**Qualitative Secondary Analysis in Practice:
An extended guide**
(with reference to concepts, contexts and knowledge claims)

A Timescapes Working Paper

Sarah Irwin and Mandy Winterton

2011

Timescapes Working Paper Series No.7

UNIVERSITY OF LEEDS   LONDON SOUTH BANK UNIVERSITY   CARDIFF UNIVERSITY PRIFYSGOL CAERDYDD   THE UNIVERSITY OF EDINBURGH   The Open University   UK Data Archive

0

Qualitative Secondary Analysis in Practice:
An extended guide
(with reference to concepts, contexts and knowledge claims)

A Timescapes Working Paper

Sarah Irwin and Mandy Winterton

CONTENTS

1. Introduction

The aim of this paper is to share some of the ideas which guided and evolved during the Timescapes Secondary Analysis (SA) project, and some of the lessons we learned on the way. In it we hope to offer an extended and practical guide to those undertaking secondary analysis of qualitative data, but we also hope it will be of interest to other qualitative data analysts (primary as well as secondary). In our discussion we also describe our approaches to working across data sets, and working with longitudinal data. This extended guide evolved in the writing of a brief Timescapes Series Methods Guide on Secondary Analysis. In seeking to provide a resource for secondary analysts there seemed value in offering a more expansive and more fully exampled account of lessons we have learned which we hope may be of interest to prospective secondary analysts. It may be, too, that in reflecting on issues confronting secondary analysts, primary researchers depositing the data for possible re-use will think in some detail about the kinds of metadata and guidance they might make available. Whilst we have a range of secondary analysis project outputs and publications describing our strategies and analyses these are by their nature diverse. We hope that bringing together some of our approaches in one document might offer an informative account of a secondary analysis undertaking 'in practice'. (This paper is very lightly referenced and we refer the reader to our other work for a review and orientation to available literature on secondary analysis. See especially Irwin and Winterton 2011a).

ESRC Timescapes is a 5 year long programme of work centring on 7 primary qualitative longitudinal research projects, run by project teams in 5 different universities.[i] The projects were independently conceived, and many were in place before Timescapes commenced. They were brought together and cohere due to a range of shared substantive interests, in biography, life course, and life course transitions, familial relationships and inter-generational dynamics. Crucially they are all qualitative longitudinal projects and have concerns with social processes as temporal and dynamic, and a range of methods oriented to exploring time in a range of dimensions, including experiential, biographical, and historical. The projects, then, were working with some common themes. They also committed to a shared programme of activities alongside the independent project objectives. This shared programme included, crucially, depositing data in the Timescapes Archive as well as developing supporting activities, such as devising ethical protocols and guidance on best practice; and undertaking various secondary analysis activities. For example, the secondary analysis team ran a two-day residential meeting with participation of team members from all projects. Also, the Timescapes programme as a whole has undertaken some common ventures (e.g. a collection of articles from across Timescapes projects, exploring a common theme, Edwards and Irwin 2010). Additionally individual projects have coordinated paired meetings to explore inter-project secondary analysis possibilities. As a separate and distinct undertaking Timescapes has also included a dedicated Secondary Analysis Project. This was run within a two year period, across the final stages of the overall project, by the current authors. This meant that the Secondary Analysis Project was running concurrently with primary projects and was dependent on them for supplying data directly. The timing of the SA project meant that some projects were still generating data and most were still in the stages of preparing data for archiving, as well as undertaking their analysis and writing. Some of our secondary analysis experiences and practices would not, then, directly mimic what many secondary users will experience. We seek to offer here commentary on what we believe to be generally useful lessons. We do so through a grounded account which builds on examples of our research in practice.

Our discussion will steer a course between the detailed 'doing' of things and the relative abstraction of academic discussion of concepts and methods. We seek a middle way in a paper which is primarily practically oriented and uses a series of briefly sketched examples to illustrate some of our

philosophy in practice and our varied analytic strategies. The paper is organised as follows, and roughly relates to how one might go about doing a secondary analysis. We commence with a brief consideration of getting started with data re-use, getting a handle on a new data set, and the nature of data within it. We reflect on issues of sample structure and issues of context and the need to engage with the conditions of data production. We consider how analysts might set about building an in-depth understanding of data sets where these may be too large for an analysis of all available data. We then explore different analytic strategies we have pursued. Here we consider diverse possibilities for organising available data, and different ways of thinking about links between the particular and the general. Through our examples we reflect on a range of issues including how we built analyses across independently conceived data sets and developed strategies for analysing qualitative longitudinal data.[ii]

2. Orienting to the archived research project

2.1. Archives, ethics and data re-use

Different archives, research projects and originating teams will have different protocols to be observed in undertaking data re-use and secondary analysis. Diverse forms of data, conditions of confidentiality and levels of access to data will render data differently amenable to re-use by others. The Timescapes Guides on Archiving, and on Ethics, will provide discussion of these matters. Here we assume prospective data re-users have secured access to data, and that originating researchers' (and participants') permissions are in place, ethical protocols and re-use procedures are agreed. This extended guide focuses on issues the analyst confronts having accessed the data.

