Application of Rigour and Credibility in Qualitative Document Analysis: Lessons Learnt from a Case Study

Leanne M. Wood
Griffith University - Australia, leanne.wood@griffithuni.edu.au

Bernadette Sebar
Griffith University - Australia, b.sebar@griffith.edu.au

Nerina Vecchio
Griffith University, Gold Coast, Queensland, Australia, n.vecchio@griffith.edu.au

Follow this and additional works at: https://nsuworks.nova.edu/tqr

Part of the Community-Based Research Commons, Entrepreneurial and Small Business Operations Commons, Public Health Commons, Quantitative, Qualitative, Comparative, and Historical Methodologies Commons, and the Rural Sociology Commons

Recommended APA Citation

Application of Rigour and Credibility in Qualitative Document Analysis: Lessons Learnt from a Case Study

Abstract

This paper probes functions and processes of qualitative document analysis (QDA), a method widely used in case study research. It firstly demonstrates the application of a QDA framework to inform a case study of women entrepreneurs in rural Australia; and provides insights into the lessons learnt, including strengths and limitations of QDA. Secondly, the paper provides guidelines for novice researchers seeking to use thematic analysis in a QDA process, arguing for rigour in naming assumptions and explicitness about the procedures employed. The paper contributes to discussion in the literature that positions QDA not only as a convenient tool, but as a method embedded in a conceptual framework integral to the credibility and rigour of the qualitative "story" and what makes that story feel "right" to both researcher and reader (Corbin & Strauss, 2008).

Keywords

Qualitative Document Research, QDA, Documents, Content Analysis, Grounded Theory, Thematic Analysis, Social Constructionism, Case Study

Creative Commons License

Acknowledgements

Application of Rigour and Credibility in Qualitative Document Analysis: Lessons Learnt from a Case Study

Leanne M. Wood, Bernadette Sebar, and Nerina Vecchio
Griffith University, Gold Coast, Queensland, Australia

This paper probes functions and processes of qualitative document analysis (QDA), a method widely used in case study research. It firstly demonstrates the application of a QDA framework to inform a case study of women entrepreneurs in rural Australia; and provides insights into the lessons learnt, including strengths and limitations of QDA. Secondly, the paper provides guidelines for novice researchers seeking to use thematic analysis in a QDA process, arguing for rigour in naming assumptions and explicitness about the procedures employed. The paper contributes to discussion in the literature that positions QDA not only as a convenient tool, but as a method embedded in a conceptual framework integral to the credibility and rigour of the qualitative "story" and what makes that story feel "right" to both researcher and reader (Corbin & Strauss, 2008). Keywords: Qualitative Document Research, QDA, Documents, Content Analysis, Grounded Theory, Thematic Analysis, Social Constructionism, Case Study

Introduction

There are few aspects of our lives that are not recorded in some form, from official statements of birth, death and milestones in between; government and organisational directives, policy, legislation, maps and minutes; personal correspondence and photographs; marketing material; the ephemera of bus tickets and sales receipts; and the narratives generated by media, both printed and via electronic channels such as Facebook that bridge both the personal and public spheres. Such documents, and myriad other examples, are what Olson (2010) calls "the by-products of human activity" (p. 318) — written, visual, and physical material that reveals the complexities of our lives as individuals and societies. They provide evidence of the way in which we represent and organise ourselves; they reveal how we strive to make sense of our past and inform our future. Often, they play an active role in influencing those processes of sense-making.

This methodology paper probes functions and processes of qualitative document analysis (QDA), a method widely used alongside interviews and observation in case study research. It aims to contribute to the literature in two ways. Firstly, it presents examples of several QDA frameworks from across a diverse range of disciplines and approaches. These provide context for a case study of the application of a QDA framework and the consequent choice of analytic method used to inform research into women entrepreneurs and community wellbeing in rural Australia. Secondly, the paper provides guidelines for novice researchers seeking to use a particular kind of analysis (thematic analysis) in a QDA process, arguing for rigour in naming assumptions and explicitness about the procedures employed.

The role of documentary evidence is largely inconspicuous in most areas of qualitative research. Documents tend "to enter and to leave the 'field' in relative silence" (Prior, 2010, p. 417), often unobtrusively complementary to other methods in case study research (Olson, 2010). The documentary realities of social settings and the rich potential of documents as a

way of understanding social practices are often neglected in qualitative accounts of such environments (Coffey, 2014).

QDA provides a systematic methodological process for eliciting meaning from documentary evidence (Bowen, 2009). The interpretative process is both recursive and reflexive as the investigator moves between concept development, sampling, data collection, data analysis and interpretation. QDA is thus an emergent process focused on the search for underlying meanings, themes and patterns, rather than "a rigid set of procedures with tight parameters" (Altheide, Coyle, De Vriese, & Schneider, 2008, p. 127).

As such, QDA can be considered an "umbrella" descriptor for a systematic, reflexive process that may employ a range of document analysis methods. Scanning articles across the broad range of disciplines within which QDA is used — for example, medicine and health, business and management, communications and journalism, education, and sociological studies that cross disciplinary boundaries in any of these fields — provides examples of QDA utilising a range of methods, including variants of thematic analysis, content analysis and discourse analysis. These examples range from almost quantitative procedures (such as Hughes' (2004) coding and summary of duty statements and key selection criteria for public health nutritionists), through to ethnographic and grounded theory approaches to coding and analysis, such as Purtle's (2016) exploration of post-traumatic stress disorder in US federal legislation (Coffey, 2014).

