Geography 500 Seminar Paper - Writing Through Materials, Coding, and Authoethnography

Edmund B. Molder

PhD Student

Nelson Institute for Environmental Studies

University of Wisconsin, Madison

December 15th, 2021

Written for Professor Ian Baird's Geography 500 course, fall semester, 2021

Introduction and seminar review

The topic of the seminar I led with my classmate Matthew Bettencourt on December 8th, 2021 was "Writing Through Materials" and focused on analysis of qualitative data, and transforming data into research results. Readings covered different aspects of qualitative analysis including data organization, qualitative data coding, use of software, and integrating writing into analysis. The seminar focused primarily on qualitative data coding and various approaches to coding employed by qualitative researchers. We unpacked the differences between types of code (descriptive vs. analytic), and their relationships with inductive or deductive coding methods. We also discussed the use of computer programs to automate qualitative coding or analyze large datasets, and the potential benefits and challenges these pose. During the latter half of the seminar, we split the class into three groups and assigned each a different method to code the same set of qualitative data. The following discussion revealed that the distinction between inductive-deductive methods and descriptive-analytic codes often breaks down in practice, and the challenges that arise when a rigid coding structure meets a messy set of data.

Research using qualitative coding often requires a combination of approaches to best fit the data and research question being addressed. Several recurring themes from the course were brought into discussion such as establishing rigor in qualitative research, researcher reflexivity and positionality, and the contexts in which different methodologies are appropriate or useful. The nonlinear nature of qualitative research was also raised in the discussion of recursive coding cycles and the incorporation of writing into analysis. The latter was given more attention in the assigned readings (particularly Crang & Cook ch. 9) than in our class discussion, however future

seminar leaders would benefit from greater focus here given the many linkages between coding, analytical writing the research process more broadly.

In the remainder of this paper, I review some of the main topics and concepts that structured our discussion of coding qualitative data. These include the purposes of coding, types of codes and coding approaches, the use of software for qualitative data coding, and various ways of writing through qualitative data in support of deeper analytical insight and researcher reflexivity.

Purposes of coding

Coding is the process of identifying and organizing themes in qualitative data (Cope, 2005) and thus serves three primary purposes: data reduction, data organization, and data exploration, analysis, and theory-building. Qualitative data can come in many formats such as interview transcripts, photographs, field notes, audio or video recordings, oral histories, historical documents, reports, etc. The process of coding helps researchers draw meaningful connections or patterns from these data and allows larger volumes of data to be incorporated into an analysis. An organized index of data also allows data to be queried in order to answer specific questions, serving a similar purpose to finding aids used in archival research (Cope, 2005). A basic index of data is created by labeling or assigning attributes to individual pieces of data that group them into categories (LeCompte & Schensul, 2012). Interview transcripts, for example, might be assigned attributes of "data type" (interview), "date" (date interview was conducted), "topic" (topic of interview), and "subject name" (name of the subject being interviewed). Other data types might require more attribute codes to allow more effective comparison or pattern drawing between

different cases. However, the researcher should also consider their research question and ultimate goals for the research to avoid making their analysis more complex than needed.

In addition to organizing, coding qualitative data also helps reduce data and explore higher level meanings or theory. Through the process of coding, a researcher must read and re-read their data to make judgments about how it should be classified in an emerging coding structure (Linneberg & Korsgaard, 2019) and thus construct meaning throughout the coding process. Researchers must decide how a piece of data can or should be broken apart (ie by sentence, paragraph, phrase, or down to the individual word), whether different pieces can or should be assigned to more than one code, and whether different codes ultimately refer to distinct ideas or they should be collapsed into a single category. Close reading and making judgments throughout this process can trigger analytical connections that might be missed otherwise (Linneberg & Korsgaard, 2019), and provide useful perspective for further data collection and writing. In a sense, researchers also interact with participants through reading and coding material, and can come to understand participants' views and actions from their own perspectives. Further, data reduction via coding makes it easier to access illustrative quotes that may be used in written work or while thinking about how different codes are thematically related. Regular reference back to the original data is important for ensuring a rigorous coding process and that the researcher is acknowledging alternative interpretations of data that might arise through coding.