2.2. The project research design

Secondary users need to understand the purpose and contours of the project they draw on. How, why and by whom was the research done? What was the social and historical context of the research? What was the theoretical context of the research? Some of the relevant information may be stored within the archive and/or publically available elsewhere (in published journal articles, end of award reports, project websites, working papers etc).

There are a range of kinds of project metadata which can be documented within Qualitative Data Archives to provide further orientation to the data by secondary users. We summarise some of these briefly elsewhere (Irwin and Winterton 2011a) and both the Timescapes Archive and ESDS Qualidata offer minimum standards and guidance about useful metadata which will facilitate re-use. The Timescapes Secondary Analysis Project was usually working with data that had not yet been fully prepared for the Archive, because of concurrent timing. There are debates about the extent of metadata primary researchers might supply, lest this pre-judge or frame what subsequent analysts do with the data. As secondary analysts our view is that more, rather than less, contextual information is helpful. Standard descriptors will include notes on the project outputs, its research design, the sampling decisions and final sample structure (including 'hidden', e.g. self selection, specificities). They would also include an overview of what data is being made available for re-users, and relevant contextual data (Bishop 2006). A descriptive profile of each participant is not 'standard' practice but very helpful for orienting re-users to the sample profile, and individual cases, albeit at a potentially surface level.

2.3. Orienting to the project data

Secondary analysts need understand the array of data available to them within archived projects especially where (as was the case with Timescapes projects) primary researchers have worked with a wide variety of methods for generating data. For example, projects generated interview data (including semi-structured interviews, oral history interviews, family based interviews), and all had a longitudinal strategy. Many collected different kinds of visual data, including having young participants complete drawing activities to represent aspects of their lives, and making use of these as elicitation devices within interview, along with them taking photographs and video footage. One project collected both oral history and contemporary (longitudinal) diary data. How do the different kinds of data within single projects articulate with one another? What are suitable strategies for analysing such data? What is the potential for working across similar types of data, produced in different contexts? Are there substantive as well as methodological lessons to be drawn? When tackling qualitative longitudinal (QL) research, secondary analysts need an understanding of the rationale underpinning the longitudinal design of projects. For example, the different periodicity of interview waves will be closely bound to project objectives and impact on the kinds of analytic insight available. Analysts will firstly need to orient themselves to the structure and content of their chosen project(s). Metadata provided by the originating project team, and other archived resources, will aid in this task (see Irwin and Winterton 2011a, and Timescapes Archiving Methods Guide). They then need to develop a detailed understanding of the project data. A superficial understanding, or cherry picking of data or cases out of context, might allow (at best) a descriptive and partial account, and risks being misleading. Developing a detailed understanding of data sets requires getting to grips with:

- The structure of the project data. If this is a multi-method study, what types of data are available and how do they articulate with one another? What was the originating rationale of the study, and method(s)?
- The structure of the sample. Analysts need understand the sampling logic, the achieved sample structure and be aware of how the sample speaks to their own research questions.

In the next sections we consider how data is embedded in the conditions of its production, and then explore how to build a grounded understanding of the data. This ordering does not strictly follow our own trajectory. Building an understanding of the contexts of data production was ongoing and a prelude to, as well as an outcome of, more detailed analysis.

3. Understanding contexts of data production

Recognising the nature of data as contextually produced is important to effective qualitative research. For us, working across projects as well as with longitudinal data, highlighted the contextual situatedness of data. The significance of research design, methods, interviewers' interests as well as the impact of specific contextual factors are particularly visible. Insight will be gained from an adequate early orientation to the project being explored, but much of the detailed manifestation of context will only be apparent through reading and interpretation of the available data. Different dimensions of context include the proximate contexts in which research participants move, research project contexts, and researcher-participant interactions to name just a few. Indeed the catch-all idea of 'context' can engender difficulties and lack of clarity. There are many different issues which fall under this heading. These are matters which have been discussed extensively, and especially in debates about the possibility of effective secondary analysis of qualitative data (see Irwin and Winterton 2011a for a review). In that paper we suggested that context is often discussed in terms of the immediate contexts of data production and in terms of wide historical, theoretical contexts in which research questions are framed. We argue there that less attention has been given to a set of

middle range issues about the conditions in which, and methods through which, data is produced and shaped. We illustrate some of the issues through concrete examples.

Within Timescapes a number of questions were asked in common across the projects. We used data from answers to such questions as an early 'way in' to project data. This exercise revealed the very embedded nature of data within project contexts, and generated some specific questions about creating meaningful comparisons across projects. We want to comment briefly on the issue of how questions are worded, framed and positioned within an interview, or how tasks set for participants are managed (e.g. visual data and elicitation strategies). These are issues relating to 'immediate' conditions of data production. Such matters shape how research participants are oriented, and influence the kinds of answers and accounts that they give. The issues are well rehearsed, indeed part of the raison d'etre of qualitative research.[iii] Qualitative researchers seek to access diverse meanings and experiences which formal similarities in questioning are more likely to miss than to capture. In qualitative interviews, for example, the interviewer will seek to understand 'where the participant is coming from', will use their wits to understand meaning in context, including the possibly varying relevance of questions to research participants. It may seem odd, then, for us to pause with the issue of how questions shape people's responses. Nevertheless we pause briefly, since familiarity with a problem does not necessarily equate to the critical reflection we urge. One issue is that the contextually embedded nature of data is not always wholly visible until it is brought into comparison with data in other contexts. It remains incumbent on us to understand precisely how data is shaped. For secondary analysts seeking to bring evidence into conversation across data sets (which may include their own primary data sets for example) this will be especially important. To illustrate, even if the same question is posed to participants re-users cannot assume this makes responses readily comparable. Minor changes in wording may orient participants in different ways, and it is important to note this when drawing on such data. If question wording is directly replicated, secondary users need to investigate further the context in which such utterances take place. Has anything relevant to the topic been raised earlier (perhaps spontaneously) in the interview, or in previous interviews or research encounters? Such concerns obviously caution against any simplistic thematic analysis. The amount of time it takes to read and interpret data with a grounded knowledge of how such data was produced, and situated within the research encounter, should not be under-estimated.