Researchers' approaches to QDA will necessarily be influenced by both their worldview (ontology), and their understanding of how we, as human beings, know about that world (epistemology). To extrapolate from Bazeley (2013) — analysis is laid on this foundation:

> Our understanding of the nature of reality and truth and whether we discover, interpret, or construct realities and truths influence our choices of topic, methods, and conclusions. (p. 1)

In QDA terms, this translates to questions such as those raised by Gorichanaz and Latham (2016): Do documents exist? *How* do documents exist? And how are we to understand them? The specific method/s chosen to undertake qualitative document analysis necessarily emerge from the researcher's framework of understanding regarding what documents are, and how they interact with the non-documentary world. As Bazeley (2013) goes on to note:

> Thinking about [these foundations] sharpens and enriches our analysis, and our understanding of these things impacts on how we assess the trustworthiness of our conclusions. (p. 1)

Being explicit about the framework underpinning a QDA method contributes to the coherence and credibility of the study: it establishes that there is methodological congruence between the worldview of the researcher/s, the aims of the research, the nature and scope of the documents under investigation, and the analytic procedures undertaken. In Altheide's work on the dynamics of media and communication (e.g., 2000), for example, he moves beyond a broad brush description of having carried out a "thematic analysis" of relevant documents to explicitly describe a framework that justifies his procedures and techniques. In the accounts of many empirical studies, however, such detail is thin.

Qualitative Document Analysis Frameworks

Researchers across a range of disciplines have sought to explicate the integral connection between their approach to documentary evidence, their QDA framework and the chosen method for data analysis. This can be seen in the following exemplar frameworks, drawn from communication theory (Altheide, 2000), sports management (Skinner, Stewart, & Edwards, 2004), library and information sciences (Gorichanaz & Latham, 2016), nursing research (Miller & Alvarado, 2005) and medical sociology (Prior, 2008).

Altheide's QDA framework suggests that it is not media content in itself that influences social life, but the way in which such content is framed, and how it is contextualised and linked to courses of action (for example, the positioning of television violence in a narrative that links it to traits such as bravery, cunning, skill and sexual attractiveness [2000]). His approach is supported by a method Altheide calls "tracking discourse," which follows specific issues, words, themes, and frames across different media, over a period of time, and across different issues. Tracing emerging connections between key terms (for instance, *fear, crime* and *youth*) can suggest ways in which reality is socially constructed over that period (Altheide, 2006).

In an alternate approach, Skinner et al.'s (2004) study of two policy and planning documents developed by a major sporting organisation distinguishes between "manifest" (or "obvious") content, and "latent" or underlying and inferred meaning. The authors describe a framework that combines latent content analysis with Derrida's theories on deconstruction and "differance," enabling deep analysis of the text and suggesting paradoxes, biases, and contradictions influenced by the political and social context. The framework provides scaffolding for their detailed exploration of key words such as "tradition" and "viability," and the way in which differing interpretations at various levels of the organisation contributed to fragmented outcomes and challenged organisational change processes.

Gorichanaz and Latham's work, in contrast, outlines a framework of "document experience" that represents a transaction between a person and an object. The object of analysis in their view is always "person plus object" (2016, p. 1118) because meaning is generated by bringing together and processing interwoven information from both sources. Documents are therefore part of a "multi-dimensional structure" that suggests analysis can be similarly multi-layered, focusing in various ways on individual documents, on parts of documents, and/or on the networks in which they are enacted (2016, p. 1129).

Finally, Miller and Alvarado (2005) and Prior (2008) are among those who differentiate between research approaches that focus on "documents as content" and "documents as topics." The first of these approaches views documents as providing firm evidence about the world that can reveal novel, confirmatory or contradictory information; provide demographic and historical context; and be used unobtrusively to extend beyond the perspectives of interviewees and others in the field (Flick, 2009). Such approaches use content analytic strategies that inductively identify themes, patterns and categories in the data (Miller & Alvarado, 2005). The second approach positions documents as social products, representing a reality reflective of wider norms and values, rather than "independently adequate" reflections of fact (Miller & Alvarado, 2005, p. 351). Using this approach, documents may be viewed:

1. as "commentary," where the role of the researcher is to infer the "how" and "why" of the social reality suggested in the document — keeping in mind that the process of gathering meaning from a document depends as much on the life experience and assumptions of the reader as the intentions of the author (Coffey, 2014). Common analytic strategies used in this approach include the use of multiple sources or types of document for triangulation; rich description;

particular attention to divergent examples and explanations; and the use of theory to inform data analysis (Miller & Alvarado, 2005); and/or

2. as actors in their own right, interacting with other human and non human actors to both mirror and generate social reality. Prior (2008) suggests that documents should be seen not only as passive "informants," but as influencing human behaviour and even identity — for example, in constructions of ethnicity (Brown, 2010) or gender (Gazso, 2004); or in the classification of patients in a hospital setting (Prior, 2003). The focus of analysis in this approach is the document itself, and how it functions in patterns and structures of social activity (Miller & Alvarado, 2005; Prior, 2008).

Miller and Alvarado (2005) conceptualise the above approaches as part of a spectrum of diverse strategies within a single framework, providing access to "events that cannot be observed, to a species of formal communication about the social world, and to social actors that generate meanings and practices" (p. 353). Given the "inherent fluidity of the world, its actors and its objects" (Prior 2008, p. 825), however, the distinctions above are not necessarily exclusive, and documents may be usefully considered in parallel as both "topic" and "resource."

The following section now turns to an example of how the last of these frameworks was applied in a qualitative research study. It briefly describes the nature of the research, explains the choice of QDA framework and, lastly, demonstrates how that framework provided an effective scaffolding for the use of QDA in this case study.