Types of codes and coding approaches

There are several approaches to coding written or textual data, and several types of codes more commonly associated with each. Broadly speaking, however, qualitative coding may be

simplified into deductive and inductive approaches. In a deductive (or 'etic' (Crang & Cook, 2007) coding approach, the researcher decides what to code for before coding the data. This approach is useful for focusing on issues known to be important in the literature and for testing or refining theory (Linneberg & Korsgaard, 2019). Deductive coding is also useful if the researcher's goal is to generalize across cases, as they can apply the same coding frame and look for similarities and differences.

Content analysis is a common deductive coding technique in which particular words or phrases are coded, either by hand or with the aid of software (Krippendorff, 2012). Where the coding rules are strictly defined, the approach is an essentially quantitative technique and the results are often analyzed to find statistical patterns or correlations (Cope, 2005). Content analysis may also be done in a more qualitative manner, with 'conventional' qualitative content analysis generating coding categories from the text itself and 'directed' content analysis drawing from relevant theory or research findings to guide initial codes (Hsieh & Shannon, 2005). In most cases, a codebook is useful for tracking newly emergent codes, examples of their use, ambiguous cases, and to establish 'intercoder reliability' (Hayes & Krippendorff, 2007) when multiple coders are working on the same dataset (DeCuir-Gunby et al., 2011; Reyes et al., 2021).

As deductive coding is primarily used to categorize words or phrases (as opposed to their implicit meanings), codes generated through this approach are commonly 'descriptive' in that they reflect themes or patterns that are obvious on the surface or stated directly by research subjects (Cope, 2005). Descriptive codes are like data labels and correspond with Cope's (2005) notion of 'manifest' messages (and resulting 'manifest codes') in that the message is blatant and obvious. 'In vivo' codes are descriptive codes that are taken directly from the text or are

common phrases found in the data. As the coding process commonly occurs recursively with codes generated and refined in each pass through the data, descriptive codes are often those created first as a means to produce an initial organizational structure to the data. Subsequent analysis and close-reading of descriptively coded text allows researchers to combine or split codes to arrive at deeper thematic ideas. Crang and Cook (2007) emphasize the ("maddeningly" (Agar, 1986: 29)) recursive nature of qualitative coding and describe coding as a process occurring in cycles. With each cycle, new codes are created and refined, and the researcher grows more familiar with the nuanced meanings associated with each piece of data. This process of developing codes over the course of analyzing a dataset is called 'inductive' (or 'emic' (Crang & Cook, 2007) coding, and codes that capture deeper thematic information within the data are called 'analytic' or 'thematic' codes. Analytic codes correspond with 'latent messages' described by Cope (2005), as they represent messages that are more implicit in a research subject's speech or the context in which the interview (or field notes, etc.) occurred.

While the pairs of coding approaches (deductive and inductive), code types (descriptive and analytic), and message types (manifest and latent) might appear to represent a binary of coding characteristics, their uses are not mutually exclusive. That is to say, an analytic code (e.g. 'fighting for change') may be used in a deductive coding approach where it corresponds to a theme embedded in the research question ('how do Indigenous communities resist continued fossil fuel expansion?'). A researcher may similarly identify emergent descriptive codes (e.g. 'protest' or 'partnerships') that only become evident through an inductive coding process. In practice, a "blended" approach or "abduction" coding is common (Linneberg & Korsgaard, 2019) as inductive coding can more easily allow the voice of participants to emerge and the

development of empirically-grounded theory, while a deductive approach ensures the analysis maintains some structure. Novel themes or connections may nonetheless arise later as the researcher becomes more familiar with their data and the codes they created during early coding cycles (Fereday & Muir-Cochrane, 2006). Anselm Strauss (1987; cited in Crang and Cook, 2005) even cautioned against early definition of themes to avoid biased interpretation of later text.