A broader issue is that project designs provide less 'immediate', but probably even more fundamental, conditions of data production. Researchers' disciplinary background and concerns, their samples, research designs, techniques for orienting participants to the project, and the research methods they use all contribute to shaping emergent data. Therefore secondary analysts need to be sensitive to the specificity of sample structure, and reflect on both overt and hidden drivers of the nature of completed samples. They need to be sensitive to how participants are oriented due to the design of the research, and what facet of their experience is then engaged (and they reveal) in their interactions with the researcher (Mason 2002; Irwin 2008). The disciplinary interests of the original researchers will influence the research questions, and the fieldwork questions, and the ways these are followed up in fieldwork and in interviews: the silences that are heard and followed up, and the silences that are ignored. Furthermore, whether participants are involved as individuals or as a constituent part of a family or friendship group may influence how they respond to questions regardless of whether they were interviewed in the presence of others. Diverse methods obviously also shape data. Precisely how they do so is not always readily apparent. We illustrate this point by examining people's reflections on belonging (or not) to a generation. Within Timescapes all projects committed to asking a few questions in common, relating to perceptions of biographical change and turning points, to perceptions of historical change, and to perceptions of belonging to a generation. Our preliminary analysis of responses to the latter line of questioning suggested quite nebulous notions of generational membership (amongst mid-life, and teenage project participants for

example). However, it also appeared that *some* members of The Oldest Generation[iv] sample identified as members of a wartime generation, and referenced the war as a point of commonality. Is this a 'classic generation' (cf Mannheim) linked to shared early adult experiences and orientations forged through the Second World War? Or is it at least partly a product of the research method? The Oldest Generation study was centred on life history interviews (although other kinds of data were also collected). It seems possible that such interviews, each focusing on a life lived through most of the 20th century and thus in their historical context, might be more likely to render a particular (affirmative) response to a question on generational membership than a similar question asked in the context of studies more oriented to current experiences. Our reading here is speculative, and does not undermine the interesting evidence about perceptions of generation amongst participants. However, the general point is that we need to always recognise that participants' accounts make sense within the *particular* narrative and context in which data is generated.

In this section we have discussed some ways in which data is embedded in the contexts of its production. This can mean many diverse things, from understanding the day to day practices of researchers and the circumstances they witnessed, and the conduct of the interview, to broad questions relating to the historical and epistemological contexts in which research is conducted. It is our sense that there has been less discussion, within the literature on secondary analysis, of issues relating to the project contexts and methodologies through which data is created. We have therefore sought to illustrate some of these issues in practice and the importance of critical reflection, the need for which is very evident when we seek to work across data sets. However, rather than paint ourselves into a corner with such matters (wherein the embeddedness of data precludes meaningful cross-project comparisons of data) as qualitative data analysts we are seeking to recruit data *as evidence* towards addressing our research questions (cf Hammersley 2009). With Hammersley (2009), and as we have argued elsewhere (Irwin and Winterton 2011a), we will not find the answers *within* data sets, but in theorisation of how the data provides and links with other evidence in addressing specific research puzzles or questions. Before considering examples of some of our own analytic strategies we describe how, as re-users of data, we got to grips with complex data sets not of our own making.

4. Building a grounded understanding of the data, and situating cases

4.1. Introduction

Secondary analysts will have different purposes and the ways they read and analyse archived data will relate to diverse objectives. Geiger and colleagues (2010) draw a distinction between focusing on the form or content of archived data, as well as discussing ways in which archives themselves order material, and shape possibilities for knowledge production. Our own focus, as we have made clear, lies with the content of the data that has been made available to us. As we have also made clear we see data as 'formed' (that is made, and also structured in particular ways).

We note that the data we analysed was not 'archived' as such, rather we were provided with copies of completed transcripts and other extant data directly from the primary research projects. Even where detailed data about the research project is provided, secondary analysts can expect to confront complex and often very large volumes of data in a qualitative data archive. It may not be realistic for data re-users to read, let alone analyse, all the data produced during the course of the original research. Further, secondary analysts will be distanced from the data and need particularly to develop their familiarity and understanding. As secondary analysts we need strategies for making inroads

into the complexity of data sets and for developing a thorough understanding of the data within. We might follow a logic which is primarily deductive, primarily inductive, or a combination of both.