A Case Study: Application of Qualitative Document Analysis

The research in question focuses on the ways in which female entrepreneurs in the Mary Valley, Queensland (Australia) engage in social practices that influence the wellbeing of their community. The first author, with personal links to the region established over the past decade, draws on PhD research interests linking three diverse disciplinary areas of rurality, entrepreneurship and wellbeing; and the ways in which entrepreneurship can be repositioned within a particular spatial environment as a community-level phenomenon (Anderson & Gaddefors, 2017) — that is, exploring a shift from "what makes *firms* successful … [to] what makes *communities* successful" (Fortunato & Alter, 2015, p. 445, italics in original).

The case study utilises the Miller and Alvarado (2005) framework described above to establish a social, historical and economic context for the research. This is particularly relevant because the Mary Valley region has faced significant social and economic challenges over the past decade, largely due to the 2006 proposal for a "mega-dam" in the geographical heart of the valley. The declaration triggered a community crisis that absorbed residents and others for over three years. While a forceful and strategic campaign was ultimately successful in the battle to "Save the Mary," the long-term uncertainty and stresses of the campaign, as well as subsequent government and community strategies to "revive" the depleted community, had wide-ranging and long term impacts.

The study takes a social constructionist approach in which knowledge and meaning are considered to be historically and culturally specific, constructed through the dynamics of interaction (Young & Collin, 2004). It demanded methods capable of capturing rich, multidimensional viewpoints, both contemporary interpretations (addressed by depth interviews), but also historical insights into change in the Mary Valley that could be provided only by documents.

Documents had the potential therefore to shed light on the context within which current research participants operate, including an understanding of historical change over time and

how that has impacted on their views/actions, as well as change and development in the community over the past decade (Bowen, 2009, pp. 29-31). The role of documents as "content containers" was thus pertinent, but it was clear that the social constructionist approach was consistent with a view of documents also as "resources" to shed light on social phenomena, and as "topics" in themselves, generating as well as reflecting social reality. The Miller and Alvarado (2005) framework described above therefore provided a useful way of thinking about what contribution documents could make to this research and informed the analytic process.

Steps in undertaking a QDA process have been covered, in various levels of specificity and from a range of perspectives, by authors including Altheide (2000); Altheide et al. (2008); Bowen (2009); Chenail (2011, 2012a, 2012b); Coffey (2014); Flick (2009); May (2011); and Olson (2010). The following discussion complements such guidance by providing a case study of a "QDA-in-action" process, and identifying some lessons learnt.

Constructing the Corpus

The first step was to establish the corpus, or the selected body of documents for analysis (Flick, 2009), with which to work. The Traveston Dam crisis was the subject of a multitude of public reports (government, consultant and community); theses, academic publications and feature articles; minutes, blogs and campaign material, identified through simple web and catalogue searches. A basic search using combinations of "Mary Valley," "Traveston Dam," or "Traveston Crossing Dam" (post 2006, when the dam was announced) against full text articles in the Proquest ANZ Newstand online database turned up more than 3000 articles and letters to the editor in local and national newspapers. Research participants have provided unpublished reports and pointed to other valuable documents that would not have emerged through the usual channels.

Sifting through the sheer volume of documents, particularly news media, was a challenge that absorbed substantial resources of time and energy. One means of addressing this difficulty may have been to narrow the search terms using combinations such as "Traveston Dam" with for example, "business" or "health." This strategy would however have required a set of initial assumptions about the effects of the dam proposal to guide the choice of search terms, with the consequent risk of the results turning up more-or-less only what was anticipated, rather than the richness of the unexpected. Given the social constructionist orientation and case study approach of the research, we chose to work with the broader set of results, staying "flexible to nuances, surprises, and confusion" (Altheide et al., 2008, p. 127).

The initial inspection or "first pass" document review (Bowen, 2009) was therefore particularly important in sifting the results to achieve a corpus of more manageable size and higher quality. The first author skimmed online the news headlines and abstracts across a strong sampling of the 3000+ results, discarding obvious news wire duplicates and articles that were clearly irrelevant, such as reporting of local sports results. Simultaneously, what was effectively a form of loose, preliminary content analysis was conducted (a focus on "documents as content"), as obvious categories recurred (for example, the closure of a local primary school attributed to the proposed dam) and items were included for consideration in the corpus until no new facts or perspectives appeared to be emerging on a topic. This process resulted in 339 news article abstracts or full text items downloaded to Endnote, where a further check identified items less relevant than first thought, and further duplicates. The total corpus of documents, imported finally into NVivo 10, consisted of 426 items that included, in addition to the news articles, a collection of reports (19), submissions to government (12), minutes of meetings (31), activist group newsletters (11) as well as a small number of presentations, videos, media

releases and campaign materials.[1] While the total is not much more than one-tenth of the results turned up by the initial searches, the corpus fulfilled Bowen's (2009) requirement for "quality" rather than "quantity."

Overall, this body of documents provided a rich and diverse source of material to inform a deeper understanding of the context in which the participants live and operate their businesses, and particularly the significant social and economic upheaval and change that has characterised the Mary Valley community over most of the last decade. Choosing an analytic approach required a method that enabled the researchers to explore shades of meaning and suggest more profound insights into social phenomena. While there exists a range of possible methods (see for example St. Pierre & Jackson, who highlight a diverse array of approaches from researchers who "borrow concepts, invent approaches, and create new assemblages that demonstrate a range of analytic practices" [2014, p. 717]), thematic analysis is a frequent choice by researchers engaged in QDA. Thematic analysis recognises that there are facets of the externally verifiable social world (Fincham, Langer, Scourfield, & Shiner, 2011) discoverable through documents and that coding, carried out with both rigour and creativity, can be an effective tool to obtain trustworthy findings and offer alternative insights into that social world.

