Qualitative Data Analysis: Novelty in Deductive and Inductive Coding

Reuben Bihu

1University of Dar es Salaam

September 22, 2023

Abstract

Qualitative data analysis is a critical phase in qualitative research. One of its cornerstones is the coding process. Deductive and inductive codes are generated for breadth and depth exploration of the research topics, respectively. Deductive coding bases on ‘a priori’ of codes to which segments of texts and transcripts are assigned. Inductive coding begins with data, making segments and categories, and generating relevant codes. It is concluded from this analysis that both the deductive and inductive coding often occurs in a qualitative data analysis process. Areas of further investigation are recommended.

Qualitative Data Analysis: Novelty in Deductive and Inductive Coding

Reuben Bihu

reubimon@gmail.com

Abstract

Qualitative data analysis is a critical phase in qualitative research. One of its cornerstones is the coding process. Deductive and inductive codes are generated for breadth and depth exploration of the research topics, respectively. Deductive coding bases on ‘a priori’ of codes to which segments of texts and transcripts are assigned. Inductive coding begins with data, making segments and categories, and generating relevant codes. It is concluded from this analysis that both the deductive and inductive coding often occurs in a qualitative data analysis process. Areas of further investigation are recommended.

Key words

Qualitative data analysis, Deductive coding, Inductive coding

Introduction

Qualitative research is a data generation process which uses primary data from, or secondary textual data generated from participants. The major characteristic feature under the approach is the process deploying and manipulating data from human voice. The voice may be raw oral, written or deconstructed imagery. Qualitative data therefore may be identified as narratives in transcripts and notices which form specific patterns in texts. Traditionally, the analytical works categorise thematic patterns as segments with common meanings, likeness, or resemblance. Thematic patterns can be coded manually or by software so can be quantified into descriptive statistics11Simple estimations of measures of central tendency, e.g., mean, mode and median. Mixed-methods under software like MAXQDA can also assist generation of statistical coefficients employed in generalisable studies. The demand for qualitative coding has risen high, with many options for online and offline qualitative data gathering options available. Researchers are faced by the questions of how ontology of a phenomenon and cases can be reached under minimised logistic and technical barriers. Options for epistemological takeaways have demanded an understanding of various ways data can be gathered, how
they can be processed, coded and analysed without losing the implicit attribute of data quality in technique used. Based on the experience of using MAXQDA and NVivo software in a PhD programme, and interactive experience with early career researchers in various scholarly platforms, a need was found to put the coding experience into writing for scholars. This article therefore is designed to contribute to knowledge on data coding in qualitative data analysis processes.

Qualitative coding is preceded by accumulation of robust data packs on a phenomenon researched. The key data gathering technique in qualitative research is interviewing the participants. However, qualitative data may be gathered using different techniques as shown in table 1.

Table 1 Different techniques to gather qualitative data for analysis.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Procedure</th>
<th>Output</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview and FGD</td>
<td>Audio recording and transcription</td>
<td>Transcript</td>
<td>Bihu (2020) van Eeuwijk &amp; Angehrn (2017)</td>
</tr>
<tr>
<td></td>
<td>Field notice taking and transcription</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video</td>
<td>Discourse analysis and transcription</td>
<td>Transcript</td>
<td>Kamalu &amp; Osisanwo (2015)</td>
</tr>
<tr>
<td>Picture</td>
<td>Deconstruction and transcription</td>
<td>Transcript</td>
<td>Glaw et al. (2017)</td>
</tr>
<tr>
<td>Webpage</td>
<td>Online statements</td>
<td>Text</td>
<td>Andrew et al. (2003)</td>
</tr>
<tr>
<td>Bibliographical data</td>
<td>Bibliometric analysis</td>
<td>Text</td>
<td>Donthu et al. (2021)</td>
</tr>
</tbody>
</table>

Data coding may be initiated on data gathered through one or several techniques (cf. table 1). Implicitly, the transcripts for coding may be generated by transcribing interviews, pictures, video, and audio data. Coding may also utilize documents, web pages, and bibliographical data. Codes aid to generate summed data patterns on the themes and sub-themes attributed the researched problem (i.e., research questions). The use of software such as MAXQDA and NVivo can facilitate the coding process. For effective coding serving the real description of the topic, i.e., meeting the real hallmark of qualitative research, both deductive and inductive coding appears useful.