4.2. Deductive approaches

Analysts might sample cases for preliminary reading through making use of some summary indices. For example all Timescapes projects recorded some standardised socio-demographic data on their participants, allowing scope for subsampling according to given criteria. However, decisions here may have quite fundamental implications in shaping how a data set is read. For example, we might consider the criteria a project itself followed in respect of sampling e.g. people with particular kinds of attribute; households with different divisions of gendered labour; people with different kinds of family background. We might then choose to select out a range of cases which we take to broadly represent this diversity. However, we might find that any such representativeness or typicality resides only at the superficial level of summary indices. For us familiarising ourselves with data sets this deductive strategy was only an entry point. We then 'read outwards' across data.

Another deductive strategy might follow a logic of tackling some specific theoretically interesting aspect of population diversity, for example seeking out people in particular contexts or with particular experiences, possibly with a view to gaining general explanatory purpose from specific, situated, examples (cf. Ward Schofield 1993). Again the analyst would need a sufficient and broad knowledge of the data to situate, and analyse, specificity.

There may be times when a deductive strategy can be used for identifying theoretically interesting case studies to analyse in more depth. For example, participants in the Timescapes' Young Lives and Times project completed a questionnaire which had been used in a survey of the same age group across the authority area. Participants from the qualitative study could be identified for in-depth analysis according to how were situated with respect to wider population heterogeneity. In this example such heterogeneity was in terms not just of socio-demographic characteristics but also attitudes and expectations (Irwin 2009). It was possible to pinpoint experiences with reference to heterogeneity (e.g. of circumstances and H.E. expectations) across the wider population. This is a broadly deductive strategy for guiding case selection for detailed qualitative analysis. The link to external evidence here is in respect of understanding how cases sit with respect to population diversity and may be deemed then to offer some insight into experiences as these relate to such diversity, where it is mapped by formal (e.g. socio-demographic, or attitudinal) indices. Again we would caution that qualitative cases should only be targeted for detailed analysis in the context of a good knowledge of the qualitative data set as a whole.

4.3. Other links across qualitative and quantitative data sets

Other points of connection with extant large data sets were included in the Young Lives and Times study. For example within the UK Household Longitudinal Study young survey respondents are provided with a series of structured questions on how they feel about different aspects of their lives. Such aspects included domains (family life; friendship; school) covered in depth in YLT qualitative interviews. Young Lives participants were also requested to complete the relevant structured closed questions as completed by respondents to the (UKHLS) youth survey. This is an example of 'forcing' a point of connection between two very different kinds of data set which might, then, allow for a more expansive kind of analysis. For example, we could reflect on how individual Young Lives participants are situated with respect to a bigger picture on given summary indices, *and* gain insight into the qualitative evidence corrolary of the summary indices. The strategies we report are, in effect,

post hoc, since Timescapes and quantitative longitudinal studies have proceeded on separate tracks. Logically large survey based panel study designs would also include qualitative studies. In the meantime secondary users might usefully think about if, and how, they could build links with larger data sets. For example, Timescapes covers a range of themes which correspond with issues being explored as part of the UKHLS.

4.4.   Inductive approaches

Having discussed more deductive approaches we now consider how a more inductive approach to sampling cases within a study would seek to *build* a picture, starting with the particulars of data and cases, and then reaching outwards across the data set. An analyst might explore values for example, or expressions of identity and commitment, and extend a reading outwards from particular cases to gain a sense of how they 'fit' the bigger picture. Such an approach is iterative. The first readings build a picture of diversity. In a very large qualitative data set it may be necessary to sample, and analysts might seek to choose as wide a range of cases as possible (perhaps with reference to metadata if appropriate, and with the suitable cautions about superficiality, or perhaps with reference to reading selected material within transcripts). This then becomes a basis for undertaking detailed readings and whole case analyses. An example of such a strategy lies in Mandy's readings of young people's short 'essays' on their imagined futures within the 'Young Lives and Times', and 'Siblings as Friends' studies[v]. She built a preliminary understanding of diversity with respect to expectations, and its possible correlates (Winterton and Irwin 2011). Further evidence on socio-demographic diversity and extensive readings of qualitative evidence confirmed our initial hunch this would be a productive line of enquiry. On that basis we developed an in-depth, case based, analysis.

We tended to work by reading across single data sets as widely as we could. However, analysts may need to sample within data sets, for example, if the volume of data makes a comprehensive reading difficult. Having done so they could then read 'outwards' to test out their evolving ideas, see if they obtain in different contexts and so on. It is important to secure a fully grounded understanding of the nature and content of the data set(s) being re-used. We could accept that qualitative data is in any event not representative and ask if and why it matters which participants, or what subsample, we pick out for more detailed analysis? (cf. Geiger, Moore and Savage 2010). However, whether they are working with a comparative strategy or one based on developing theoretical cases, qualitative researchers commonly seek an understanding of how people are situated within study samples in order to grasp the contexts underpinning diversity, and hence insights into conditions and causes.We have seen in our discussion of context that any selective analysis of cases needs a thorough understanding of how such cases relate to the data set as a whole. Savage, for example, notes the risks of so-called 'juicy quote syndrome' (cited in Geiger et al 2010). We might caution also against juicy case syndrome. There are of course ways in which singular cases may be enormously informative but we need a situated sense of what they reveal (Emmel and Hughes 2009). A case may be deemed theoretically rich but we need grounds for ensuring we have grasped key features of context in properly adducing causal processes. A single case cannot 'stand' only because it beautifully exemplifies a particular theoretical or policy claim. The links which are drawn between experience, meaning and context will be more rigorous, and available to scrutiny, if we show how we have brought them into conversation or comparison with other cases. Analysts can thereby illuminate the rationale behind their choice of cases for detailed analysis, or the choice of iconic cases in published work.