Grounded theory is one methodology commonly associated with inductive coding that generally entails a process of data collection, coding (and constant comparison between new and already coded data), questioning the emerging codes, theoretical sampling, and memo-writing (Knigge, 2016; Walker & Myrick, 2006). While time-intensive, grounded theory allows qualitative researchers to develop deeper theoretical insights grounded in findings from the communities or individuals of focus in their research. Further, constant comparison among and between different codes allows verification of themes as they develop, supporting more rigorous qualitative research (Kolb, 2012). Grounded theory has been criticized as a less rigorous method of qualitative analysis (Suddaby, 2006), and scholars have proposed several ways for researchers to enhance the rigour of analyses employing grounded theory methodology (Cooney, 2011). These include validating emerging themes with their research subjects' experiences or intended meanings, sharing emergent theory with other experts to see if it 'fit' their experiences, and keeping detailed analytical memos describing their decision-making process through the duration of coding. Barney Glaser (1978; cited in Hallberg, 2006), one of the founders of grounded theory, argued that its power lies in grounding theory development completely and exclusively in data gathered from the field: "The researcher only has to expose him or herself to the research

area, and then the data, as well as the research question, will reveal itself to him or her," (Hallberg, 2006: 145; however, Suddaby (2006) contests the notion that such an approach counts as grounded theory). Critics of this approach have argued that some prior theoretical background and research question development is necessary and unavoidable (Hallberg, 2006), at the very least to expedite a field research process that might take many months without an initially defined direction.

Grounded theory notwithstanding, qualitative coding is most useful when applied to a particular end. Researchers should have a clear research question defined before beginning the coding process (and before collecting data) in order to recognize relevant descriptive or thematic codes in the data as they arise. Some scholars recommend writing the research question on each page of one's analytic notes to refer to while coding (Saldana, 2021), and regularly reflecting on how coded data might support the overall research goal. Coding is time consuming (Ngulube, 2015) and the more interview transcripts or pages of field notes one deems relevant for analysis, the more time is required to analyze it all. Saldana (2021) highlights the disagreements among qualitative methodologists as to whether or not every piece of field data is worthy of analytical consideration. Glaser (1987) argued "everything is data" and thus relevant for consideration in grounded theory research, while others contend as much two-thirds of field data may be irrelevant for answering a given research question. The amount of data one decides to include for in-depth coding will largely depend on the specifics of their coding approach, particularly the tools and/or software one employs to organize coded data and emerging thematic material.

Tools and software for coding qualitative data

Coding qualitative data can be done effectively using printed copies of qualitative data (e.g. interview transcripts) and a multi-colored set of pens or highlighters. Crang & Cook (2007) outline some important steps to prepare for coding qualitative data including making copies of all data to be coded, reading through all the data to re-familiarize oneself with the context and initial analytical ideas that arose during data collection, and starting an analytical diary or memo collection for keeping track of ideas and decisions throughout the analysis. Thereafter, coding is simply a matter of reading and annotating the data. Qualitative coding using pen and paper can be effective because the materials limit the speed at which data may be coded, and the total volume of data that may be incorporated into an analysis, leading to more time spent with the data and deeper potential insight. On the other hand, reliance on paper copies of the data can be tedious and risky if coded sheets are lost or damaged, while pens and highlighters limit the number of codes or themes a researcher can assign.

