

Qualitative data preservation and sharing in the social sciences: On whose philosophical terms?

Natasha S. Mauthner and Odette Parry

Abstract

Over the past decade, an academic debate has developed surrounding qualitative data preservation and sharing in the social sciences, and has been characterised as one between supporters and opponents of this movement. We reframe the debate by suggesting that so-called 'opponents' are not resistant to the *principle* of data preservation and sharing, but ambivalent about how this principle is being put into *practice*. Specifically, qualitative researchers are uneasy about the foundational assumptions underpinning current data preservation and sharing policies and practices. Efforts to address these concerns argue that the inclusion of the 'contexts' of data generation, preservation and reuse will adequately resolve the epistemological concerns held by the qualitative research community. However, these 'solutions' reproduce foundational assumptions by treating 'context' as ontologically *separate* from, rather than *constitutive* of, data. The future of qualitative data preservation and sharing in the social sciences is dependent on shedding its implicit unitary foundational model of qualitative research, and embracing 'epistemic pluralism' and the diversity of philosophical perspectives representing the qualitative researcher community.

Keywords: qualitative research, data preservation and sharing, epistemology

Introduction

Although the move towards qualitative data preservation and sharing has been regarded as a 'new' tradition within the social sciences, social science archives have a history that goes back many decades. In Britain, the Mass Observation archive was set up in the 1930s to create 'an anthropology of ourselves' through the collection of observations and writings about the everyday lives of ordinary people (Mass Observation 2009). In the United States, the Human Resources Area Files is an archive that was initiated in the 1950s and comprises an anthropological collection of primary, published and unpublished, ethnographic sources on selected cultures from around the world (Lagacé, 1978). There is, consequently, a tradition of social scientists drawing on these and other archival materials to explore historical, sociological and anthropological questions.

Given this established tradition of archival research within the social sciences, why has qualitative data preservation and sharing emerged as an issue for debate amongst academic researchers over recent years, and created controversy? The debate in the UK arose in the mid-1990s as a result of changes in the funding conditions of the Economic and Social Research Council (ESRC), the UK's largest funding agency for research and postgraduate training relating to social and economic issues. In 1996, the ESRC implemented a Datasets Policy (modified in 2000) which introduced two new requirements on potential and actual ESRC grant-holders. First, grant applicants were obliged to demonstrate that data similar to those they were proposing to generate did not already exist. Second, grant holders were contractually obliged to offer their research data and associated materials for archiving within three months of the end of their project (ESRC 2000). Prior to 1996, individual or teams of researchers may have considered whether to lodge their 'data' in an archive, and whether to use existing archived data in the conduct of a new study. The introduction of the ESRC policy, however, meant that qualitative researchers working within academic institutions, and certainly those seeking funds from the ESRC, had to consider, often for the first time, issues of data preservation and sharing because they were required to do so by the new policy. Since then, many more academic research funding agencies in the UK and elsewhere have introduced data preservation policies (see SHERPA 2009a).

The ESRC policy was greeted with mixed reactions by qualitative social scientists in the UK, and some publicly voiced their ambivalence by pointing to both its potential benefits and drawbacks (Alderson, 1998; Griffin, 1998; Hammersley, 1997; Mauthner et al., 1998; Parry and Mauthner, 2004, 2005). These researchers saw the value in preserving significant social scientific studies; in demystifying qualitative research by revealing its processes of knowledge construction; and in using archived studies for historical and methodological research, as well as for teaching purposes. However, they also expressed ethical, epistemological and political concerns regarding the disclosure of personal information relating to the researchers; the difficulties of providing ethical assurances to respondents about how their data might be reused in the future by third parties; the power differentials between potential data 'users' and the

data ‘suppliers’ (respondents); and, the epistemological challenges of using ‘data’ taken out of their contexts of production. These researchers were seeking to initiate a debate not so much about whether qualitative researchers should engage in data preservation and sharing activities, but rather how they could do so without compromising the ethical and philosophical integrity of their research. They were explicitly inviting the community of qualitative researchers to reflect critically on the policy, to consider what it was asking qualitative researchers to do, and to discuss its potential implications for qualitative research practice.

Since then, qualitative researchers have continued to express *ambivalence* over the issue of qualitative data preservation and reuse. The idea of preserving a scholar’s life’s work is welcomed by many (Corti, 2000). Some researchers express strong ethical imperatives vis-a-vis their participants and their funders to make accessible publicly funded research (e.g. Broom et al., 2009; Davies, 2004). Henderson, Holland and Thomson (2006), for example, explain their motivations for archiving their study:

An important [reason] was both ethical and practical. Despite a good track record on publishing and disseminating the component studies ... the sheer volume of data and the degree of investment in its collection was such that it increasingly felt irresponsible *not* to find ways of giving wider access to this unique resource. Despite familiar concerns about compromising the confidentiality we negotiated with our participants at each round, it seemed far more unethical to shy away from exploring solutions to these concerns and effectively admit that this data could not be used. We owed it to both to study participants and funders alike.

This notion of data as a community or public resource and ‘public knowledge’ has been expressed by others (Broom et al., 2009; also Silva, 2007). In contexts where data are being collected from certain communities there is an additional imperative to preserve cultural heritages, particularly for later generations (Broom et al., 2009).