We have already considered contexts of data production. However, learning the details of project data refines our understandings of such contexts, especially when we seek to work across data sets. It

is when immersed in reading and seeking to interpret project data that the detail of contextual specificity is shown in sharp relief. We move now to a consideration of data analysis.

5. Developing analytic strategies

5.1. Overview of analytic strategies

There will be very varied approaches to identifying suitable analytic strategies. As secondary analysts some may approach data with a view to gaining an overall understanding and then undertaking thematic analyses, perhaps theorising how specific themes manifest across the data sets. Many will follow a more case based analysis, often centred on individual participants. This has most typically been our approach. We have read and re-read data seeking to explore the nature and organisation of phenomena of interest to us (e.g. experiences; aspects of identity; values; motivations) and the ways these relate to diverse contexts and broader social structures and processes. We have broadly taken a realist perspective. Below we illustrate how we have sought to move between the specifics of detailed in-depth case data and evidence on general social processes.

5.2. Exploring common questions: contexts and evolving concepts

One way we worked across data sets was to take as a focus some questions which were asked in common across the different projects. This was part of a shared Timescapes commitment to explore the scope for working with some data generated around common questions, albeit these were a very minor part of each project. One of the questions asked people to recount turning points and significant events in their lives. We reviewed responses to these questions across 4 Timescapes projects. It may well be that the question is most productive within intensive case history analysis which is more readily associated with primary rather than secondary analysis (Thomson et al 2002; Holland and Thomson 2009). Nevertheless it was of interest to us partly as an early inroad into the diverse project data sets. As we have seen, exploring how these common questions 'worked' was also helpful for revealing some of the dilemmas of working across data sets. In particular it helped illustrate many ways in which context shapes data, in terms of question wording, framing, positioning, advance notice and so on, and also in respect of the wider method and research design. Subtle differences in meaning can orient research participants quite differently. Minor differences in forms of questioning or lines of enquiry can engender evidence about one facet of experience rather than another. Working across data sets very quickly puts into relief the ways data is embedded in its conditions of production.

We pursued our line of inquiry into responses to questions asked in common across the projects (these formed a small component of project interview schedules and were incorporated in different ways, and with varying degrees of integration within project designs). As secondary analysts working with different data sets, one aim was to see if we could reach beyond project specifics and generate meaningful comparisons across projects, given the challenges described above. In our readings of data across projects we developed an interest in the differing kinds of accounts of turning points people provided. Interpreting this diversity presented challenges due to the varied project contexts and modes of questioning. For example, some projects and interviewers asked about significant events in peoples lives whilst others asked participants about turning points, or a singular turning point, and sometimes these questions were brought together. These framings may have had an internal project logic, but they tend to generate different kinds of responses. For example, a turning point to us suggests some perception of a redirection of one's life, and may generate very different kinds of responses to a question about significant events. We looked closely then at

accounts where people were asked about, or appeared to orient to, turning points as referencing some kind of biographical redirection. However, it is important to be aware that any straightforward comparison across projects is a non-starter, even where some element of standardised questioning was agreed in principle.

On our reading of the data (and this was a preliminary and not an extensive analysis) we understood perceived turning points were strongly linked, for some project participants, to their capacity to act to redirect their lives. This was not a general pattern, far from it, but in some contexts such as experiences of divorce we saw evidence of women describing in quite stark terms their own agency and decision making around redirecting their lives in some more positive way (this is drawn from evidence in the Work and Family Lives Study[vi]). Amongst others, notably in The Oldest Generation sample, there was a less vivid picture of agency. Where we believed some participants to be oriented to turning points, as a redirection of their lives, their accounts hinged less on personal capacity and more on experiences outside these participants' own personal control. Why might this be so? Bringing together data from different projects reveals some issues in understanding the bases on which we are comparing data. For example, the more limited account of agency in the older people's accounts of turning points could arise for different reasons. It might reflect the nature of turning points described. It might stem from social historical circumstances in which older people felt they were less able to be authors of their own lives. With very differing implications, it might be that the perspective of later life offers a more sociological view of lives in context: for these older participants distance (and sometimes death of close others) puts agency in context (Irwin, Bornat and Winterton 2011). Or, just as difficult for the analyst, it might be an outcome of the research study design. For example, it is plausible that the focus of studies which were primarily researching current experiences and behaviours (such as the Work and Family Lives study) encourages participants to foreground their own personal strategies and agency, in contrast to those in The Oldest Generation, a study centring on oral history, in which participants in later life may themselves orient more to their lives in social and historical context. Finally, and especially since we are dealing with a small number of individuals, it is possible that the patterning we observe is simply something we have conjured up, as secondary analysts, an outcome of our own desire to find order rather than disorder.