Methodology

This review used selected documents published on the qualitative data analysis techniques, particularly, on ‘deductive coding’ and ‘inductive coding’ in qualitative data analysis. I adopted a methodology used in Bihu (2021). The Boolean logic was used to generate documents on ‘deductive qualitative coding’ and ‘inductive qualitative coding’. Document management was assisted by Zotero software. The review sought to answer the following questions:

How deductive coding is implemented in qualitative data analysis?
How inductive coding is implemented in qualitative data analysis?

Results

Deductive and Inductive Qualitative Data Coding

As researchers navigate through narratives and texts attributed to research topics (with generation of themes and sub-themes), conceptually both deductive and inductive coding is a stage in a research process. Deductive and inductive research reflects paradigms, i.e. positivism and post-positivism (Barney & Saleem, 2008) take place especially in evaluation research projects (Thomas, 2006). The general relationship appears as in figure 1.
Figure 1 Deductive and inductive coding in qualitative research. NB: Both deductive and inductive coding occur in a single data analysis phase, associated with breadth and high-depth codes, respectively. Arrows reflect on the richness of data (i.e., codes), in contrast to conceptual implications of the processes.

Research problems in qualitative data analyses give arrays of related questions which may be studied using one or several techniques (cf. Table 1); through scrutinizing qualitative texts and transcripts for similarities and relevance. Grouping of segments in texts and transcripts is done by attributing specific groups to specific codes to obtain its measures of frequency and other descriptive statistics. While deductive coding process may show thin attribution of codes to segments in texts and transcripts, inductive coding leads to thick descriptions on the problem (cf. figure 1). Inductive coding involves generation of more sub-themes emerging from the parent themes and sub-themes to the deeper of the case.

Deductive Qualitative Coding (DQC)

Deductive analysis involves generation of analytical themes from the main topic or research problem, research questions, and emerging key assumptions of what data may entail. Principally, the researcher begins coding with prior knowledge and conceptual analysis of the research problem. This is one of the key elements of an evaluation research project (Thomas, 2006). It is a top-down process moving from the general topic to specifics, with predetermined codes, and finding excerpts that fit those codes.

Fereday and Muir-Cochrane (2006) indicate that a researcher should first develop ‘a priori’, i.e., a template of codes, based on research questions and theoretical concepts. The analyst starts with a code and finds excerpt(s) that fit(s) the code. Then, they add that a researcher should also check the applicability of the code to the raw data, summarise data and identify initial themes, apply template of codes and addition coding, connect the codes and identify themes; and corroborate and legitimise coded themes (to identify sub-themes through familiarity). For example, deductive coding using software like MAXQDA allows a researcher to execute coding processes with prior thought codes, and then attributing the codes to the relevant segment in the text window. This approach ensures that the coding processes identify themes that explore the research questions in breadth covering all relevant key concepts.

Inductive Qualitative Coding (IQC)

Inductive analysis refers to approaches that primarily use detailed readings of raw data to derive concepts, themes, or a model through interpretations made from the raw data by a researcher (Thomas, 2006). This is effective in exploratory research. Fereday and Muir-Cochrane (2006) encourage a detailed study of themes, and further study of their similarities and differences. Consequently, a theory may emerge from the data patterns. Inductive coding serves several purposes identified by Thomas (2006):

1. To anchor condensed brief summary format of extensive and varied raw text data onto a theme, keyword or phrase for easy identification of categorical segments in texts;
2. To ease linking categorical segments of texts from raw data to research objectives; and
3. To theorise on the underlying data patterns in the text data.