Clearly, qualitative researchers can see important intellectual, ethical and cultural reasons for wanting to preserve and share data. Nevertheless, some qualitative social scientists in the UK and elsewhere remain sceptical and continue to express ethical and epistemological concerns (Bishop, 2005; Boddy, 2001; Broom et al., 2009; Corti, 2000; Fielding, 2003; Heaton, 2004; Henwood and Lang, 2005; Thompson, 2000). Most recently, Broom et al.’s (2009: 1167) Australian study of social scientists’ attitudes towards archiving and reusing data suggests “significant reticence among qualitative researchers to the establishment” of a qualitative data archive. Moore (2007) notes that in the UK, the ‘perceived injunction to archive data has been met with resistance by recalcitrant researchers who are wary of the implications of depositing data, and the possibilities of reusing data.’ Moreover, Mason (2007) points to a ‘culture of uneasy suspicion that there might be something ethically, morally or

epistemologically dubious about ‘re-using’ qualitative data.’ Certainly, levels of sharing and re-use of qualitative data have been low in the UK (Fielding, 2003; Heaton, 2004; Henwood and Lang, 2005); there has been a general reluctance to draw on material created by other research teams (Thompson, 2000); and, most secondary studies have involved re-use of researchers’ own data (Heaton, 2004). Indeed, some argue that researchers should reuse their own data rather than that of others (Bell 2004).

There has been a tendency within the literature to interpret these ‘dissenting’ voices as *resistance* and *opposition* to the *principle* of data archiving and reuse. This explains why much of the debate over the past decade has focused on the possibility, feasibility and desirability of creating and using qualitative research archives in the social sciences (e.g. van den Berg, 2005). However, we interpret these voices as expressing *ambivalence*: they are supportive of the *principle* of data archiving and reuse, but cautious, hesitant and uneasy about how this principle is being put into *practice*. Specifically, we suggest, they are concerned about taking part in data preservation and sharing practices on *foundational* terms for two reasons. First, because foundationalism represents a *particular* philosophical theory of knowledge that not all qualitative social scientists subscribe to. Foundationalism assumes that there are basic, self-evident or foundational beliefs that require no justification and that are premised on a strict ontological distinction between an external objective material world and an internal subjective human world. Foundationalism is rooted within a Cartesian understanding of our relationship to the world, premised on separation and detachment, rather than relationality and engagement, as our ontological way of being. ‘Data’ (knowledge) are conceptualised as separate from the subjectivities that generate them, and independent of the relational and intersubjective contexts that give rise to them. ‘Data’ and knowledge are understood as representative in character, and are seen to be produced through the ontological separation of the knower from the known. Second, while foundationalism represents one of many perspectives on knowledge and its production, it claims ‘epistemic sovereignty’ (Rouse, 1996a, 1996b) and a controlling moral authority over knowledge. It advocates ‘epistemic monism’ (Mauthner, 2009) over ‘epistemic pluralism’ (Healy, 2004) and claims epistemic supremacy for itself while denying the epistemic status and legitimacy of other perspectives on knowledge. In this sense, foundationalism is a form of ‘epistemic ideology’ in that it confers upon itself a normative and hegemonic status. Our argument is that contemporary data preservation and sharing debates, discourses, policies and practices are embedded within these foundational understandings of knowledge and its production, and that this creates tensions for academic qualitative researchers who draw on *other*, non- or post-foundational, theories and practices of knowledge production.

These concerns over the philosophical terms of engagement with data preservation and sharing activities echo those of the early critics of qualitative archiving. Indeed, many would suggest that discussions over the past ten years have successfully addressed the concerns through, in particular, the inclusion and analysis of the ‘contexts’ of data generation, preservation and reuse. Our

argument, however, is that these theoretical and practical ‘solutions’ neither allay the concerns of many, nor address what are more deep-rooted philosophical tensions. This is because they continue to consider the question of how to create and use qualitative archives in the social sciences within terms that are *implicitly* foundational.

The purpose of our article is to illustrate how the qualitative data preservation and sharing ‘project’ within the social sciences has been developed and promoted on implicit foundational terms. In the first part, we point to the foundational genealogy of this project and indicate how it is growing out of broader international data preservation and sharing movements that concern primarily quantitative data, and that are premised and promoted on foundational terms. In the second part, we discuss what we call ‘modernist’ approaches to qualitative data archiving and reuse. We suggest that these approaches draw on implicit foundational assumptions, discourses, and justifications in promoting qualitative data preservation and sharing practices, and in ‘fixing’ the epistemological challenges that arise out of this position. In the third part, we examine what we call ‘postmodern’ approaches to qualitative data archiving and reuse. These more recent developments draw on postmodern notions of multiplicity in promoting and supporting qualitative data preservation and reuse. They dismantle the epistemological challenge by suggesting that all researchers, whether ‘primary’ or ‘secondary’, can generate legitimate interpretations of data. We discuss these contributions and indicate how they too are based on implicit foundational terms. We conclude the article by suggesting that the future of qualitative data preservation and sharing in the social sciences is dependent on moving away from ‘epistemic monism’ and an implicit unitary foundational model of qualitative research, towards developing policies and practices that embrace and accommodate ‘epistemic pluralism’ (Healy, 2004) and the diversity of philosophical perspective and approach amongst qualitative researchers.