These questions may not be resolvable in the extant data sets. Certainly they were not designed with such questions in mind. Are there ways in which we can test out the alternative interpretations with reference to the project data sets? We could seek to generate hypotheses which make a virtue of cross project working. For example, we might seek parallel events or turning points (bereavement, divorce) and explore the extent to which mid life, and later life participants provide similar or different accounts. In this way we seek to isolate and exclude from consideration non-relevant causal or contextual factors, and develop a more precise account of the intersection of contexts and participants' accounts. Or we might seek out other kinds of evidence on experiences of agency and constraint within the projects, and build a broader conceptualisation of the contexts in which agency and constraint are perceived. We might work across projects to interrogate perceptions of agency and constraint in extremely diverse circumstances of wealth and poverty (the Intergenerational Exchange project offers a distinctive lens here given its focus on socially excluded, mid-life grandparents). These strategies are essentially comparative ones in which we are seeking to refine our understanding of the contexts which give rise to particular experiences and accounts, working within projects and then making a virtue of possibilities for cross-project analysis. In the next example we discuss further issues in working across data sets, and develop our argument that we need to work with evidence 'in translation' across projects, as a necessary grounds for any sort of comparative analysis.

5.3.   Exploring common themes: translating concepts and evidence across different data sets

The Timescapes projects had a number of substantive areas of interest in common, albeit ones quite generally construed, for example biographical change, life course transitions, family relationships, along with an interest in temporal processes. We came to projects with broad conceptual questions which we then sought to refine on reading available data on specific substantive themes. Settling on themes which could be analysed *across* data sets was itself a challenge since the commonalities were typically at a general level (see Irwin and Winterton 2011b), more so than had been anticipated in the original secondary analysis project design. This presented challenges given the particular framing of the Timescapes secondary analysis project. It is nevertheless quite possible that secondary analysts more generally will also often need to refine their questions in line with the available data, which may not be exactly as they predict or desire it.

Projects supplied us with their fieldwork materials including interview schedules across waves. Through a reading of these secondary materials and extensive transcribed interview data across projects we identified some specific (if still general) themes which were common across subsets of projects. This process was iterative, and both researchers read data following up different lines of inquiry. We documented these within a discussion paper which we circulated to project teams to invite their reflections and feedback, and subsequently met with all project teams to discuss issues relating to secondary analysis of project data.[vii] The discussion paper now exists as a Timescapes working paper (Irwin and Winterton 2011b). We settled on a set of questions relating to gender in contexts of parenting young families, a theme of 3 of the Timescapes projects and an area of interest to us as secondary analysts. We focused on the longitudinal data to which we had access, across two of the projects.[viii] Through readings of data by project, and across interview waves, we evolved a series of more precise questions and areas of potential interest.

To illustrate some of our thinking here we take as an example some of our analysis of data on gender and time stress in the family lives of parents with young (primary school age) children. [ix] An aim here was to consider if and how we could work in a way which would allow a meaningful analytic 'conversation' across differently constituted data sets. We developed an inductive approach exploring data from one project (Work and Family Lives), reading and re-reading cases and building an understanding of how a phenomenon of interest (experiences of time pressure) manifest across the sample, and how individual cases were situated in this respect, and in respect of each other. We evolved our focus having explored different avenues within the data broadly relating to gender and values. Although 'inductive', our understanding of gendered experiences of time pressure connected to a broader conceptual knowledge rooted in wider evidence and research (on gender, time pressure and work life stress).

We refined our understanding by exploring women's and men's experiences of time pressure across different household circumstances, with varied divisions of labour in doing paid work and care. This informed an analysis of asymmetry and inequality of gendered experiences of time pressure across different contexts within the Work and Family Lives data set. We first read across all transcripts of adult participants in households where they were co-residing with their partner or spouse. (We later read outwards to include some lone parent interviews). On this basis we sought to 'map' individual cases with reference to the specific dimensions of experience in which we were interested. Our understanding of pattern is built inductively with reference to the content of the data. An advantage is enhancing our understanding of the organisation of diverse experiences of time pressure, and how individuals are 'situated' with respect to this. A disadvantage is that the way we organise the data (cases) relates to a very specific dimension of interest so we need ensure it does not become an

overarching organising framework for other kinds of questions we may want to ask. We sought to use the available data to explore accounts of time pressure amongst women and men when they have significant paid work commitments and children, across different divisions of household labour evident within the data set. For example, we compared the accounts of working women who were partnered by working men, with the accounts of working women whose partners took on quite extensive practical care commitments. On our readings, the evidence related well to extant evidence on experiences of gender and time pressure. For example, women across diverse circumstances appear more likely to *manage* the work of work life balance than do men, and generally appear more prone to experiencing pressure, particularly when they have extensive paid work commitments.

