Theory and the breadth-and-depth method of analysing large amounts of qualitative data: a research note

Rosalind Edwards[1] · Emma Davidson[2] · Lynn Jamieson[2] · Susie Weller[3]

Abstract

This research note builds on a previously published discussion of the 'breadth-and-depth' method for working with extensive amounts of secondary qualitative data, to consider the way that theory can be used and developed as part of this method. We illustrate potential deductive, inductive, and abductive logics of the relationship between theory and data that can be pursued using the method, but note that in reality research analysis rarely proceeds along such clear categorical lines. Rather, qualitative researchers are more likely to pursue a flexible retroductive logic and analytic practice that the breadth-and-depth method also can accommodate.

Keywords Breadth-and-depth method · Secondary qualitative data · Theory · Retroductive analysis

1 Introduction

In this research note we build upon our published discussion of how to work with extensive qualitative data, to consider the relationship between the methodology and theory. Our article 'Big data, qualitative style: a breadth-and-depth method for working with large amounts of secondary qualitative data' was published in *Quality &Quantity* (Davidson et al. 2018). It puts forward an original 4-step approach for working with a large corpus of qualitative data, undertaking an iterative analysis that can deal with substantial amounts of data but nonetheless retain the rigour, integrity and nuance of qualitative approaches. The article has been accessed over 6500 times at the time of writing, and citations are growing. As a result of the publication we have been invited to present papers on the breadth-and-depth method in the UK and internationally.

Some of the questions that we have been asked are concerned with the theoretical substance of the breadth-and-depth method and the various ways users of the method can work

✉ Rosalind Edwards
r.s.edwards@soton.ac.uk

1   SSPC, University of Southampton, Southampton SO17 1BJ, UK

2   Centre for Research on Families and Relationships, University of Edinburgh, 23 Buccleuch Place, Edinburgh EH3 9LN, UK

3   Clinical Ethics and Law (CELS), University of Southampton, Southampton SO17 1BJ, UK

with theory. In this research note we indicate and illustrate the potential ways in which users of the method may use and develop theory. It is not our intention to advocate for any particular relationship between the breadth-and-depth method and theory, or to engage in debates about the relationship of their relational logics. Rather, this note provides pointers for researchers to consider explicitly their own theory routes working with the breadth-and-depth method.

2 The breadth-and-depth project and method

The breadth-and-depth approach allows researchers to manage and analyse a large volume of qualitative secondary data, yet retain the distinctive order of knowledge about social processes, context and detail that is the hallmark of rigorous qualitative research. It resulted from an ESRC National Centre for Research Methods methodological project that examined the possibilities for developing new procedures for working across multiple sets of archived qualitative data: http://bigqlr.ncrm.ac.uk/. The project was a response to two methodological developments. One is the increasing practice in the UK and more widely of archiving qualitative data from primary research projects for sharing and reuse, raising the prospect of merging qualitative data from several discrete studies. The other methodological development concerns the use of computational processing tools to manipulate enormous amounts of data speedily. This can prioritise quantitative knowledge by default, so our methodological project was an attempt to insert quality into the quantity default of 'big data' and to retain a form of analysis where computational text mining of large volumes of qualitative data sits alongside and is equal with 'deep data' research approaches.

The breadth-and-depth approach we developed in this methodological context involves four steps:

1. An overview survey of archived qualitative data that is available in an archive or from several archives, selection of data from relevant projects, and construction into a merged corpus for the secondary study;
2. Recursive surface thematic mapping of the contents of the large merged qualitative data set through keyword analysis using computational techniques, to indicate themes where it might be fruitful to conduct investigations;
3. Sampled preliminary examinations based on step two, working with short extracts of data that contains the keyword themes to see if the content speaks to the secondary researchers; and
4. Working with whole cases for in-depth interpretive analysis.

The method is an iterative process where steps 3 or 4, for example, may throw up issues that require return to step 2 or even step 1 (Davidson et al. 2018; Edwards et al. 2019).

Our intention in developing the breadth-and-depth method was methodological—providing a step-by-step process that any researcher or team of researchers could use whatever the theoretical logic, substantive topic and nature of qualitative data for their project. Nonetheless, within the 4-step process, each step is driven by theoretically informed research questions and a theoretical framework that underlies the secondary study. This includes the 'design' of the data assembly that purposively brings data together for comparison or complementarity, and the logic of the keyword themes that are mapped. The units of analysis or cases that are selected as short extracts and then for in-depth analysis from the multiple

archived small-scale qualitative studies that form the secondary corpus will relate to the theoretical standpoint as well as the substantive research topic and questions of the secondary study. Further, at the purpose level, key benefits of secondary analysis of 'big qual' include the ability to scope out new research questions that could not be answered by individual projects, in particular allowing the possibility of stronger theoretical generalizations about how social processes unfold, and the testing of ideas against evidence. Thus both the process and purpose of the breadth-and-depth method of secondary analysis of big qual raise issues of the relationship between theory and data.