A variety of computer-assisted qualitative data analysis software (CAQDAS) packages have been developed to assist qualitative researchers in managing their data and the coding process. These can shorten analysis timeframes and provide more thorough and rigorous coding and interpretation (Basit, 2003; Jones, 2007), but are not a panacea for qualitative analysis. Most CAQDAS requires at least a modest learning curve, for example. The ability to quickly and easily create new codes can also be counterproductive and lead to a coding process that is more mechanical, less in-depth, and less reflective about meanings and connections with other codes or themes (Linneberg & Korsgaard, 2019). Saldaña (2021) recommends students (and researchers in general) new to qualitative coding attempt manual coding before using software in

order to learn the fundamentals of different coding approaches. Using Microsoft Word or other word processing software can be a useful coding tool as it allows more flexibility than paper but does not require learning entirely novel software (Hahn, 2008). For large datasets and largely inductive coding approaches, however, CAQDAS can be very useful for developing, managing, and visualizing coding hierarchies, relationships, and emergent theoretical insights grounded in the data (Welsh, 2002).

It is worth noting that coding with pen and paper and with CAQDAS are both manual approaches to qualitative analysis. That is, "CAQDAS itself does not actually code the data for you; that task is still the responsibility of the researcher" (Saldaña, 2021: 28). The researcher is also responsible for drawing meaning from coded data and themes to produce rich, empirically grounded research. There is growing interest in the use of software or programming methods to automatically code qualitative data and other "natural language" texts (Crowston et al., 2010; Crowston et al., 2012; Lennon et al., 2021). Dempster and Woods (2011), for instance, describe the use of Transana software to automatically code paired audio and video data about the 2008 financial crisis. Further, "topic modeling" is a machine learning method increasingly used in communications research in which large volumes of text data are algorithmically processed to identify themes within the data according to word association and frequency (Baumer, 2017; Blei, 2012). While useful for processing "big data" (e.g. text data automatically scraped from the web), the themes are not necessarily meaningful or useful for addressing fundamental research questions until they have been reviewed by a researcher. Such automated methods thus must be used in combination with the same close reading, judgement, and interpretation required of manual coding approaches.

CAQDAS are particularly useful for combining multiple data formats into a single analysis. Dempters and Woods (2011), for example, emphasize Transana's power to analyze large volumes of audio, video, and text data simultaneously. While some formats may be less commonly (or usefully) coded than others, ethnographic researchers commonly produce or incorporate field notes (Tessier, 2012), interview transcripts (Cope, 2010), photographs (Chapman et al., 2017), drawings (Bland, 2012; Moskal, 2017; Taussig, 2009), or historical documents (Bowen, 2009) into their work, and can benefit from synthesizing the insights provided by such disparate data sources. CAQDAS offers one powerful and efficient tool for coding such a variety of data formats.

Writing through materials

While the readings covered a range of topics, the in-class discussion and activity focused primarily on qualitative coding. Writing autoethnography and montage were both discussed in Crang and Cook (2007) in addition to writing through coded materials, and I would have liked to spend more time discussing how writing may be incorporated into qualitative research as an aid to analysis. Given their broad scope, an entire class period each could have been devoted to writing and coding/analysis, though the interconnected nature of these themes makes them a good fit together, both as writing functions while doing research and assembling a finished piece of writing. Saldana (2021), Linneberg & Korsgaard (2019), and Crang & Cook (2007) emphasize that keeping an analytical journal while coding is necessary for tracking thematic ideas as they develop, and the process of writing itself can nudge the researcher to think through coded materials more deeply and make connections they might otherwise miss (Liamputtong, 2009).

Autoethnography is a particular qualitative method that merges common tools of ethnography such as participant observation with reflexive autobiographical writing (Jones, 2016). Autoethnography is highly reflexive in that it "seeks to describe and systematically analyze personal experience in order to understand cultural experience" (Ellis et al., 2011), and as such, is both process and product of research. Autoethnographic writing is distinct from analytical memo writing one might do while coding qualitative research as autoethnography is more often written while the researcher is "in the field" collecting data rather than analyzing data to generate meaning. But both analytical memo writing and autoethnography serve to turn the focus back on the researcher and their thoughts, emotions, ideas, and experiences during research. As a research method, autoethnography in particular can serve to dissolve the researcher-subject boundary by placing the researcher reflexively within the research context and forcing reimagination of how researchers represent themselves and the objects and/or people of study (Butz & Besio, 2009). As coding qualitative data is a significant component of qualitative research--and where much of the meaning and theoretical concepts resulting from research are developed--writing analytical memos can and should be autoethnographic to the extent that they encourage deep self-reflection by the researcher regarding their experience and decisions made while coding. Like most other methodological distinctions in qualitative research, the boundaries between collecting, analyzing, and "writing up" data and research results are never clear (Tracy, 2013). Rather, researchers find meaning in qualitative data by "writing the meaning into being".