The international data preservation and sharing movement

The development of the qualitative data preservation and sharing project must be understood within the context of a broader international movement towards the preservation, sharing and reuse of a wide range of ‘data’ useful, but not necessarily collected, for research purposes (e.g. data generated via administrative records) (OECD 2007; ESRC 2008; UKRDS 2008, 2009; RIN 2008). While the issue of qualitative ‘data’ is occasionally raised within these documents, the main focus of these international-level discussions is large quantitative datasets. This global context is critical because it constitutes the genealogy of qualitative data preservation and sharing and has played a key part in shaping associated policies, practices and discourses.

Databases of all kinds are increasingly being seen as ‘an essential part of the infrastructure of the global science system’ (OECD 2007: 3; see also ESRC 2008; UKRDS 2008, 2009). This move is linked to the availability, development and application of advanced computing and information technologies over recent years, which have resulted in enormous growth in the volumes of

research and other data being generated, and have facilitated and enhanced the possibilities of data sharing. Thus, the development of national capability and capacity for the management of digital research data is increasingly regarded as a key component of the national e-infrastructure in the UK (HMSO 2004; OSI 2007) with other advanced economies also starting to make significant investments in this area (UKRDS 2009). This in turn is leading to funding organisations 'increasingly requiring grantees to deposit their raw research data in appropriate public archives or stores, in order to facilitate the validation of results and further work by other researchers' (SHERPA 2009b).

The creation of larger scale national and international facilities for data access and sharing are seen to have the potential to create 'a high quality research infrastructure' (UKRDS 2009: 1) which will bring great benefits to individuals and society as a whole (see OECD 2007; ESRC 2008; UKRDS 2008, 2009; RIN 2008). Specifically, data preservation and sharing are seen to bring the following economic, scientific, ethical and professional benefits:

Economic investment and cost-effectiveness of valuable resources: Data sharing and better exploitation of existing data sets are understood to add value and increase the returns from public investment and reduce costs by avoiding duplicating data collection efforts. As the UKRDS states: 'Research data have become a valuable resource that needs to be maintained for future access and re-use if we are to reap the full benefits of the UK's investment in research ... The ability to share research data, minimising the need to repeat work in the laboratory, field or library, thus saving time and effort' (UKRDS 2009:1).

Public access to publicly-funded research: The principle of making publicly funded research data openly available and accessible to the public is endorsed (OECD 2007).

Scientific innovation: Data sharing is understood to provide a means for scientific innovation by promoting new research, new questions, new methods and new insights. Topics that were not envisioned by the initial investigators can be explored and new research areas can be identified for collaborative study. The ability to retrieve and compare data from multiple sources easily, can lead to the testing of new or alternative methods, and to 'powerful new insights' (UKRDS 2009: 1). The UKRDS report notes that 'research data will increasingly be the starting point for new research as well as a key output' (UKRDS 2009: 1).

High quality research through the creation of new datasets: Combining datasets from different departments, agencies and sources allows the creation of new datasets which can facilitate high-quality, policy-relevant research by providing a fuller picture rather than analysing separate pieces of a jigsaw (ESRC 2008).

Transparency in scientific inquiry: Data sharing is understood to reinforce open scientific inquiry thereby improving methods of data collection and measurements through the scrutiny of others.

Intellectual capital: Funding agencies see themselves as having significant stewardship roles for the substantial investment they make in the British scientific infrastructure, and are seeking to preserve inventories of data sets they have funded. Writing about the Medical Research Council, Davis (2004: 212) notes that: ‘With the retirement of a generation of scientists came the realisation that resources of incalculable intellectual value were in danger of being lost for want of an adequate asset management policy’.

Training: Data depositories are seen to provide an important resource for training in research by enabling new researchers to utilise existing data.

Reduction of burden on participants: Data preservation and reuse are understood to reduce the burden on respondents caused by multiple data collection efforts.

Greater data security: Data storage and preservation are seen to reduce the information security risks associated with maintaining duplicated datasets in more than one location.

The key benefits of data preservation and sharing are articulated in, for example, the Social Sciences and Humanities Research Council of Canada Research Data Archiving Policy:

Sharing data strengthens our collective capacity to meet academic standards of openness by providing opportunities to further analyse, replicate, verify and refine research findings. Such opportunities enhance progress within fields of research as well as support the expansion of inter-disciplinary research. In addition, greater availability of research data will contribute to improved training for graduate and undergraduate students, and, through the secondary analysis of existing data, make possible significant economies of scale. Finally, researchers whose work is publicly funded have a special obligation to openness and accountability. (SSHRC 2009)

These discourses promoting the economic, scientific, ethical and professional benefits of data preservation and sharing suggest that the collection, preservation, sharing, exploitation, access, availability, retrieval, comparison, and combination of ‘data’ are *unproblematic* activities. They do so, we suggest, because *ontologically* ‘data’ are implicitly viewed in foundational terms: ‘data’ are seen as being ‘out there’, separate from the conditions of their production and from the subjectivities through which they are produced. The meaning of ‘data’ is understood as intrinsic to them and independent of the contexts that give rise to them.