3 Potential relationships between theory and data in the breadth-and-depth method

The different ways of seeing and explaining the social world, of making knowledge claims of various levels of complexity and of conceptualising substantive research issues, that comprises 'social theory' come together in various relationships with the material constructed or generated from or within the social world that comprises 'data', including big qual. In practice this is often quite a complex process, an issue we pick up on later. Nonetheless, methodologists describe three ideal-type logics of the relationship between theory and data in qualitative research: deductive, inductive and abductive, and some also include a fourth, retroductive (see e.g. Åsvoll 2014; Blaikie 2007; Kennedy and Thornberg, 2017; Timmermans and Tavory 2012). There are some differences of emphasis between researchers in how they portray these model logics as well as debates about the relationship between them. Nonetheless, there are some common features in discussions of each of the models which we outline briefly below, and then note how each may be enacted within the 4-step breadth-and depth method of working with large amounts of qualitative data. To illustrate, we briefly refer to a substantive example that served for our development of the methodology working with a corpus of qualitative longitudinal projects: shifts in vocabularies and practices of care and intimacy over time by gender and age cohort.

3.1 Deduction

In the deductive model of the relationship between theory and data, researchers will start from an existing theory, and adopt it as an analytical lens to guide attention to detail in a specific substantive field. The logic of movement is from the general theory to the specific empirical; taking data, applying a theoretical framework to them, and then using that theory to deduce a 'why' explanation for the empirical findings. Adopting the substantive example of shifts of vocabularies and practices of care and intimacy, researchers following a purely deductive logic could start from a theory or hypothesis that predicts convergence in men's and women's practices over time and hence vocabularies of care and intimacy, and deductive analysis would involve looking for instances of care and intimacy that are convergent over time to fit or falsify the theory. (As we will note later, though, in practice qualitative researchers are usually more open to unexpected patterns in data and rarely follow a purely deductive logic.)

In a deductive approach to research utilising the breadth-and-depth method researchers can undertake step 1—the identification of sets of archived data suitable for secondary analysis, driven by the issue of whether or not each discrete data set and their construction into a merged corpus enable the hypothesis or theory to be applied to the specific

substantive focus of the project. The mapping of keywords and identification for themes of step 2, gaining a sense of the data corpus, would be guided by the theoretical framing. In a weak form of deductive reasoning, alertness to the relevance of particular words as 'keywords' or particular word co-location and clustering as 'themes' would be guided by the theoretical framework. In a strong form, searches would proceed using predetermined word, word clusters and co-locations. The extracts of text for step 3 investigation would be those that show particular promise of throwing light on how the theory works in relation to the study's research questions, and the selection of materials for the in-depth analysis of step 4 similarly showing how theoretical 'why' processes play out in richer detail. A singularly-minded deductive approach to the breadth-and-depth method is likely to work through the steps in order, as a guided and straightforward process, rather than needing to move backwards and forwards between them, with the aim of testing ideas about evidence. There remains the possibility, however, that by the end of step 2 researchers may realise that their chosen corpus cannot reveal anything relevant to their hypothesis and need to return to step 1.

3.2 Induction

The logic of an inductive model is that the meaning interpreted from data is the basis for inferring theoretical statements about the nature of the social world and generalisation of substantive findings. The logic of movement is from the specific empirical to the general theory. Researchers are as open as possible to theory emerging from data without preconception about the outcomes. So, in our practices and vocabularies of intimacy and care example a strictly inductive approach might lead to dispensing with any hypothesis or theory about any shifts by gender and age cohort, or (more often) loosely working inductively within the deductively derived categories of gender and age cohort.

Thus the step 1 precursory understanding of the nature and quality of the available sets of archived qualitative data will be shaped by an inductive logic that is fairly wide ranging, locating data on a broad topic area that fits with the secondary analysis research topic and the potential for generating theory. Step 2 mapping across the corpus would involve identifying frequent words and themes that are generated by the outcomes of text mining or automated semantic analysis without a framing of the parameters and nature of the search. The next step, 3, would be to undertake preliminary examination of features of the recursive mapping that occur together frequently, following the regularities that offer the potential for inferring theory about the identified 'what' and 'how' issues through into greater depth in step 4. An inductive logic is also likely to involve a more iterative approach to the breadth-and-depth method as interesting features and patterns are identified in steps 3 and 4. In turn, that may require a return to step 1 and identification of more, different, data to merge into the corpus for exploring processes in other contexts in an effort to develop stronger theoretical generalisations about how social processes unfold.