Conclusion

Over the course of this semester, we examined a number of qualitative methods used to gather data for ethnographic and human geography research. Data gathered through these

methods, however, are rarely clean, organized, and clearly meaningful to others unfamiliar with the research context. Analyzing data through qualitative coding is a key component of qualitative research as it provides the researcher a means to examine their data closely and develop theoretically rich narratives about its meaning and its broader significance. In this paper, I reviewed deductive and inductive coding approaches, descriptive and analytic code types, and a variety of more in-depth coding methodologies that variously draw from inductive, deductive, and computational approaches. I also discussed the tradeoffs offered by CAQDAS in coding qualitative data. Namely, CAQDAS is an efficient tool for organizing and managing, coding hierarchies and qualitative data of varying formats, but can make it more difficult to achieve the deep familiarity with qualitative data as afforded (required) through slower pen-and-paper coding. Researchers new to coding qualitative data should learn the fundamentals using paper and pen or basic word processing software before diving into CAQDAS. Analytical memo-writing throughout the coding process is widely recognized as critical to effective qualitative research (Saldaña, 2021) because it forces reflection about emerging themes and theory development and helps solidify potential connections between disparate pieces of data. Autoethnographic writing, both in contrast to analytical memo-writing and as a reflexive form of writing about the coding process, is another process-focused method of writing through materials that can provide deeper, often critical insight on a researcher's positionality throughout the conduct of qualitative research. While my experience conducting empirical qualitative research is limited, I began this course with an M.A. in anthropology and some familiarity with qualitative coding approaches, CAQDAS, and ethnographic research methods. At the course's conclusion, however, I feel simultaneously less knowledgeable about qualitative methodologies and more

equipped to do empirical qualitative research. I know far more about archival research methods, focus groups, and filmic approaches than at the semester's start, yet each week's focus on reflexivity, positionality, and ethics of human subjects research forces me to recognize that the knowledge gained reveals even greater knowledge gaps. Writing this paper has been a challenging exercise in acknowledging these knowledge gaps, at least those related to qualitative data coding, and I feel much more familiar with the subject as a result. I look forward to broadening my familiarity with other methods of qualitative inquiry and applying these methods for real qualitative research in the field.

References

Agar, M. H. (1986). Speaking of ethnography (Vol. 2). Sage.

Basit, T. (2003). Manual or electronic? The role of coding in qualitative data analysis. Educational research, 45(2), 143-154.

Baumer, E. P., Mimno, D., Guha, S., Quan, E., & Gay, G. K. (2017). Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence?. Journal of the Association for Information Science and Technology, 68(6), 1397-1410.

Belgrave, L. L., & Seide, K. (2019). Coding for Grounded Theory. In A. Bryant & K. Charmaz, The SAGE Handbook of Current Developments in Grounded Theory (pp. 167–185). SAGE Publications Ltd.

Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative research journal.

Butz, D., & Besio, K. (2009). Autoethnography. Geography Compass, 3(5), 1660-1674.

Chapman, M. V., Wu, S., & Zhu, M. (2017). What is a picture worth? A primer for coding and interpreting photographic data. Qualitative Social Work, 16(6), 810-824.

Cooney, A. (2011). Rigour and grounded theory. Nurse researcher, 18(4).

Cope, M. (2005). Qualitative methodologies for human geographers. Ed. Iain Hay.

Cope, M. (2010). Coding transcripts and diaries 27. Key methods in geography, 440.