‘Modernist’ approaches to qualitative data preservation and sharing

The development of qualitative data preservation and sharing policies and practices has taken place within the context of these broader international discourses, and has adopted similar rationales for promoting qualitative data preservation and sharing in the social sciences. For example, Corti and

Thompson (2004: 341) suggest that data reuse 'is resource efficient, optimising the use of pre-existing material'. Archived data are also seen to offer the possibility of conducting comparative research, and providing historical, methodological and teaching resources. Arguments are also put forward concerning intellectual capital, and the importance of saving intellectual resources by 'rescuing' classic studies and archiving contemporary ones (Corti, 2000).

The scientific rationales for data preservation and sharing have also been taken up. It is specifically this notion that qualitative data archiving and reuse can be justified on 'scientific' grounds that we want to focus on here. The implicit argument is that qualitative data preservation and sharing lead to 'better' science through scientific transparency, scientific innovation, and scientific accuracy. Data archiving and reuse are understood to promote the kind of scientific transparency and quality control that, it is argued, is much needed in qualitative research: 'If we are to accept the label "scientist", then we should adopt the scientific model of opening up our data to scrutiny, and the testing of reliability and validity. The quality of social research is highly variable, and in the UK there are no quality control standards for qualitative studies' (Corti, 2000: 28). Scientific innovation, it is suggested, is made possible through data archiving and reuse because 'new questions' can be asked of 'old data', data can be approached in ways that were not originally addressed, and new substantive themes, findings, perspectives and interpretations can be generated allowing for new understandings of the data (Bornat, 2003, 2005, 2006; Corti and Thompson, 2004; Bishop, 2007). Scientific accuracy is understood as being facilitated through archival work because, compared to empirical field-based research, it can increase the 'reliability', 'validity', 'verifiability' and 'plausibility' of the research through 'distance', 'detachment' and 'objectivity' (Corti, 2000; Corti and Thompson, 2004; Fielding 2000). Fielding (2000: 21), for example, argues that secondary analysis can lay claim to greater plausibility since the analytic interests of the contemporary researchers will not have shaped, influenced and 'distorted' the data collected towards particular analytic purposes. He argues that: 'We generally regard data as more convincing the less the researcher has had to intervene directly in order to elicit them', and that a volunteered statement is generally regarded as more reliable than one in response to a direct question from the researcher. This view is echoed by Bornat (2005: 13) who argues that when using archived studies the 'interview texts become decontextualised so the secondary analyst is working with less data. This may allow a more 'objective approach'.

The arguments put forward implicitly draw on foundational assumptions. 'Data' are clearly conceptualised as entities that are separate from the researchers and the contexts that generated them. Naturalist and realist assumptions are made that the social world has a pure and 'natural state', that is independent of the researcher(s), and that can be best be understood and captured through minimal interference, disturbance and contamination from the researcher(s). The involvement and closeness that the 'primary' researchers have with 'their' 'data' is understood to undermine their ability to be objective and impartial. Through

their distance and detachment, 'secondary' researchers are seen as being more objective and, implicitly, better able to grasp 'the truth'. The 'scientific' case being made here is that 'distance from the data' is epistemologically advantageous.

This position, however, has been criticised for its foundational assumption that 'data' are 'out there', separate from the conditions and contexts in which they are generated, carriers of inherent meaning and 'truths', and commodities that can be unproblematically reused out of context (e.g. Hammersley, 1997; Mauthner et al., 1998). The response to this criticism has been to dismiss this epistemological challenge by suggesting that it can be 'fixed'. Specifically, it is suggested that secondary researchers can better contextualise the 'data' and understand their meaning by adding 'meta-data' and 'meta-documentation' about the research process, and as much of the 'original' context as possible (e.g. fieldnotes, diaries, memos, etc.), and where feasible by consulting with the original researchers. As Fielding writes, 'If the debate over epistemological issues relating to secondary analysis tells us anything, it is that it is very important that archived materials include as much information about the context of the original data collection as possible' (Fielding, 2000: 23, 2004).

Some suggest that adding context can go some ways towards recreating the experience of 'being there' and can therefore largely compensate for *not* 'being there' (e.g. Bishop, 2006, 2007; Corti, 2000; Fielding, 2000, 2004; Heaton, 2004). Others argue that, epistemologically, it may never be the same as 'being there' and that all secondary analysis is therefore likely to be more insecure than primary analysis because of 'the problem of decontextualisation of the interview' (Bornat, 2005: 30). Fielding, for example, suggests that if there are no other equivalent data available using archived data is useful. However, for data that still have contemporary relevance, he suggests that researchers should consider collecting primary data 'where they will enjoy the advantages of having the first bite at the cherry' (Fielding, 2004: 104). The implication is that while the context is not essential, the richer the overall data and context are, the better the resulting analyses will be. In other words, this is a pragmatic position in which second-hand is seen as second-best but as nevertheless worthwhile, good enough, and epistemologically possible and meaningful (Bishop, 2005; Corti, 2000; Fielding, 2000, 2004;). This position is supported by suggestions that these epistemological limitations are not unique to archival research. They are prevalent within other research practices that social scientists routinely engage in, such as team and collaborative research, and indeed they are generic to all research in that we can never fully grasp the entirety of the context of a study (Fielding, 2000, 2004).