Coding begins with close readings of text and consideration of the multiple meanings that are inherent in the text. The evaluator then identifies text segments that contain meaningful units and creates a label (i.e. a code) for a new category to which the text segment is assigned. Additional text segments are added to the categories to which they are relevant (Thomas, 2006). When software is used, e.g., MAXQDA, it eases the analytical processes (i.e. substitutes ordinary excel spreadsheet and manual ways) by allowing retrieval and exportation of the qualitative text matrices and quantified segments for presentations. This may serve for judgments in analytical processes based on the data frequency and percentiles attributable to given patterns.

Discussion

In practice, qualitative data analysis most often utilises both deductive and inductive coding. When software is used, e.g., MAXQDA, the deductive and inductive coding brings several codes and quantified text segments on a software interface. The common technical aspects of both deductive and inductive analyses in-
clude checking the logical consistency, subjective interpretation, and adequacy of key messages in transcripts and text segments (Fereday & Muir-Cochrane, 2006). The implicit consideration of logical consistency is the technical aspect of ensuring the highest degree of clarity between analytical themes and coding options. The subjective interpretation requires that ‘a priori’ must base on subjective meanings to the researcher. Nonetheless, the postulate on logical consistency is that researcher’s constructs must be consistent with those found in common-sense experience. This analysis is confined to the view of subjective meaning of individual researchers. However, the core attribute of logical consistency is the ability to generate codes which can be attributed to other researchers’ experiences.

Different researchers have suggested different stages which the analysts have to undergo in qualitative analysis (Braun & Clarke, 2006; Erlingsson & Brysiewicz, 2017; Thomas, 2006). The scholarly convergences in terms of coding are identified in five logical stages. First, the processes require preparation of raw data files. This involves data cleaning. If the software is used, documents are assigned with proper names.

Second, there should be close reading of text to discern similarities and differences in text themes. Close reading of texts allows familiarisation of the researcher with data and the hermeneutic spiral. The researcher looks into the meaning unit by understanding higher and lower levels of abstractions close to the text and the content. Keeping the aim in focus, the researcher writes down the initial impressions, and embraces intuition pertaining to the key messages of the text and reactive cognitive aspects.

Third, the researcher segments the text into different themes. Segmentation of the text into themes requires dividing up the text into different meaning units. The researcher navigates through several transcripts and texts many times to identify conceptual schema in specific segments, and divides up segments into specific themes. Themes should express an underlying meaning; the latent content (i.e., interpretative level) found in two or more segments.

Fourth, the researcher assigns the identified segment to its categorical code. The text is therefore reduced into segments while preserving the core meaning. Segments are grouped into different categories of themes based on their coherence in meanings. A category is assigned to a label or name, i.e., a code (or a label to a category in inductive coding) to suit the description of the segment meaning unit in about one or two words long, e.g., work experience, work routine, reading distracters, etc. Codes are short expressions that are factual sounding. They are therefore organized into a category when they are describing different aspects, similarities or differences of segments in texts.

Last, the researcher does a continuous revision and refinement of the category system. This may lead to overarching themes (i.e., higher level abstraction) which reflects more on the interpreted latent meaning of the text segment. Codes are grouped into categories that answer the research questions. The researcher compares and appraises codes to determine which codes belong together, thereby forming a category. Therefore, codes dealing with the same issue are categorised together.

Conclusions

Deductive and inductive coding may co-occur in qualitative analysis in practice. Deductive coding process rigor is dependent on researchers’ capacity to manage the topic knowledge in broader known perspectives. It is significantly covering width of the skimmed and conceptualized themes (superficially), i.e., may be thin in attribution to description of the problem. Deductive codes are often based on theory, existing literature, or research questions. Inductive coding in contrast, starts with data, goes through grouping text segments into themes, and developing codes from themes.

Recommendations

I have found that a clear separation between deductive and inductive coding in a qualitative data analysis seem lacking. Further researchers may focus on establishing clear margins between inductive and deductive coding processes. This is particularly important for understanding the extent to which the data-set itself may guide for building up the wider theory.
Nonetheless, the level of involvement between the co-occurring deductions and inductions in coding seem to need more space for debate and research. What topics should weigh more of deductions and inductions in this regard, presents a dilemma. Scholarly debates and research are therefore needed, particularly basing on the methodologies of data analysis that may be deployed.

Reference