Analysis

The process of establishing an initial corpus of documents for the thematic analysis provided opportunities to identify emerging broad topic areas relevant to the study. Using NVivo 10, the first author began to capture these in open or descriptive codes such as "small business networks" and "property resumptions," and to create a timeline of key events in the Valley over the relevant period. At this early stage of analysis, the aim was not to focus too obsessively on code names or definitions, and to avoid preliminary assumptions about how codes related to each other. The intent was rather to begin to develop a broad historical, social and economic context for the research by becoming deeply familiar with the documentary data.

To provide one example: with the eventual abandonment of the dam proposal in 2009, nearly 500 Mary Valley properties that had been resumed in readiness for construction became a significant liability for the Queensland Government (e.g., Strachan, 2011). There was considerable delay in returning the properties to the open market, with various arrangements including lease-backs to former owners and rentals to new residents operating in the interim until full return to private ownership of all the acquired properties in April 2015 (Courtney, 2015). Exploring the documents simply as "content containers" revealed that this sequence of events was generally agreed across the documentary corpus, but further insights into the complexity of the situation in the Valley at the time required a focus on "documents as topics," with context analytic strategies that sought to explore firstly, the "how" and "why" of the social reality inferred in the sources, as well as the role of the documents as social actors in their own right (Miller & Alvarado, 2005).

As the analysis progressed, therefore, codes were grouped around phenomena emerging in the documentary and interview data ("social impacts of the dam"), and concepts from White's (2010) conceptual framework for wellbeing which was guiding the research at this stage ("community goals and aspirations"), as well as other concepts from the literature ("rural idyll"). The codes and clusters were re-worked, using mind maps to reflect a growing understanding of context, and informed by strategies like those suggested later in this paper by Becker (1998) and Bazeley (2009). The codes were gradually formalised with the development of a codebook that captured definitions and "inclusion" and "exclusion" examples of each code.

---

[1] As would be expected, the corpus has continued to expand over time as the research has progressed.

This too was an iterative process, as deeper insights were gained into the data and definitions were revised accordingly (DeCuir-Gunby, Marshall, & McCulloch, 2011).

Coding and memos started to capture not only what was explicit, but also what was inferred in the data. To return to the above example, exploring the source and purpose of relevant documents suggested divergent views and layers of complexity in the post-dam environment not immediately apparent in the text. With the termination of the dam project, government media releases immediately announced a buyback scheme for previous owners. Seven months later, the government acknowledged the "relatively low interest" in the scheme (Queensland Government, 2010), and there were no further media releases on the Mary Valley buybacks from the then government for the remainder of its term. The absence, sparseness, or incompleteness of documents can also be suggestive (Bowen, 2009) — absence raising in this case the question of whether there might have been a lack of good news to be shared. Investigating the newsletter of a local community activist group shed further light on such an interpretation:

> There is a lot of uncertainty in the community at the moment and people want answers … As far as we can tell, the Queensland Government is still working out what they will do. (Save the Mary River Coordinating Group Inc., 2009, p. 3)

Paradoxes, contradictions and "negative cases or outliers" also became rich sources for analysis (Bazeley, 2009). In the absence of official information, local journalists continued to trace the Valley's story, providing emotive accounts of an emerging "economic, social and spiritual wasteland" (Fynes-Clinton, 2011a). Government disinterest and poor management were clearly identified as the source of emerging social problems, with articles quoting local concerns about the "rentals" living cheaply in government-owned properties — "ferals," drug addicts or thieves with no commitment to the community and no intention of staying long term (Fynes-Clinton, 2011b). The underlying message is of a community under siege from undesirable "outsiders," and at the continuing mercy of an uncaring government after three years of trauma. Alongside this account of "paradise lost" (Anonymous, 2010) however, ran a parallel discourse based on resilience and a commitment to rebuilding the community. The minutes of a local community group, for example, revealed clear themes of reclaiming the "heart and soul" of the Valley by involving the whole community; repairing and healing; and overcoming distrust, fear and anger (Mary Valley Renewal Team, 2009).

To return again to the Miller and Alvarado (2005) framework, viewing these documents also as social actors in their own right provided an additional lens to interrogate the data and added deeper dimensions to the analysis. The Fynes-Clinton (2011b) news article cited above, for example, could be posited to generate as well as capture community fragmentation; and to further embed negative views of government. Similarly, press releases of the incoming Queensland Government in 2012 (of a different political persuasion) are noteworthy in their number and persuasive messaging about the "debacle of massive proportions" instigated by the previous government (Queensland Government, 2012) compared to their new consultative and proactive approach, evidenced by regular reports of successful property transfers into private ownership and new business start-ups. As social actors, it could be surmised that such documents played a role in creating hope of a fresh start in the community (or alternatively, further embedding political scepticism, depending on the perspective of the reader).

The iterative nature of the analysis allowed these "educated hunches" to be investigated across data sets, one informing and/or supporting the other; and raising further questions for exploration. For example, the strong desire for an inclusive, participatory community engagement process to drive the future of the Valley had elements of paradox: set against the

discussion above about the unwelcome newcomers, it prompted the exploration of local perceptions of "community" through depth interviews, as well as revisiting literature on social inclusion and exclusion.

As the analysis deepened and possible connections between themes emerged, informed by the literature, the bones of a coherent "story" or model started to develop. This too was iterative, reworked over and over and using mind maps, charts and tables as thinking tools to find meaning and coherence in the evolving narrative. The insights that emerged from the documentary examples above became strands in a much more complex account of the contemporary and historical factors influencing women who currently live and run enterprises in the Mary Valley, and the ways in which they contribute to the wellbeing of their community.

The application of QDA in the case study above suggests the breadth of purposes to which QDA may be put, and the consequent but necessary challenges of recording key research processes and decision-making to ensure the legitimacy and rigour of the findings. Corbin and Strauss (2008) note that there is never only one story that can be constructed from qualitative data. Individual researchers focus on different aspects of data, interpret things in various ways, and identify a range of meanings, and the same analyst might look at the same data differently at different times. This, they suggest, is part of the "art" of qualitative research: the ability to construct a cogent, explanatory story from data, "a story that 'feels right' to the researcher" (Corbin & Strauss, 2008, p. 47).