The problems with this position are two-fold. First, the argument that 'distance from the data' is an epistemological *disadvantage* directly contradicts the position (often espoused by the same authors; see above) that 'distance from the data' is an epistemological *advantage* for secondary researchers using archived data. Hence, Bornat (2005: 13) maintains:

Even if the intimacy of the interview does not always survive it may have an effect on what is discussed and how the dynamics of an interview develop. Not “being there” reduces the interview to the text alone and changes the way the text can be analysed ... The interview texts become decontextualised so the secondary analyst is working with less data. This may allow a more “objective” approach but it excludes all the additional data on tone, emotion, expression and body language that an original interviewer can remember assuming that interviewer is also the data interpreter, unless these are included in the transcript and supporting notes.

Second, while this position is seen as ‘fixing’ the epistemological challenge, it fails to do so because the underlying ontological conception of ‘data’ remains a foundational one. ‘Data’ are still conceptualised in foundational terms as, for example, ‘respondent narratives’ that are separate from the researchers who generated them and from the original ‘context’ of the study. Whatever context is added, it is still understood as ontologically separate from ‘data’.

‘Postmodern’ approaches to qualitative data preservation and sharing

The modernist position acknowledges that archival (‘secondary’) and empirical (‘primary’) research are epistemologically distinct activities, but nevertheless argues that it is possible to put, at least some of, the original context back in. A more recent ‘postmodern’ position has shifted attention away from the ‘original’ or ‘primary’ contexts in which data are produced and highlighted the importance of the contemporary contexts in which archival materials are used and explored (e.g. Moore, 2007). This argument implies that archival and empirical research are epistemologically equivalent rather than distinct activities. The epistemological challenge is effectively dismantled by doing away with the epistemological distinction between field-based and archival qualitative research, and associated conceptual and terminological distinctions between ‘old’ and ‘new’, or ‘primary’ and ‘secondary’ data, analysis and researchers. It is suggested that, like historians, literary critics and other disciplines that use archived materials, social scientists ‘recontextualise’ the data through the reflexive production of data in the *contemporary* research project (Moore, 2007). Data are constructed anew in the process of a new project undertaken by subsequent investigators (Bornat, 2003). Researchers are no longer seen as using ‘past’ or ‘old’ data but rather as appropriating data in a contemporary context, and constructing new relationships between researchers and the data in the context of a *current* project (Bornat, 2003, 2005, 2006; Andrews, 2008; Mason, 2007; Moore, 2007). Moore (2007), for example, argues that contemporary researchers can generate new meanings, understandings and interpretations of data by bringing new theoretical lenses and questions to these materials and ‘recontextualising data’ in a contemporary context. She argues that reuse is more productively understood as a process of recontextualising data and that ‘attending to the reflexive production of data in the contemporary research project may offer more hopeful possibility for reuse’. Andrews (2008: 92)

similarly suggests that ‘there is never a definitive interpretation of data’ and that ‘subsequent readings of material we, or others, have gathered invariably bring with them a new layer of understanding. But no interpretation is ever final; our current framework is itself one which will change over time’ (2008: 90).

This position, we suggest, is as problematic as the modernist argument because it too remains located within an implicit foundational tradition. These authors usefully draw our attention to the contemporary contexts of the ‘secondary’ researchers, but in doing so they leave behind the ‘primary’ researchers and revert to foundational notions of ‘data’ as narratives that are ontologically separate from the ‘original’ researchers and contexts that gave rise to them. Moore (2007), for example, firmly opposes ‘the notion of pre-existing data’ which are ‘collected, gathered or found’, and equally firmly aligns herself to the view that ‘data ... is created and co-produced in the research process, between researcher and respondent’. However, her discussion suggests that in the processes of recontextualisation carried out by ‘secondary’ researchers, the ‘primary’ context of what she terms the ‘data’ is not considered. In other words, contemporary researchers are recontextualising decontextualised data. The *implied* ontological unit of analysis – for Moore does not define what she means by ‘data’ – that the contemporary researchers are working with remains a foundational notion of decontextualised ‘pre-existing data’ – data taken out of the original and archival contexts through which they have been created, and put into a contemporary context:

A new research project provides a new context for the creation and emergence of ‘data’, particularly through the contemporary production of the relationship between researcher and data ... secondary analysis is not so much the analysis of pre-existing data; rather secondary analysis involves the process of recontextualising, and reconstructing, data (Moore, 2007).

Attempts to address the epistemological challenge of data preservation and sharing are all variations on an ‘add context and stir’ approach. Whether we contextualise data within the original contexts of ‘data production’, the archival contexts of ‘data preservation’, and/or the contemporary contexts of ‘data reuse’, ‘data’ are still conceptualised in implicit foundational terms in which ‘context’ remains ontologically separate from, rather than constitutive of, ‘data’. Both ‘modernist’ and ‘postmodern’ arguments, we suggest, are still working with foundational ontological conceptions of ‘data’. They therefore fail to address the foundational norms that are implicit within the contemporary qualitative data preservation and sharing movement and the ambivalence that these norms are creating amongst qualitative researchers.

