

# The Secondary Analysis of Archived Qualitative Data

by

**Louise Corti and Paul Thompson**

*[Pre-print. Published as Corti, L. and Thompson, P. (2012) 'Secondary analysis of archive data' in C. Seale, G. Gobo and F. Gubrium and D. Silverman (eds.) (2004) Qualitative Research Practice, (ed.), London: Sage Publications Ltd and as second edition in J. Goodwin (ed.) SAGE: Secondary Data Analysis London: Sage Publications Ltd.]*

## ***Introduction***

Archived qualitative data are a rich and unique, yet too often unexploited, source of research material. They offer information that can be reanalysed, reworked, and compared with contemporary data. In time, too, archived research materials can prove to be a significant part of our cultural heritage and become resources for historical as well as contemporary research. In this chapter we explore the methodological, ethical and theoretical considerations relating to the secondary analysis of such qualitative data.

There is a well-established tradition in social science of reanalysing quantitative data. Nor is there any logical intellectual reason why this should not be so for qualitative data. However, among qualitative researchers there is no similar research culture encouraging the reanalysis of data collected by other researchers. Until very recently there has been a striking lack of discussion of the issues involved and very little published 'evidence' of the benefits and limitations of such an approach. In this chapter we outline some ways of re-using data, and discuss their strengths and weaknesses, drawing on various examples including a case study by Paul Thompson. Thompson's reflections of his own personal experiences of reanalysing data from his pioneering oral history study conducted in the 1970s, provide the reader with first-hand evidence of both the difficulties and gains which resulted. It was these early experiences of re-use which led Thompson to establish for the National Life Story Collection at the British Library, and then ESRC's national archiving centre for qualitative data, Qualidata.

Finally, we summarise the ways archived qualitative research data are best presented so as to be of maximum potential use to other researchers. Experience is drawn from two qualitative data services, Qualidata in the UK and the Murray Research Center in the U.S. We discuss the organisation, preparation and documentation of data, and also legal and ethical issues.

## **1. Defining qualitative data and the implications for re-use**

Qualitative data are collected across a range of social science disciplines, often with varying techniques or emphasis; but typically aiming to capture lived experiences of the social world and the meanings people give these experiences from their own perspectives. Often a diversity of methods and tools rather than a single one are encompassed. The types of data collected vary with the aims of the study and the nature of the sample. Samples are most often small, but may rise to 500 or more informants. Such data include interviews – whether in-depth or unstructured, individual or group discussion – fieldwork diaries and observation notes, structured and unstructured diaries, personal documents, or photographs. Thus any one study may yield a wide range of data types for archiving. Moreover most of these types of data may be created in a variety of formats: digital, paper (typed and hand-written), audio, video and photographic.

In Britain the decades since 1950 witnessed an unprecedented flowering of social research: in the growth of its influence, in the spread of its themes, and in the development of its methods – both quantitative and qualitative. From the 1960s into the 1970s sociology was not only an exceptionally popular subject with students, but was also given more national research resources than at any time before or since. This enabled social researchers to carry out studies of a thoroughness unlikely ever to be equalled.

Just one example is Peter Townsend's in-depth investigation into the nature and status of older people's institutions in postwar Britain. The publication resulting from this research, *The Last Refuge* (1962) was considered a pioneering piece of research when it was published in 1957 and attracted much publicity for its focus on an important and hitherto neglected area of policy, and also for its methodology and its policy recommendations. But Townsend's meticulously preserved fieldwork descriptions of old people's institutions and accompanying interviews, now archived and available to researchers at the University of Essex, are in the long run equally significant: not only a unique historical record, but equally a rich, multi-layered resource for institutional research which seeks to explore the meaning and nature of institutional life across both micro and macro levels of analysis (Charlesworth and Fink 2001).

Clearly the scope and format of data usually determine its potential for secondary analysis. For example, data from a research study that collected, recorded and transcribed 100 in-depth interviews and documented detailed field notes, particularly when based on a clear sampling strategy, are much more likely to be useful than a set of hand-written interview notes from 20 brief semi-structured interviews. But as with all archived material, sometimes the most striking discoveries come from re-examining material which hitherto has not been thought worth researchers' attention. Thus Zeitlyn argues that field photographs and more recent recordings by camcorders are among the most prolific largest and least exploited resources produced by anthropology (Zeitlyn 2000).