This is not to deny the responsibility to tell the "story" in a way that also rings true to the reader. The flexibility of thematic analysis, while central to the value of the approach, can equally lead to inconsistency and a lack of coherence, particularly where the epistemological position and conceptual framework for the study is not made explicit (Nowell, Norris, White, & Moules, 2017). Establishing a reader's confidence in the findings requires the author to demonstrate the trustworthiness of the research process (Golafshani, 2003; Lincoln & Guba, 1985): its credibility or "truth value"; applicability or "transferability"; consistency or "dependability"; and its neutrality or "confirmability" (Lincoln & Guba, 1985; Wesley, 2014). Achieving these four criteria of trustworthiness requires the researcher to be explicit about the process used to interpret the evidence, the rationale for that process, and to be open with one's data, so that findings may be verified (Lincoln & Guba, 1985).

Observations from the Case Study: Rigorous Qualitative Document Analysis

There is an established and comprehensive body of literature addressing thematic analysis generally as an analytic tool (e.g., Corbin & Strauss, 2008; Miles & Huberman 1994; Saldaña, 2009), and the following discussion is clearly informed by such authors. Drawing on this scholarship, and the learnings of the above case study, we offer the following guidance on a QDA process based on thematic analysis. The process comprises four overlapping, interwoven "steps":

- establishing the initial corpus of documents, based on relevance to the research purposes;
- "open coding" that identifies broad topic areas in the data;
- "theoretical coding," that clusters open codes into themes and concepts; and
- creating a coherent story, that connects themes emergent from the data and the literature in a meaningful way.

While the "steps" are necessarily linear in presentation, the process is in reality both recursive and reflexive, drawing on an iterative process of document identification, concept development, data collection, data analysis and interpretation (through coding or other forms

of analysis) and representation that interweaves emergent insights from both the documentary evidence and other methods, such as depth interviews or participant observation (see Figure 1).
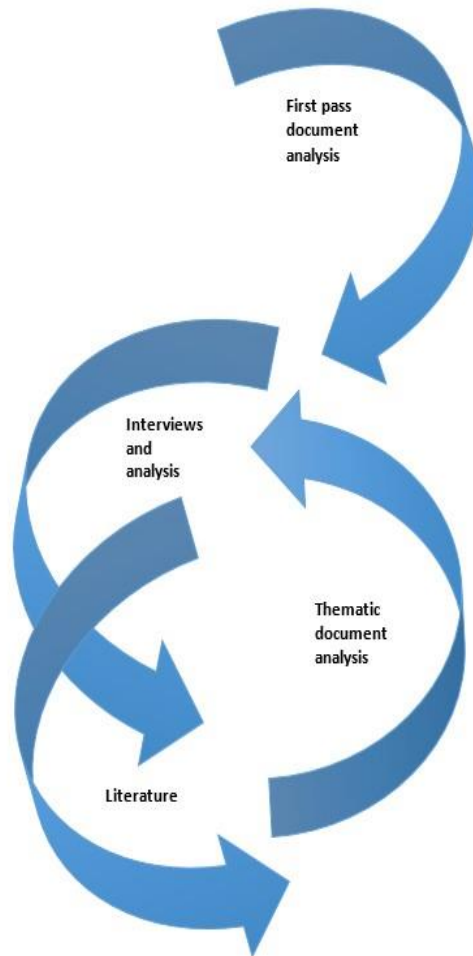

Figure 1: Iterative process of data collection and analysis

Establishing the Corpus

  Constructing the corpus is a multifaceted task that requires attention to database and other search terms; an initial inspection of the search results for relevance; and careful consideration to determine the most significant documents for the purpose of the research. In studies that require a broad scan of documentary evidence, the key is to establish a corpus that fulfils the requirement for "quality" rather than "quantity":

> Although it is a legitimate question, the concern should not be about "how many"; rather, it should be about the quality of the documents and the evidence they contain, given the purpose and design of the study. (Bowen, 2009, p. 7)

The choice of relevant search terms, as demonstrated in the case study above, will often entail a certain amount of trial and error, using one or more search terms in combination with others (Pickering & Byrne, 2014) or making justifiable decisions about the form, spelling or synonyms to be used (e.g., Sointu, 2005). This trial and error phase will overlap with what Bowen (2009) calls a "first pass" document review, which is particularly important in achieving a corpus of manageable size and high quality.

The "first pass" step discards obvious duplicates and articles that are clearly irrelevant for the purposes of the research. This is consistent with a process that Yin (2009) calls "triaging" the materials by their apparent significance to the inquiry; and also with O'Reilly and Parker's discussion of the need to be pragmatic and flexible, while gathering sufficient depth of information to answer the research question (2013).

The development of the corpus is often an iterative process, with the body of documents frequently reshaped as analysis progresses and gaps in the material or analysis appear (Flick, 2007). Altheide et al. (2008) describe for example how the analysis of their initial documents would lead to the discovery of other themes and usages, and consequently to further searches of other sources or with different search terms.

Open and "theoretical" coding

Given the often cyclic nature of corpus development, there is frequently no clear breakpoint between this first step and later stages of more focused analysis. Equally, the process of "coding" — the methods by which data are broken down, conceptualised, and put back together in new ways (Flick, 2009) — is an iterative one. While the rationale, terminology and procedures for various approaches to coding vary from author to author (see Flick, 2009 for a useful summary of three established approaches), these procedures should be neither "clearly distinguishable … nor … temporally separated phases in the process. Rather, they are different ways of handling textual material between which the researchers move back and forth if necessary and which they combine" (Flick, 2009, p. 307). The process of interpretation, however, starts as close to the data as possible, and moves progressively to the more abstract formulation of concepts and the relations between them, supported by memos that describe and explain the rationale for decision-making (Flick, 2009).