Conclusions

The archiving of qualitative research data is increasingly becoming a matter of national policy and practice in the United Kingdom, United States, Canada and Europe (Corti, 2000; Cribier, 2005; Fink, 2000). In Australia, the recently established Australian Qualitative Archive (AQuA) is promoting enhanced data

preservation, access, and secondary analysis for the international qualitative research community (Broom et al, 2009). In New Zealand the issue of saving and sharing quantitative and qualitative research data is being debated, but as yet, there are no established policies or practices (Davis, 2004). For over a decade, qualitative social scientists have been expressing concerns over this emerging data preservation and sharing movement. The tendency within academic debates, however, has been either to attempt to 'fix' these concerns or to dismiss them, largely, we suggest because they have been seen as opposing and threatening the larger project of 'qualitative data preservation and sharing'. We argue, however, that these concerns are not challenging the principle of data preservation and sharing, but rather the philosophical or scientific terms on which qualitative researchers are being asked to participate in these practices. Specifically, we contend that it is the implicit foundational assumptions and norms underpinning qualitative data preservation and sharing activities that are proving problematic for many qualitative researchers.

We want to reframe the debate as one that is not about whether, as qualitative researchers, we embrace or oppose the data preservation and sharing movement, but rather *on what philosophical and scientific terms we engage with it*. To date, the qualitative data preservation and sharing project has been implicitly informed by a foundational model of qualitative research. This reflects its genealogy and emergence out of a much broader and international movement that is overwhelmingly concerned with quantitative data. Thus, a particular philosophical approach to research is being defined as normative and imposed onto the qualitative research community. Moreover, researchers are expected to comply with this norm regardless of their own philosophical orientation and approach. Data preservation and sharing discourses and policies, as currently conceptualised, are therefore creating a set of normative and dominant expectations about the philosophical orientation of qualitative research.

At the level of qualitative research practices, however, researchers are engaging with these normative expectations in different ways. For researchers working in applied settings (e.g. research designed to directly address practical issues), the conditions and contexts of their research and funding arrangements may leave them with little scope to specify the philosophical terms of their research. Furthermore, for foundationally-inclined qualitative researchers, taking part in data preservation and sharing activities on foundational terms will present few difficulties. Researchers drawing on non- or post-foundational philosophical perspectives, however, may experience tensions. For some scholars, these tensions may not constitute significant obstacles to their involvement in data archiving and reuse. They may recognise the epistemological tensions and limitations inherent in this approach, but are pragmatic and argue that these practices can nevertheless lead to research that is 'good enough'. Others are developing approaches and practices drawing on post-foundational theories of knowledge, though they themselves may not characterise their approaches in this way (e.g. Akerstrom et al.,- 2004; Gillies and Edwards, 2005; McLeod and Thomson, 2009; Savage, 2005). What is more problematic is the engagement

in practices that are explicitly framed as ‘post-foundational’ or ‘reflexive’, while implicitly drawing on foundational assumptions without recognising or acknowledging the contradictions involved.

This problem is not unique to the area of qualitative data archiving and reuse. It illustrates much wider tensions within the field of qualitative research between its post-foundational aspirations and its use of practices that continue to be informed by an implicit foundational tradition (see Mauthner and Doucet, 2008). For these reasons we view the data preservation and sharing debate, and the concerns it raises, as highly instructive. This is because it points to broader issues within the field of qualitative research concerning the need for qualitative researchers to be more ‘reflexive’ not only about the ontological and epistemological assumptions that inform their own research, but more generally about the assumptions that implicitly inform the practices they engage in (Doucet and Mauthner, 2008; Mauthner, 2009; Mauthner and Doucet, 1998, 2003, 2008). This latter form of reflexivity, what Bourdieu terms ‘epistemic reflexivity’, goes beyond the individual scholar and takes as its focus of analysis ‘scientific practice’, modes of knowledge production, and the ‘epistemological unconscious’ underpinning these (Bourdieu and Wacquant, 1992). In this article, we have attempted to engage in this form of epistemic reflexivity, through critical and reflexive analysis of the emerging data preservation and sharing movement within qualitative research. This analysis, we argue, suggests that the future success of qualitative data preservation and sharing in Australia, the UK and elsewhere is likely to depend on our ability to shed its foundational norms, and embrace ‘epistemic pluralism’ and the diversity of philosophical traditions within qualitative research thereby allowing qualitative researchers to participate in data preservation and sharing activities on their own philosophical terms.

References

- Akerstrom, M., Jacobsson, K. & Wasterfors, D. (2004) ‘Reanalysis of previously collected material’. In: Seale, C., Gobo, G. and Gubrium, J. F. (eds) *Qualitative Research Practice*, London: Sage.
- Alderson, P. (1998) ‘Confidentiality and consent in qualitative research’, *British Sociology Association Network*, No. 69, 6-7.
- Andrews, M. (2008) ‘Never the last word: revisiting data’. In: Andrews, M., Squire, C. and Tamboukou, M. (eds) *Doing Narrative Research*, London: Sage.
- Bell, C. (2004) ‘Doing sociological research: the genre of ‘owning up’’, *International Journal of Social Research Methodology*, 7(1), 29-33.
- Bishop, L. (2005) Is secondary analysis second best? A case study of reusing qualitative data. Paper Presented at the CRESC Methods Workshop: Reusing Qualitative Data, University of Manchester, September.
- Bishop, L. (2006) ‘A proposal for archiving context for secondary analysis’, *Methodological Innovations Online*, 1(2): Available at: <http://sirius.soc.plymouth.ac.uk/~andyp/viewarticle.php?id=26>. Accessed 11 July 2009.