### *The availability of qualitative data for re-use*

From the early days of social research onwards there have always been a minority of researchers who deliberately chose to publish their interview data. Classic instances run from Henry Mayhew's interviews with the London poor in the mid-19<sup>th</sup> century through to the much more sustained life story genre in North American anthropology (from Leo Simmons to Oscar Lewis and Sidney Mintz) and the two briefer waves of life story research in Chicago sociology (as in the work of Clifford Shaw and Helen Hughes) (Thompson 2000). Some of these – such as the Chicago sociologists – also deliberately archived their interviews. There are also exceptionally large archives of ethnographic research data in Scandinavia. But in such researchers have remained a minority. This is partly because of the defensive position which qualitative social research has had in relation to the much better-funded tradition of quantitative social research, in which informants' words disappear behind statistical findings and speculations. But it is also because even qualitative researchers can be more interested in offering their own interpretations in their papers, articles and books addressed to academia, rather than presenting their original sources for others to evaluate.

Thus not surprisingly, if we take a look across the world in an attempt to identify qualitative data sources that could be openly consulted, we immediately encounter problems. The first is the absence in most countries of any national effort to either gather together or draw attention to existing research sources. The second is the lack of infrastructures and also of agreed practical procedures for preparing, storing and disseminating qualitative data. Throughout the world there are innumerable archives which collect (mainly historical) qualitative material, as well as a large number of sound archives and ethnographic archives, but there are few common descriptive standards, access to many collections is poor, and there are no integrated resource discovery tools.

One of the earliest and perhaps best known sources in Britain is the collection of papers resulting from the 1930s social research organisation, Mass-Observation. These were established as a notably well-organised and accessible public archive at the University of Sussex in the early 1970s,

and since then have attracted a notably high number of researchers (Sheridan 2000). More typically, other data collections that were retained were stored as in-house research resources, such as the Berkeley and Oakland cohorts collected from the 1920s to the 1990s at the Institute for Human Development at Berkeley. It has also been not unusual for the papers of eminent scholars, sometimes representing a lifetime's research, to be transferred on retirement to their local university archives. Such papers may include not only primary research data, but also administrative documents about the research, such as grant proposals and correspondence. A collection may also contain 'secondary' sources utilised for a particular research study, such as newspaper clippings, organisational or medical records. University archives may sometimes broader evidence of theoretical or intellectual processes in the institution, for example the development of ideas within a key social science department, as was used by Platt for her account of the Chicago school of sociology (Platt 1996). More generally, however, attempts to archive qualitative research material were both rare and unsystematic.

It was only from the late 1980s that any sustained concern with archiving qualitative research data began to develop. Paul Thompson had earlier seen the importance of archiving the material from his national study of 'Family Life and Work Experience before 1918' (Thompson 1975), a unique and unrepeatable set of 444 interviews with men and women born before 1906. These interviews were kept in a special room within the Sociology Department at Essex, and as a result became the basis of a series of books and articles by visiting scholars. But this was an informal archive with no secure future, and this led him to be concerned about the lack of any national archival facilities for in-depth interviews. The National Sound Archive had only just been established, and was then without an oral history section, focussing mainly on music and nature.

In 1987 Paul Thompson, with the support of Asa Briggs, set up the National Life Story Collection as an independent charitable trust within the British Library National Sound Archive, and at the same time the NSA appointed Rob Perks as its first oral history curator. The result has been a dramatic improvement in the archival situation. In the one hand NLSC has been generating new life story history projects creating public research resources, such as the recording projects on artists and sculptors jointly with the Tate Gallery, on holocaust survivors, on the financial elite of the City of London (Courtney and Thompson, 1996), on the book trade, the post office, North Sea oil workers, and on workers of British Steel (Dein and Perks 1993). On the other hand, NSA has become a national repository for the tape recordings from social research interviews. It is this cross-disciplinary approach which is unusual, for across the world the oral history community now has a professional interest in preserving tape recordings gathered from oral history interviewing projects, and indeed in the U.S. this archiving goes back, for example at Columbia University Library and at the Bancroft Library at Berkeley - for over forty years. In Europe more recent oral history archives across Europe include in Germany the oral history archive of "German Memory" based in Hagen, which comprises some 1,500 life history interview recordings with witnesses of time periods from East and West Germany (Leh 2000); and in Hungary, the 1956 Institute holds oral history interviews, trial records and photographs dealing with research relating to the 1956 Hungarian Revolution, its development and subsequent effects (Lux 2000).