3.3 Abduction

Abduction looks for and explores potential 'what', 'why' and 'how' explanations, seeking or abducting theory through the identification of theory gaps and data anomalies. The abductive model involves an iterative filling of a theoretical gap in a particular substantive field, putting together theories from quite different fields, moving back and forth between data and theories, making comparisons and interpretations, and rethinking and refining

best possible, plausible explanations. It is a logic of movement that actively seeks out and moves from general theory gap to specific empirical puzzle (in light of existing theories) to novel theory explanation. Taking our example of researching shifts in vocabularies and practices of care and intimacy over time by gender and age cohort, in the context of theories that predict gender convergence alongside other theories that account for continuing gendered inequalities of care and intimacy, this could mean systematic 'inductive attentiveness' to surprising evidence of, say, younger cohort men and women displaying more heavily gendered differences in their practices and vocabularies, alongside 'deductive attentiveness' to those where it is the case.

An abductive logic in step 1 of the breadth-and-depth method will seek out the potential of the available relevant archived data sets for throwing up unusual phenomena that have the potential to fill a theoretical gap and to stimulate the abductive process. Step 2 is likely to involve bringing together searches based on lists and also open-ended searching for patterns and regularities. But in the latter case, the regularities would enable the pointing up of where there were irregularities and unexpected coincidences to pursue. Such anomalies could be followed up with close attention to examination of data extracts in step 3, to see if there were indications of plausible 'what' and 'why' theoretical explanations. Indeed, abduction is often applied in qualitative interview research through close attention to passages of text, with ideas about meaning and significance abducted from the segment. Cases for deeper examination in step 4 would be chosen because step 3 indicated that they held out the promise of insightful comparative differences and unusual features to enable the building of new theories or pulling together of two quite different theories to fill a gap in knowledge. Recursive moving between steps is a feature, for example if provisional and depth examinations in steps 3 or 4 showed up a surprising juxtaposition of features, this would entail revisiting the keyword themes search of step 2 to look for relevant irregularities that would then point to places in the corpus to explore initially in provisional and selectively in greater depth, which in turn might suggest other anomalies to pursue.

4 Retroductive relations between logics: a conclusion

Logics of inquiry are often idealised, sanitised versions of the way qualitative research proceeds. In reality theory and data analytic processes may be quite messy, as we have indicated at points in our discussion above. Deductive researchers often are open to re-thinking and challenging theory, while inductive researchers are never a completely uninformed *tabula rasa*. Inductive logic can involve an element of deduction when working with the data, such as a prior orientation, or deducing further research questions to explore during analysis. Deductive logic can allow testing of conclusions at different stages of the research process, where in terms of the breadth-and-depth method theory may be identified through induction applied in steps 1–4, and then steps 2 and 3 repeated deductively using that theory. Further, all three of the relationships between theory and data outlined above may be going on simultaneously. Abductive logic purposefully utilises unusual features of deductively or inductively generated findings to develop plausible explanations and generate new theory. In practice or as planned, there are also what are referred to as retroductive logics to the relationship between theory and data, which posit complementary or overarching combinations of deductive, inductive and abductive in the oscillation, backtracking and creative process that is social research [see various perspectives on this in Åsvoll (2014), Chiasson (2001), Kennedy and Thornberg (2017) and Timmermans and Tavory (2012)].

Our own work as a team of four researchers developing and utilising the breadth-and-depth method has been more along in practice retroductive lines as we sought to chart new territory, developing and then applying our method for working with extensive amounts of secondary qualitative data (e.g. Edwards et al. 2019).

What we hope that we have made explicit through this research note, whatever the logic—following pure deductive, inductive or abductive approaches, or a purposeful or eclectic retroductive process—the iterative steps of the breadth-and depth-method of analysing extensive amounts of qualitative data can encompass flexibly a range of articulations between theory and data.

References

Åsvoll, H.: Abduction, deduction and induction: can these concepts be used for an understanding of methodological processes in interpretive case studies? Int. J. Qual. Stud. Educ. **27**(3), 289–307 (2014)

Blaikie, N.: Approaches to Social Enquiry: Advancing Knowledge, 2nd edn. Polity Press, Cambridge (2007)

Chiasson, P.: Abduction as an aspect of retroduction. In: Bergman, M., Queiroz, J. (eds.) The Commens Encyclopedia: The Digital Encyclopedia of Peirce Studies, (**online publication**). (2001) https://www.commens.org/encyclopedia/article/chiasson-phyllis-abduction-aspect-retroduction.

Davidson, E., Edwards, R., Jamieson, L., Weller, S.: Big data, qualitative style: a breadth-and-depth method for working with large amounts of secondary qualitative data. Qual. Quant. **53**(1), 363–376 (2019)

Edwards, R., Weller, S., Jamieson, L., Davidson, E.: Search strategies: analytic searching across multiple data sets and within combined sources. In: Hughes, K., Tarrant, A. (eds.) Qualitative Secondary Analysis, pp. 79–100. Sage Publications, London (2019)

Kennedy, B.L., Thornberg, R.: Deduction, induction and abduction. In: Flick, U. (ed.) The SAGE Handbook of Qualitative Data Collection, pp. 49–64. SAGE, London (2017)

Timmermans, S., Tavory, I.: Theory construction in qualitative research: from grounded theory to abductive analysis. Sociol. Theory **30**(3), 167–186 (2012)