As documents are triaged for relevance, broad topics will start to become apparent. These can be captured through the use of "open" (e.g., Corbin & Strauss, 2008) or "descriptive" codes (e.g., Saldaña, 2009) that sort the data at a low level of abstraction (for example, "entrepreneurial networks") and start to provide early insight into the research topic, captured through the use of memos. Open coding can be somewhat akin to brainstorming, with what often becomes a multiplicity of codes. At these early stages of the analysis, the researcher wants to inductively "open up the data to all potentials and possibilities contained within them" (Corbin & Strauss, 2008, p. 160); although this does not preclude broad codes based on what could be anticipated in the data, or reflecting key categories in the literature (e.g., Beekhuyzen, Nielsen, & von Hellens, 2010). Bazeley (2009) notes that a priori categories or themes present no problem as long as they are identified and declared as such: the researcher can still be flexible and open to nuance and different emphases in the data. In their study of unauthorised online file sharing activities, Beekhuyzen et al. (2010) provide examples of codes drawn both from the literature (privacy, security, file sharing, Digital Rights Management) and from the data ("death of the album," "understanding of the law on file sharing").

As the researcher becomes increasingly intimate with the data, patterns and interrelationships start to become more evident (Saldaña, 2011); and open codes begin to be grouped around concepts in the theory being used (Beekhuyzen et al., 2010), and around relevant phenomena or "themes" emerging from the data — "like knots in the webs of our experiences, around which certain lived experiences are spun" (van Manen, 1998, p. 90). The coding names for the new groups in this more "theoretical" phase of coding are usually more abstract than those used in the first step (Flick, 2009) — for example, Beekhuyzen et al.'s "theoretical" codes include "challenging hegemony," "power relations," and "rationality in behaviour" (2010). Often this involves re-examining previously coded data, and recoding at the more theoretical level.

Creating a Coherent Story

As the coding journey winds from descriptive "labels" to more abstract "theoretical" themes, gaps appear in the corpus and new documents are added; memos proliferate and become more insightful; and possible relationships between codes and themes across both documentary and empirical data begin to emerge. The actual strategies or procedures that researchers use to discover and explore these relationships, however, are rarely described in publications of findings — partly because, as discussed above, this is the nexus between art and science: "not simply a technical matter, but a theoretical and conceptual issue that cannot be codified or abstracted into concrete rules of practice" (Gibson & Brown, 2009, p. 128). There are, however, "tricks of the trade" (Becker, 1998) that can assist the analytic process, including using comparison and pattern analysis ("describe, compare, relate") to refine and explore relationships between themes; utilising divergent views and negative cases to challenge generalisations; and visually capturing ideas and relationships using matrices, graphs, flow charts and models (Bazeley, 2009).

Richards warns of the danger of "garden path analysis" at this stage — the pleasant pathway, scattered with attractive themes, that leads nowhere (cited in Bazeley, 2009). Gathering these "attractive themes" into a coherent whole requires methods of discovery (or heuristics) based on "deductive" (what we conclude from established facts and evidence), "inductive" (what is inferred from the particular to the general, based on evidence and an accumulation of knowledge) and "abductive" reasoning (the educated "hunch" about the most likely explanation, based on an iterative dialogue between theory and data) (Järvensivu & Törnroos, 2010; Saldaña, 2009). This is the process of reconnecting the data in new ways (Flick, 2009) to create a meaningful and coherent story, a model or theory that is deeply connected to and draws from the data, and is informed by the literature (Bazeley, 2009).

Limitations

In the case study described above, the documentary evidence undoubtedly provided content, informing a timeline of events and identifying key actors and activities in the period of interest. There was little in the corpus that directly addressed the role of female entrepreneurs in the Mary Valley, and the documentary evidence could have been considered limited in this sense. Where documents are perceived and used solely as containers for content, issues of sample size, access, bias and relevance can be significant concerns (Bowen, 2009; Owen, 2014).

In this research, however, the purpose of the QDA process was to provide a broad historical, social and economic context for the research, suggesting threads of possibility to explore through other forms of data gathering, and providing insights into aspects of change and development in the Mary Valley over the past decade that supplemented, complemented and sometimes challenged the interview data.

In this way, the use of QDA was methodologically congruent with the social constructionist worldview of the researchers and the chosen analytic framework. As Coffey (2014) notes, acknowledging context can make some of the challenges of working with documents — such as the "often complex and hidden relations between documents, the selectivity of documentary accounts, the prescriptive and often formulaic structure of documents" — the very characteristics that make documents a rich source of data. "Limitations" become "considerations" that inform the analysis. It only becomes problematic where the researcher's approach is not explicit, and where the QDA findings appear as if by magic, unsupported by sufficient discussion of rationale, purpose, analytic procedures or decision-making.

Conclusion

It is clear that qualitative document analysis is used across wide range of disciplines, from various epistemological viewpoints, and for diverse research purposes. To return to the commonly accepted view of QDA as an emergent process focused on a search for underlying meanings, themes and patterns (e.g., Altheide et al., 2008), the above discussion suggests that there is no "one size fits all" approach to rigidly determine the nature of that search, and nor should there be. In this paper the application of QDA in the context of a case study provides argument for rigour in naming the assumptions underpinning the process, and explicitness about the methodological procedures employed.

In making this case, the article has contributed to the body of literature addressing the potential of qualitative document analysis as a methodological tool offering rich insights into complex social practices and environments. It provides a case study of "QDA-in-action" that demonstrates the value of a framework in sensitising the researcher to the different ways in which documents may be approached, and highlights some of the implications this may have for data collection and analysis. Finally, it draws from the learnings of the case study to offer some practical guidance on the application of a thematic analysis process in QDA.

