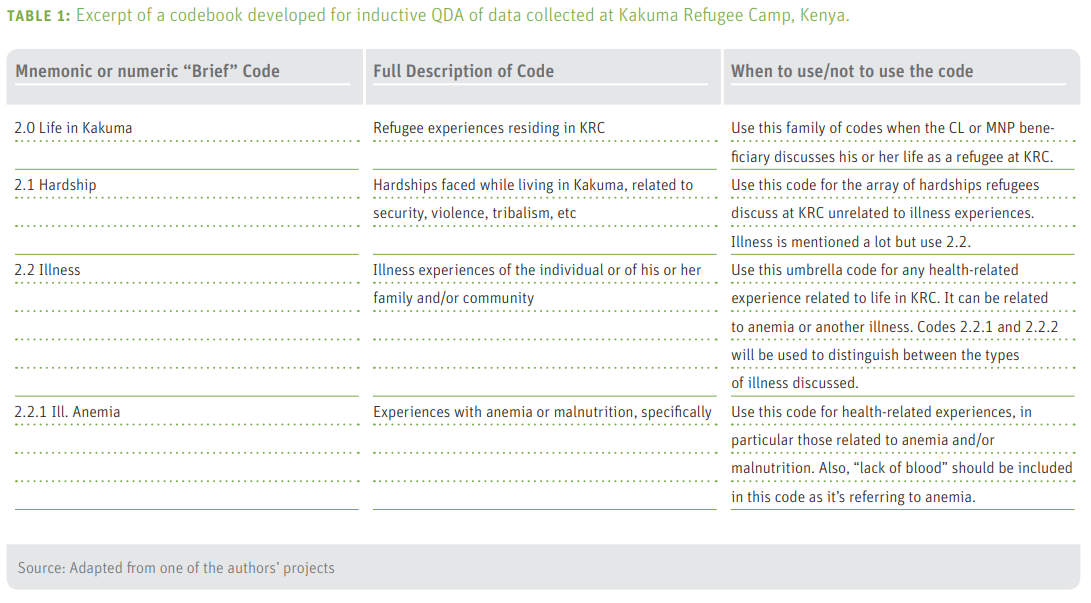
Such a codebook helps make clear to all involved in analyzing data the precise definition of each code, and this helps with increase inter-coder agreement when multiple people are involved in the coding process.

Codebook development can be time intensive because it often involves multiple revisions as codes are often changes throughout the coding process.

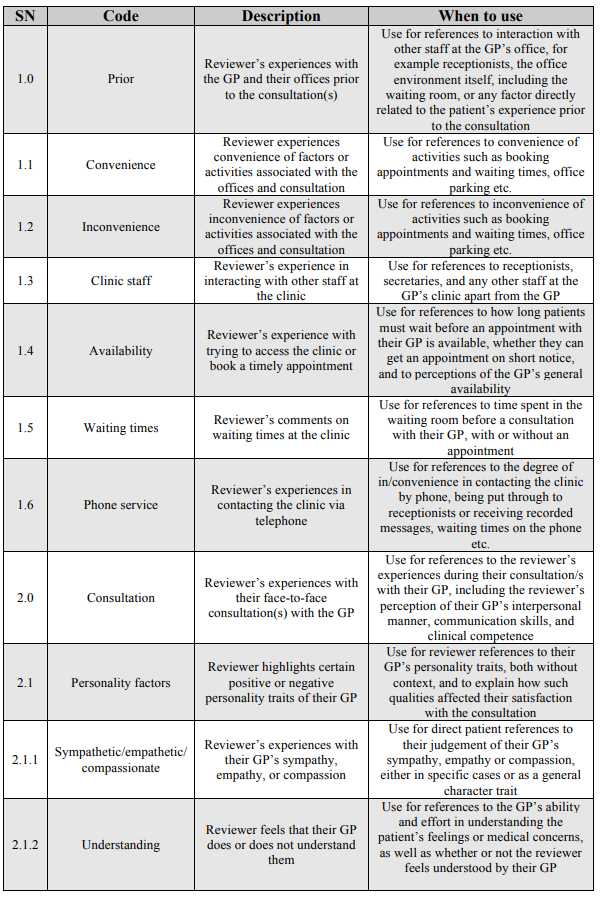
Below is an example from DeCuir-Gunby, Marshall, and McCulloch (2011).



Kodish and Gittelsohn (2011) present the following example of a codebook.



Below is a section of the codebook Kilby (2014) developed for a review of general practitioners in Norway.



**References**

DeCuir-Gunby, J.T., P.L. Marshall, A.W. McCulloch. (2011). Developing and using a codebook

for the analysis of interview data: an example from a professional development

research project, Field Methods 23, pp. 136–155.

Impedovo, M.A., Ritella, G., & Ligorio, M.B. (2013). Developing codebooks as a new tool to analyze students’ ePortfolios. International Journal of ePortfolio, 3, 161-176.

Kilby, G. (2014). A qualitative content analysis of onlne GP reviews in Norway. Master’s Thesis, Department of Health Management and Health Economics, University of Oslo.

LeCompte, M.D. (2000). Analyzing qualitative data. Theory Into Practice, 39(3), 146–154.

Moore, M. T., and B. W. Griffin. 2006. Identification of Factors that Influence Authorship Name Placement and Decisions to Collaborate in Peer-Reviewed, Education-Related Publications. Studies in Educational Evaluation 32 (2): 125–35.

Stephen Kodish & Joel Gittelsohn (2011). Systematic Data Analysis in Qualitative Health Research: Building Credible and Clear Findings. Sight and Life.