We then sought to bring the evidence into conversation with another data set, specifically Men as Fathers. Here we began with a more deductive way of entering into the project data set, starting with cases which looked potentially 'productive' with respect to our questions, based on project metadata (supplied to us by the team) about household divisions of labour. However, having followed this more deductive strategy we then read 'outwards', to give confidence we were interpreting the evidence in a way which was consistent with other cases/evidence within the sample. We chose for in-depth analysis examples of men in diverse circumstances so we could compare their experiences *within* the MaF project data set. Here we looked in-depth at a range of circumstances, but took a particular interest in men who worked extensively *and* desired extensive practical hands on care of their young children. Why? Because they were men in circumstances where we hypothesised some interesting light could be shed on general social processes. Echoing the strategy for Work and Family Lives data, we sought to explore circumstances which might generate less conventional outcomes, specifically lower time pressure for working women in Work and Family Lives; and greater time pressure for men in Men as Fathers. In effect we sought to bring evidence into comparison on the basis of *translating* our questions, and emergent hypotheses, to a new project context, as dissimilarity in project designs and samples meant that we could not simply ask identical questions across them. Rather we focused on the contexts in which time pressure is, and is not, experienced by women and men and brought these into comparison within, and across, projects. The evidence speaks to the very entrenched nature of the difficulties women in particular need confront in reconciling paid work and familial commitments. What we drew from the data here was somewhat speculative and based on small samples. Whilst our efforts here were partial they serve to illustrate a broader point: that secondary analysts need to be creative and critical in conceptualising how to *translate* evidence between differently constituted data sets.

5.4 Qualitative longitudinal analysis: longitudinal case studies and the social structuring of diverse trajectories

In our third example we specifically consider the analysis of qualitative longitudinal data. The fact that we have not yet tackled this in our discussion reflects some of the difficulties of undertaking qualitative longitudinal secondary analysis in conjunction with working *across* Timescapes projects. The projects all had their own specific rationales behind their longitudinal designs and there was no immediately apparent 'dovetailing' of longitudinal research questions that we were able, within our timeframe, to adduce evidence towards. We therefore come to a single project, Young Lives and Times. We came to this as secondary analysts, Mandy having had no role and Sarah a partial one in the primary project, and having previously worked with some of the early data (Irwin 2009).

In undertaking our analysis of longitudinal data here we developed case profiles (cf. Thomson 2007). We developed our longitudinal case based analysis with reference to social diversity, exploring the

interplay of specific influences on participants' (educational) identities and expectations, and how these evolve differently, through time, across social groups. We undertook a case based analysis with reference to diversity within the qualitative sample as this maps onto significant, class related, groupings across the population. As secondary analysts a sample structure may not be as we would wish it so there is need to maximise its potential. For example, Young Lives and Times had a sample quite heavily weighted towards middle class youngsters and towards privately educated youngsters. The majority of the sample expected to go to university and we focused on this grouping in this analysis. There was nevertheless some diversity within this grouping (for example in parents' educational backgrounds). We read much of the available material, including longitudinal interview data from the ages of 14 to 17/18. We then selected for in-depth analysis a spread of cases chosen strategically to illuminate diversity in family background and resources. This spread was very revealing also of diversity in young people's *temporal* experiences of family, school and peer influences in their evolving orientations to higher education (Winterton and Irwin, in preparation). The link to external evidence here is through working up theoretical links between individual cases, the pattern into which they fall, and extant evidence on relevant processes.

The case based approach helped us to explore the dynamics of social inequalities operating at a biographic, micro, level of social experience, self perception and interactions with significant others. Our analysis confirmed the appropriateness of organising the cases with reference to family (class related and higher education) background and youngsters' perceptions of their parents' expectations, important dimensions around which the data cleaved. We used these as axes around which to organise further analysis of the linked influences of school, teachers, friends and other more contingent factors. In contexts where parents were in middle class occupations and had been to university it was also the case that parental expectations, school contexts and friendship influences were aligned, and pulling in the same direction. Here youngsters held assured expectations of going to H.E. *throughout* their teenage years. Other youngsters expecting to go to university had no family higher education background. There was an interesting division within this grouping. Amongst some youngsters the evidence showed how expectations were acquired and firmed up through their teenage years, in part through the influence of parental resources, expectations and investment in private education. Amongst other youngsters we saw greater contingency in their expectations. Here familial H.E. influences were 'weaker' and peer group and school level influences tended to pull in different directions over time. Expectations here were more subject to vagaries, and in a state of flux. Such vagaries are not random but have a logic and influence which is structured by circumstance and background. Overall, then, the evidence reveals the interplay of different influences over time and how these underpin, or render uncertain, evolving ideas about going to university amongst youngsters from different backgrounds.

We suggest, then, that detailed longitudinal case based analyses may orient us to the particular, but it simultaneously reveals how the interplay of these factors over time varies by social background and circumstance, and provides a revealing lens on the temporal, biographical confluence of processes shaping inequalities. A case based longitudinal analysis organised with reference to how diverse (here class related) experiences are situated, and evolve over time, offers a powerful resource in theorising the structuring of inequality. Whilst our resulting arguments about the shaping of diverse trajectories are related to a small sample, they can be tested and refined by exploring them across different contexts.