Since the 1990s new technology has led to new possibilities for sharing qualitative data, especially through online resources and databases. For example in anthropology, Paul Stirling's longitudinal research on two Turkish villages between 1949 and 1994 has been electronically archived. The resource consists of over thirty volumes of field notes, and a selection of his other field materials, including photographs, video and audio tapes (Zeitlyn 2000). Of the many approaches to digitisation, we see three as particularly promising. First, web-based multi-media samplers of key qualitative materials can provide 'edited highlights' to illustrate the potential of the collection for research and teaching. Second, thematic sets of interviews, across topic or time can be drawn together by interweaving raw data collections and related methodological information into an intellectual and substantive framework. Finally, the mark-up of textual documents enables a range of data types and formats, to be accessed via the Web, as demonstrated by the *Edwardians On-line* project, undertaken by Qualidata (Barker, 2002).

Other examples come from field of socio-linguistics. Since 1990 the Institute of Psychology at the University of Zurich has archived video-based interview narratives from psychotherapy sessions in order to investigate the linguistic structures within psychotherapeutic communication. The collection includes digitized audio-visual data, patient card files, transcripts of therapy sessions, extracted narratives, research publications and student research papers that are integrated into a central database (the JAKOB database) to make it accessible for research purposes (Luder 2000). Equally, there are corpora of speech that are archived. The Spoken Language Corpus of Swedish at Göteborg University (Allwood 2000) is a growing corpus of audio-visual spoken language samples from several languages taken from naturalistically occurring interactions from as long ago as the early 1980s. The goal of the corpus is to capture differences in pronunciation, vocabulary, grammar and communicative functions.

### ***Towards national archiving***

Nevertheless in Britain, as indeed in most western countries, until recently no infrastructure existed for the systematic archiving and dissemination of qualitative data from social science research. The Economic and Social Research Council (ESRC; then SSRC) had already recognized from very early on in 1967, the value in retaining the most significant machine-readable data from the empirical research which it funded by establishing a Data Archive. Since the 1970s, social science data archives across the world have typically acquired a significant range of data relating to society, both historical and contemporary, from sources including surveys, censuses, registers and aggregate statistics. Equally, these centres of expertise have established networks of data services for the social sciences which foster co-operation on key archival strategies, procedures and technologies.

Thus crucial survey data, can be re-analysed by other researchers, and the money spent on research has become not only an immediate outlay but an investment for the future. There was, however, a significant gap in this policy in that qualitative data were rarely acquired, even when much interview data became transcribed in word processed form. When a small pilot study commissioned by the ESRC was carried out by Paul Thompson in 1991 (Thompson 1991), it was revealed that ninety percent of qualitative research data was either already lost, or at risk, in researchers' homes or offices. However the ten percent 'archived' were found not to have the basic requirements of an archive, such as physical security, public access, reasonable catalogues, with recorded material or listening facilities. It was further calculated that it would have cost at least £20 million to create a resource on the scale of that at risk. For the older British sociological material, moreover, the risk was acute, and the need for action especially urgent. This was borne out by the very recent destruction of research data on the classic British community studies of Banbury (Stacey, 1974); on race and conflict in Sparkbrook (Rex and Moore, 1967); and the unique longitudinal study on child-rearing, consisting of over 3,000 interviews recorded over thirty years by John and Elizabeth Newson (1963, 1968, 1976, etc) – all lost after the retirement of these researchers.