Crowston, K., Liu, X., & Allen, E. E. (2010). Machine learning and rule‑based automated coding of qualitative data. proceedings of the American Society for Information Science and Technology, 47(1), 1-2.

Crowston, K., Allen, E. E., & Heckman, R. (2012). Using natural language processing technology for qualitative data analysis. International Journal of Social Research Methodology, 15(6), 523-543.

DeCuir-Gunby, J. T., Marshall, P. L., & McCulloch, A. W. (2011). Developing and using a codebook for the analysis of interview data: An example from a professional development research project. Field methods, 23(2), 136-155.

Dempster, P. G., & Woods, D. K. (2011, January). The economic crisis though the eyes of Transana. In Forum Qualitative Sozialforschung/Forum: Qualitative Social Research (Vol. 12, No. 1).

Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: an overview. Historical social research/Historische sozialforschung, 273-290.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International journal of qualitative methods, 5(1), 80-92.

Hahn, C. (2008). Doing qualitative research using your computer: A practical guide. Sage.

Hallberg, L. R. (2006). The "core category" of grounded theory: Making constant comparisons. International journal of qualitative studies on health and well-being, 1(3), 141-148.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. Communication methods and measures, 1(1), 77-89.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, *15*(9), 1277-1288.

Jones, M. L. (2007). Using software to analyse qualitative data.

Jones, S. H. (2016). Handbook of autoethnography. Routledge.

Kolb, S. M. (2012). Grounded theory and the constant comparative method: Valid research strategies for educators. Journal of emerging trends in educational research and policy studies, 3(1), 83-86.

Krippendorff, K. (2018). Content analysis: An introduction to its methodology. Sage publications.

Knigge, L. G. (2016). Grounded theory. International Encyclopedia of Geography: People, the Earth, Environment and Technology: People, the Earth, Environment and Technology, 1-5.

LeCompte, M. D., & Schensul, J. J. (2012). Analysis and interpretation of ethnographic data: A mixed methods approach (Vol. 5). Rowman Altamira.

Lennon, R. P., Fraleigh, R., Van Scoy, L. J., Keshaviah, A., Hu, X. C., Snyder, B. L., ... & Griffin, C. (2021). Developing and testing an automated qualitative assistant (AQUA) to support qualitative analysis. Family Medicine and Community Health, 9(Suppl 1), e001287.

Liamputtong, P. (2009). Qualitative data analysis: conceptual and practical considerations. Health promotion journal of Australia, 20(2), 133-139.

Linneberg, M. S., & Korsgaard, S. (2019). Coding qualitative data: A synthesis guiding the novice. Qualitative research journal 19(3): 259-270. https://doi.org/10.1108/QRJ-12-2018-0012

Ngulube, P. 2015. Qualitative data analysis and interpretation: systematic search for meaning, in Mathipa, ER & Gumbo, MT. (eds). *Addressing research challenges: making headway for developing researchers*. Mosala-MASEDI Publishers & Booksellers cc: Noordywk, pp. 131-156

Reyes, V., Bogumil, E., & Welch, L. E. (2021). The Living Codebook: Documenting the Process of Qualitative Data Analysis. Sociological Methods & Research, 0049124120986185.

Strauss, A. L. (1987). Qualitative analysis for social scientists. Cambridge university press.

Suddaby, R. (2006). From the editors: What grounded theory is not. Academy of management journal, 49(4), 633-642.

Taussig, M. (2009). What do drawings want?. Culture, Theory & Critique, 50(2-3), 263-274.

Tessier, S. (2012). From field notes, to transcripts, to tape recordings: Evolution or combination?. International journal of qualitative methods, 11(4), 446-460.

Tracy, S. J. (2019). Qualitative research methods: Collecting evidence, crafting analysis, communicating impact. John Wiley & Sons.

Welsh, E. (2002, May). Dealing with data: Using NVivo in the qualitative data analysis process. In Forum qualitative sozialforschung/Forum: qualitative social research (Vol. 3, No. 2).