6. Conclusion

This paper relates closely to our brief Methods Guide on secondary analysis, but provides a more detailed and more fully exampled account of our work on the Timescapes Secondary Analysis Project which ran during the final two years of Timescapes. We have discussed the difficulties of, and strategies for, familiarising ourselves with data sets not of our own design or making. In so doing we hope to have also provided a spur to primary researchers who deposit data for re-use to reflect fully on what metadata and guidance they could usefully archive alongside their data. We have discussed issues relating to context, and how accessing this needs be understood as a conceptual as well as a practical issue. It is a qualitative research commonplace to say that data is embedded in the contexts of its production (so then contexts are embedded, often hidden, in data). We have argued the importance of ongoing critical judgement about the contextually embedded nature of data. However, this should not hamper, but be part and parcel, of evolving substantive analyses. We also sought to develop some strategies for working across data sets and thinking productively about how we can test out and refine our evolving concepts by bringing data sets into conversation, and by appropriately *translating* evidence between them. We also illustrated briefly an analysis of micro level qualitative longitudinal data where we could explore diversity with reference to, and as a lens upon, social structural inequalities.

The discussion we have provided is a guide, not a recipe. We have tried to distil some principles in the hope of facilitating general lessons, or at least to have offered an account which will have value for others tackling qualitative secondary analysis. The process of 'distilling principles' is partly post hoc. We did not start from the principles and then apply them, but rather evolved them, continually moving between practice and critical reflection. We have sought to reflect on principles too, more than offer prescriptions. There are many diverse ways to tackle analysis, and we can only offer reflection on the strategies we found useful. We hope through our examples to have illustrated some of the challenges for secondary analysts, as well as suggesting some 'routes through' to effective analysis.

Secondary analysis is a challenging undertaking. It is time consuming. It requires great persistence in ensuring an adequate understanding of details which may be tacit for primary researchers. It can be frustrating for researchers to become 'users' when they may be more used to controlling the who, what and why of research design and data. In conjunction, it may be a risky course to follow in so far as outcomes are uncertain yet time commitments can be extensive. However, as many insightful secondary analyses stand testament, there is a depth of social scientific insight and progress which can be achieved. There are a range of reasons researchers might seek to undertake secondary analysis, and with enhanced technology there is now an outstanding set of qualitative data resources readily available for exploration and analysis. We hope to have provided here some helpful guidance, and encouragement, to would-be secondary analysts.

References

Bishop, L. (2006) 'A proposal for archiving context for secondary analysis', *Methodological Innovations Online* 1 [2].

Edwards, R. and Irwin, S. (eds.) (2010) Special edition: lived experience through economic downturn in Britain, *Twenty First Century Society*, Journal of the Academy of Social Science 5 [2].

Emmel, N.D. and Hughes, K. (2009) 'Small –N access cases to refine theories of social exclusion and access to socially excluded individuals and groups', in Byrne D and Ragin C (eds.) *The SAGE handbook of Case-Centred Methods.* Sage. London

Geiger, T., Moore, N and Savage, M. (2010) '*The archive in question*', CRESC Working Paper Series, WP 81.

Hammersley, M (2009) 'Can we re-use qualitative data via secondary analysis? Notes on some terminological and substantive issues', *Sociological Research Online* 15 [1], at http://www.socresonline.org.uk/15/1/5.html

Holland, J & Thomson, R. (2009) Gaining perspective on choice and fate: Revisiting critical moments, *European Societies* 11 [3]: 1-19

Irwin, S. (2008) 'Data analysis and interpretation. Emergent issues in linking qualitative and quantitative evidence', in Hesse-Biber, S. and Leavy, P. (eds) *Handbook of emergent methods*, New York: Guilford Press.

Irwin, S. (2009) 'Family contexts, norms and young people's orientations: researching diversity', *Journal of Youth Studies* 12 [4]: 337-54.

Irwin, S., Bornat, J. and Winterton, M. (2012) 'Timescapes secondary analysis: comparison, context and working across data sets', *Qualitative Research,* (in press) (special edition: Advancing methods and infrastructure for qualitative longitudinal research in the UK: The Timescapes Experience;).

Irwin, S. and Winterton, M. (2011a) *Debates in qualitative secondary analysis: critical reflections*, Timescapes Working Paper no. 4.

Irwin, S. and Winterton, M. (2011b) *Timescapes data and secondary analysis: working across the projects*, Timescapes Working Paper no. 5.

Mason, J. (2002) 'Qualitative interviewing: asking, listening and interpreting', in May, T. (ed) *Qualitative research in Action*, London: Sage.

Thomson, R. (2007) 'The qualitative Longitudinal Case History: Practical, methodological and ethical reflections', *Social Policy and Society* 6 [4]: 571-582

Thomson, R., Bell, R., Henderson, S., Holland, J., McGrellis, S. and Sharpe, S. (2002) 'Critical moments: choice, chance and opportunity in young people's narratives of transition to adulthood', *Sociology* 36 [2]: 335-54.

Ward Schofield, J. (1993) 'Increasing the generalizability of qualitative research', in Hammersley, M. (ed) *Social Research. Philosophy, politics and practice*, London: Sage.

Winterton, M. and Irwin, S. (2011) 'Youngster's expectations and context: secondary analysis and interpretation of imagined futures', in Winterton, M. Crow, G. and Morgan-Brett, B. (eds) *Young lives and imagined futures: insights from archived data*, Timescapes Working Paper no. 6.