As the range, nature and complexity of what might be called a "document" expands with the ever more pervasive influence of digital technologies, and the lines blur between public or private, primary or secondary, and over time and space (Coffey, 2014), it will become even more pressing to ensure that documentary evidence of our social existence is carefully and thoughtfully represented.

References

Altheide, D. L. (2000). Tracking discourse and qualitative document analysis. *Poetics, 27*(4), 287–299

Altheide, D. L. (2006). Terrorism and the politics of fear. *Cultural Studies ↔ Critical Methodologies, 6*(4), 415-439.

Altheide, D. L., Coyle, M., DeVriese, K., & Schneider, C. (2008). Emergent qualitative data analysis. In S. N. Hesse-Biber & P. Leavy (Eds.), *Handbook of emergent methods* (pp. 127–151). New York, NY: Guilford Press.

Anderson, A. R., & Gaddefors, J. (2017). Entrepreneurship and context: Is entrepreneurship research out of context? Dilemmas with (non) contextualised views of entrepreneurship. *Journal of Asia Entrepreneurship and Sustainability, 13*(4), 3–9.

Anonymous. (2010, 13 Nov). Trouble in paradise as eyesores infect valley. *Sunshine Coast Daily*, p. 14.

Bazeley, P. (2009). Analysing qualitative data: More than 'identifying themes'. *The Malaysian Journal of Qualitative Research, 2*(2), 6–22.

Bazeley, P. (2013). Foundations for thinking and working qualitatively. In P. Bazeley (Ed.), *Qualitative data analysis practical strategies* (pp. 3-31). Thousand Oaks, CA: SAGE Publications Ltd.

Becker, H. S. (1998). *Tricks of the trade: How to think about your research while you're doing it*. Chicago, IL: University of Chicago Press.

Beekhuyzen, J., Nielsen, S., & von Hellens, L. (2010). The NVivo looking glass: Seeing the data through the analysis. Paper presented at *5th International Conference on Qualitative Research in IT & IT in Qualitative Research* (QualIT2010), Brisbane, Australia

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40.

Brown, A. L. (2010). Counter-memory and race: An examination of African American scholars' challenges to early twentieth century K-12 historical discourses. *Journal of Negro Education, 79*(1), 54–65.

Chenail, R. J. (2011). Conducting qualitative data analysis: Reading line-by-line, but analyzing by meaningful qualitative units. *The Qualitative Report, 17*(1), 244–247. https://nsuworks.nova.edu/tqr/vol17/iss1/12

Chenail, R. J. (2012a). Conducting qualitative data analysis: Managing dynamic tensions within. *The Qualitative Report, 17*(Article 4), 1–6. https://nsuworks.nova.edu/tqr/vol17/iss2/2

Chenail, R. J. (2012b). Conducting qualitative data analysis: Qualitative data analysis as a metaphoric process. *The Qualitative Report, 17*(1), 248–253. https://nsuworks.nova.edu/tqr/vol17/iss1/13

Coffey, A. (2014). Analysing documents. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 367–380). London, England: SAGE Publications Ltd.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: techniques and procedures for developing grounded theory*. Thousand Oaks, CA: SAGE Publications, Inc.

Courtney, P. (2015, 19 Apr 2015). Mary Valley revival: Last of properties sold during Traveston dam project back in private hands. *ABC News [online]*. www.abc.net.au/news/2015-04-19/mary-valley-revival-last-of-properties-sold-back-in-private-hand/6402514

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.

Fincham, B., Langer, S., Scourfield, J., & Shiner, M. (2011). *Understanding suicide: A sociological autopsy*. Basingstoke, UK: Palgrave Macmillan.

Flick, U. (2007). *Designing qualitative research*. London, England: SAGE Publications.

Flick, U. (2009). *An introduction to qualitative research* (5th ed.). London, England: SAGE.

Fortunato, M. W. P., & Alter, T. (2015). Community entrepreneurship development: an introduction. *Community Development, 46*(5), 444-455.

Fynes-Clinton, M. (2011a, 2 July). Mary Valley homes rot as State Government dithers. *The Courier-Mail,* p. 7.

Fynes-Clinton, M. (2011b, 7 Feb.). Dead end — Locals beat the dam, but the Mary Valley town of Kandanga is losing the war against official neglect. *The Courier-Mail,* p. 22.

Gazso, A. (2004). Women's inequality in the workplace as framed in news discourse: refracting from gender ideology. *The Canadian Review of Sociology and Anthropology, 41*(4), 449–473.

Gibson, W. J., & Brown, A. (2009). *Working with qualitative data*. London, England: SAGE Publications, Ltd.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report, 8*(4), 597-606. https://nsuworks.nova.edu/tqr/vol8/iss4/6

Gorichanaz, T., & Latham, K. F. (2016). Document phenomenology: A framework for holistic analysis. *Journal of Documentation, 72*(6), 1114–1133.

Hughes, R. (2004). Employers' expectations of core functions, credentials and competencies of the community and public health nutrition workforce in Australia. *Nutrition & Dietetics: The Journal of the Dietitians Association of Australia, 61*(2), 105–111.

Järvensivu, T., & Törnroos, J.-Å. (2010). Case study research with moderate constructionism: Conceptualization and practical illustration. *Industrial Marketing Management, 39*(1), 100–108.

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

Mary Valley Renewal Team. (2009, 19 Nov.). Meeting notes.

http://sites.google.com/site/maryvalleyrenewal/minutes-of-meetings

May, T. (2011). *Social research: Issues, methods and process* (4th ed.) Maidenhead, UK: McGraw-Hill Education/Open University Press.