- Bishop, L. (2007) 'A reflexive account of reusing qualitative data: Beyond primary/secondary dualism', *Sociological Research Online*, 12(3), Available at: <http://www.socresonline.org.uk/home.html>. Accessed 11 July 2009.
- Boddy, M. (2001) *Data Policy and Data Archiving: Report on Consultation for the ESRC Research Resources Board*, Bristol: University of Bristol.
- Bornat, J. (2003) 'A second take: Revisiting interviews with a different purpose', *Oral History*, 31(1), 47–53.
- Bornat, J. (2005) 'Recycling the evidence: Different approaches to the reanalysis of gerontological data', *Forum: Qualitative Social Research*, 6 (1), Art. 42.
- Bornat, J. (2006) 'Secondary analysis of one's own and others' data'. Paper given at ESRC seminar Practice and Ethics in Qualitative Longitudinal Research, University of Leeds, January.
- Bourdieu, P. & Wacquant, L. (1992) *Invitation to Reflexive Sociology*. Cambridge: Polity.
- Broom, A, Cheshire, L. & Emmison, M. (2009) 'Qualitative researchers' understandings of their practice and the implications for data archiving and sharing', *Sociology*, 43(6), 1163-1180.
- Corti, L. (2000) 'Progress and problems of preserving and providing access to qualitative data for social research', *Forum: Qualitative Social Research*, 1(3), Art. 2.
- Corti, L. & Thompson, P. (2004) 'Secondary analysis of archive data'. In: Seale, C., Gobo, G. and Gubrium, J. F. (eds) *Qualitative Research Practice*, London: Sage.
- Cribier, F. (2005) 'The value of qualitative data and their archiving: The French situation', *Forum: Qualitative Social Research*, 6 (2), Art. 35.
- Davis, P. (2004) 'Saving and sharing research data: Issues of policy and practice', *Social Policy Journal of New Zealand*, 21, 211-215.
- Doucet, A. & Mauthner, N. (2008) 'What can be known and how? Narrated subjects and the listening guide', *Qualitative Research*, 8 (3), 399-409.
- ESRC (2000) 'Dataset Policy'. Swindon, UK: Economic and Social Research Council. **Error! Hyperlink reference not valid..** Available at: http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Images/DataPolicy2000_tcm6-12051.pdf. Accessed: 15 August 2009.
- ESRC (2008) Memorandum from the Economic and Social Research Council in Response to the Ministry of Justice Consultation on the Use and Sharing of Personal Information in the Public and Private Sectors. Available at: <http://www.justice.gov.uk/docs/Economic-and-Social-Research-Council-Data-Sharing-Response.pdf>. Accessed: 15 August 2009.
- Fielding, N. (2000) 'The shared fate of two innovations in qualitative methodology: The relationship of qualitative software and secondary analysis of archived qualitative data', *Forum: Qualitative Social Research*, 1(3) Art. 22.

- Fielding, N. (2001) 'Automating the ineffable: Getting the maximum from archived data' Paper presented at Celebrating Classic Sociology, Essex, July.
- Fielding, N. (2004) 'Getting the most from archived qualitative data: Epistemological, practical and professional obstacles', *International Journal of Social Research Methodology*, 7(1), 97-104.
- Fink, A. S. (2000) 'The role of the researcher in the qualitative research process. A potential barrier to archiving qualitative data' *Forum: Qualitative Social Research*, 1(3), Art. 4.
- Gillies, V. & Edwards, E. (2005) 'Secondary analysis in exploring family and social change: Addressing the issue of context', *Forum: Qualitative Social Research*, 6 (1), Art. 44.
- Griffin, C. (1997) 'Archiving qualitative datasets: Social research and the politics of control', *Radical Statistics*, 66, 44-47.
- Hammersley, M. (1997) 'Qualitative data archiving: Some reflections on its prospects and problems', *Sociology*, 31(1), 131-42.
- Heaton, J. (2004) *Reworking Qualitative Data*, London: Sage.
- Healy, S. (2004) 'A 'post-foundational' interpretation of risk: Risk as 'performance'', *Journal of Risk Research*, 7(3), 277-296.
- Henderson, S., Holland, J., & Thomson, R. (2006) 'Making the long view: Perspectives on context from a qualitative longitudinal (QL) study', *Methodological Innovations Online*, 1(2), Available at: http://erdt.plymouth.ac.uk/mionline/public_html/viewarticle.php?id=29&layout=html. Accessed 6 July 2009.
- Henwood, K. & Lang, I. (2005) "Qualitative Social Science in the UK: A Reflexive Commentary on the 'State of the Art'." *Forum: Qualitative Social Research*, 6(3), Art. 48.
- HMSO (2004) Science and Innovation Investment Framework 2004-2014. London: Her Majesty's Stationary Office. Available at: http://www.hm-treasury.gov.uk/spending_sr04_science.htm. Accessed 12 March 2009.
- Lagacé, R. O. (1978) 'The HRAF probability sample: Retrospect and prospect', *Cross-cultural Research*, 14, 211-229.
- Mason, J. (2007) "Re-using' qualitative data: On the merits of an investigative epistemology", *Sociological Research Online*, 12 (3). Available at: <http://www.socresonline.org.uk/home.html>. Accessed 7 June 2009.
- Mauthner, N. (2009) 'Breaking the mould: Feminist research collaborations and their scientific, ethical and political counter-practices'. Paper presented at Collaboration and Duration: A Celebration of the Work and Research Practices of Janet Holland, London South Bank University, December.
- Mauthner, N., Parry, O. & Backett-Milburn, K. (1998) 'The data are out there, or are they? Implications for archiving and revisiting qualitative data', *Sociology*, 32(4), 733-45.