Miles, M., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed). Thousand Oaks, CA: SAGE.

Miller, F. A., & Alvarado, K. (2005). Incorporating documents into qualitative nursing research. *Journal of Nursing Scholarship, 37*(4), 348–353.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis. *International Journal of Qualitative Methods, 16*(1), 1-13. https://journals.sagepub.com/doi/pdf/10.1177/1609406917733847

O'Reilly, M., & Parker, N. (2013). 'Unsatisfactory saturation': A critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research, 13*(2), 190–197.

Olson, M. (2010). Document analysis. In A. J. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of case study research* (pp. 319–321). Thousand Oaks, CA: SAGE Publications, Inc.

Owen, G. T. (2014). Qualitative methods in higher education policy analysis: Using interviews and document analysis. *The Qualitative Report, 19*(26), 1–19. https://nsuworks.nova.edu/tqr/vol19/iss26/2

Pickering, C., & Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Education Research and Development, 33*(3), 534–548.

Prior, L. (2003). *Using documents in social research*. London, England: SAGE Publications Ltd.

Prior, L. (2008). Repositioning documents in social research. *Sociology, 42*(5), 821– 836.

Prior, L. (2010). Documents in health research. In I. Bourgeault, R. Dingwall, & R. de Vries (Eds.), *The SAGE handbook of qualitative methods in health research* (pp. 417-432). London, England: SAGE Publications Ltd.

Purtle, J. (2016). "Heroes' invisible wounds of war:" Constructions of post- traumatic stress disorder in the text of US federal legislation. *Social Science & Medicine, 149*, 9–16.

Queensland Government. (2010). Time to register interest in Mary Valley buy back scheme ends today [Press release]. http://statements.qld.gov.au/Statement/2010/6/30/time-to-register-interest-in-mary-valley-buy-back-scheme-ends-today

Queensland Government. (2012). Deputy Premier orders new strategy for Mary Valley [Press release]. http://statements.qld.gov.au/Statement/2012/4/19/deputy-premier-orders-new-strategy-for-mary-valley

Saldaña, J. (2009). *The coding manual for qualitative researchers*. London: SAGE Publications Ltd.

Saldaña, J. (2011). A survey of qualitative data analytic methods. In *Fundamentals of qualitative research* (pp. 89–138). Cary, NC: Oxford University Press.

Save the Mary River Coordinating Group Inc. (2009, 21 December). The Dambusters: newsletter of the Save the Mary River Coordinating Group Inc. www.savethemaryriver.com/drpl/node/106

Skinner, J., Stewart, B., & Edwards, A. (2004). Interpreting policy language and managing organisational change: The case of Queensland Rugby Union. *European Sport Management Quarterly, 4*(2), 77–94.

Sointu, E. (2005). The rise of an ideal: tracing changing discourses of wellbeing. *The Sociological Review, 53*(2), 255–274.

St. Pierre, E. A., & Jackson, A. Y. (2014). Qualitative data analysis after coding. *Qualitative Inquiry, 20*(6), 715-719.

Strachan, S. (2011, 18 Nov). Mary Valley land values halved. *The Gympie Times*, p. 3. www.gympietimes.com.au/news/mary-valley-land-values-halved/1177542

Van Manen, M. (1998). *Researching lived experience: Human science for an action sensitive pedagogy* (2nd ed.). Ontario, CA: The Althouse Press.

Wesley, J. J. (2014). The qualitative analysis of political documents. In B. Kaal (Ed.), *From text to political positions: Text analysis across disciplines* (pp. 135-160). Amsterdam; Philadelphia: John Benjamins Publishing Company.

White, S. (2010). Analysing wellbeing: A framework for development practice. *Development in Practice, 20*(2), 158–172.

Yin, R. K. (2009). Collecting case study evidence. In *Case study research: Design and methods* (4th ed., pp. 99-126): SAGE.

Young, R. A., & Collin, A. (2004). Introduction: Constructivism and social constructionism in the career field. *Journal of Vocational Behavior, 64*(3), 373–388.

Author Note

Leanne Wood is a doctoral candidate in Public Health and Health Promotion, School of Medicine at Griffith University, Australia; and a researcher for a large community organisation. Her research interests include community wellbeing particularly in rural communities, and qualitative methods for enhancing participant voice. Her recent publications have appeared in the *Journal of Family Studies* (2017), *Parity* (2017), and *Children and Youth Services Review* (2016). Correspondence regarding this article can be addressed directly to: leanne.wood@griffithuni.edu.au.

Dr Bernadette Sebar is a Senior Lecturer in Public Health and Health Promotion, School of Medicine at Griffith University, Australia. Her research interests focus on the impact of discourse and power on understandings of health and health behavior, with particular interests in gender and health, women and work, and feminist theory. Her recent publications appear in *International Journal of Drug Policy* (2017), *Health & Social Care in the Community*, (2016) and the *Journal of Health Psychology* (2016). Correspondence regarding this article can also be addressed directly to: b.sebar@griffith.edu.au.

Dr Nerina Vecchio is a Senior Lecturer in the Griffith Business School, Griffith University, Australia. Her research focuses on how public and private organisations can better serve the needs of frail and disabled individuals and their carers residing in the community, with recent publications appearing in *Health & Social Care in the Community* (2015), *Health Care Management Review* (2014) and the *Australian Health Review* (2013). Correspondence regarding this article can also be addressed directly to: n.vecchio@griffith.edu.au.

Article Citation

Wood, L. M., Sebar, B., & Vecchio, N. (2020). Application of rigour and credibility in qualitative document analysis: lessons learnt from a case study. *The Qualitative Report, 25*(2), 456-470. https://nsuworks.nova.edu/tqr/vol25/iss2/11