- Mauthner, N. & Doucet, A. (1998) 'Reflections on a voice-centred relational method: Analysing maternal and domestic voices'. In: Ribbens, J. and Edwards (eds), *Feminist Dilemmas in Qualitative Research: Public Knowledge and Private Lives*, London: Sage.
- Mauthner, N. & Doucet, A. (2003) 'Reflexive accounts and accounts of reflexivity in qualitative data analysis', *Sociology*, 37(3), 413-431.
- Mauthner, N. & Doucet, A. (2008) "Knowledge once divided can be hard to put together again": an epistemological critique of collaborative and team-based research practices', *Sociology*, 42(5), 971-985.
- Macleod, J. & Thomson, R. (2009) *Researching Social Change: Qualitative Approaches*, London: Sage.
- Mass Observation (2009) A brief history. Available at: http://www.massobs.org.uk/a_brief_history.html. Accessed 1 August 2009.
- Moore, N. (2007) '(Re)using Qualitative Data', *Sociological Research Online* 12(3). Available at: <http://www.socresonline.org.uk/home.html> [Accessed 15 March 2009].
- OECD (2007) *Principles and Guidelines for Access to Research Data from Public Funding*. Paris: OECD. Available at: <http://www.oecd.org/dataoecd/9/61/38500813.pdf>. Accessed 16 March 2009.
- OSI (2007) 'Developing the UK's e-infrastructure for Science and Innovation'. Office for Science and Innovation – OSI e-infrastructure Working Group. Available at: <http://www.nesc.ac.uk/documents/OSI/report.pdf>. Accessed 16 March 2009.
- Parry, O. & Mauthner, N. (2004) 'Whose data are they anyway? Practical, legal and ethical issues in archiving qualitative research data', *Sociology*, 38(1), 139-152.
- Parry, O. & Mauthner, N. (2005) 'Back to basics: who re-uses qualitative data and why?', *Sociology*, 39(2), 337-342.
- RIN (2008) *Data Sharing Review: Consultation on the Use and Sharing of Personal Information in the Public and Private Sectors*. Response from the Research Information Network, February 2008. Available at: <http://www.justice.gov.uk/docs/Research-Information-Network-Data-Sharing-Response.pdf>. Accessed 19 March 2009.
- Rouse, J. (1996a) Beyond epistemic sovereignty. In: Galison, P. and Stump, D. J. (eds) *The Disunity of Science: Boundaries, Contexts and Power*, Stanford: Stanford University Press.
- Rouse, J. (1996b) *Engaging Science: How to Understand its Practices Philosophically*, Ithaca: Cornell University Press.
- Savage, M. (2005) Revisiting Classic Qualitative Studies. *Forum: Qualitative Social Research*, 6(1), Art. 3.
- SHERPA (Securing a Hybrid Environment for Research Preservation and Access) (2009a) Research funders' open access policies. Available at: <http://www.sherpa.ac.uk/juliet/>. Accessed 5 July 2009.

- SHERPA (2009b) Research funders' open access policies. Medical Research Council. Available at: <http://www.sherpa.ac.uk/juliet/index.php?fPersistentID=8>. Accessed 5 July 2009.
- Silva E. B. (2007) 'What's [Yet] to be seen? Re-using qualitative data' *Sociological Research Online*, 12(3), Art. 4.
- SSHRC (Social Sciences and Humanities Research Council) (2009) Research Data Archiving Policy. Available at: http://www.sshrc-crsh.gc.ca/site/about-crsh/policy-politiques/statements-enonces/edata-donnees_electroniques-eng.aspx. Accessed 1 August 2009.
- Thompson, P. (2000) 'Re-using qualitative research data: A personal account', *Forum: Qualitative Social Research*, 1(3), Art. 27.
- UKRDS (2008) *The UK Research Data Service Feasibility Study Report and Recommendations*. Available at: <http://www.ukrds.ac.uk/resources/>. Accessed 25 March 2009.
- UKRDS (2009) *The Data Imperative: Managing the UK's Research Data for Future Use*. Available at: <http://www.ukrds.ac.uk/resources/>. Accessed 25 March 2009.
- Van den Berg, H. (2005) 'Reanalyzing qualitative interviews from different angles: The risk of decontextualization and other problems of sharing qualitative data', *Forum: Qualitative Social Research*, 6(1), Art. 30.