In 1994, with support from the ESRC, the first UK qualitative data archiving project on a national scale was established at the University of Essex. Its first task was a rescue operation aiming to seek out the most significant material created by research from past fifty years. The second was to work with the ESRC to implement a Datasets Policy (ESRC 2002) to ensure that for current and future projects the unnecessary waste of the past did not continue. Qualidata was not set up as an archive itself, but as a clearing house and an action unit, its role being to locate and evaluate research data, catalogue it, organize its transfer to suitable archives across the UK, publicize its existence to researchers and encourage re-use of the collections. (Corti, Foster and Thompson 1995; Thompson and Corti 1998).

Qualidata established procedures for sorting, processing and listing both raw data and accompanying documentation (metadata); systematically describing studies for web-based resource discovery systems; for establishing appropriate access; and for training in the re-use of qualitative data (Corti 2000). By 2002, Qualidata had acquired, processed and catalogued some one hundred and forty datasets, and catalogued a further one hundred and fifty already housed in archives across the UK. Surviving 'classic studies' data from key researchers were also rescued, including well-known British single projects such as John Goldthorpe et al's *The Affluent Worker* (1962); Stan Cohen's *Folk Devils*

and *Moral Panics* (1967); and the entire life's work of pioneering UK researchers such as Peter Townsend's *Family Life of Old People* (1955), *The Last Refuge* (1962) and *Poverty in the UK* (1979) and Paul Thompson's life-history interview studies of *The Edwardians* (1975) and *Families, Social Mobility and Ageing, an Intergenerational Approach* (1993).

From 2001, Qualidata began a new life as a specialist unit housed within the UK Data Archive (UKDA) at the University of Essex, with a focus on acquiring and distributing digital data. Qualidata's emphasis on providing easy quick and access to data has led to the development of the Edwardians Online project (Barker 2002), an online resource that offers thematic access to a collection of oral history interviews with people who lived in Edwardian Britain. The multi-media resource integrates the original text transcripts, digital sound-bites of the original audio tapes, background material concerning the original research study, and details of publications based on secondary studies of the interview texts. This resource has provided a model for the digitization and interactive online provision of 'classic collections' based on qualitative data for research and teaching resources.

In the US, there is also a centre that has been systematically gathering qualitative as well as quantitative research data in order to make it available to other social science researchers. Founded in 1976, the Murray Research Center: a Center for the Study of Lives is a national repository for social and behavioural science data on human development and social change, with special emphasis on the lives of American women (James and Sorenson 2000). The archive holds more than 270 data sets with a wide range of topics, samples, and designs. Many of these studies include in-depth interviews or surveys with some open-ended questions. The Center holds a major collection of longitudinal studies of mental health, including Glueck And Glueck's (1968) *Crime Causation Study*, some material from the Institute of Human Development, and Lewis Terman's (1954) *Life Cycle Study of Children of High Ability*. In the area of racial and ethnic diversity, an important study is Brunswick's (1980) *Harlem Longitudinal Study*.

Finally, over the past few years there have been a number of other initiatives across the world that have sought to establish national archiving projects for qualitative research data. At the time of writing, a small-scale Czech Archive of Qualitative Data has been recently established at Masaryk University; proposals are being prepared in Germany and Switzerland for qualitative archival resource centres; and national (survey-based) Social Science Data Archives in Finland, Netherlands, Denmark and Canada are conducting feasibility work for extending the scope of their own collecting.

This build up of qualitative data resources has thus encouraged the uptake of secondary analysis. It also reflects some of the efforts invested in promoting or re-packaging data collections to meet researchers' needs. And as resources grow, so experiences of secondary research have begun to find their place in the academic domain.

## **2. Approaches to re-using qualitative data**

The re-use of qualitative data provides an opportunity to study the raw materials of recent or earlier research to gain both methodological and substantive insights. Because new data are usually expensive to collect, using existing sources may cut costs as well as avoiding duplication of effort. Nevertheless there has been a noticeable silence on this issue among the qualitative research community. While there are a large number of published texts describing different styles of qualitative interviewing (e.g. Burgess 1982), and fewer on how to analyse and interpret interview material (Plummer 1983; Silverman 2000), there are none that have appraised the secondary analysis of qualitative data. Methodological handbooks recommend that newcomers to the qualitative method should undertake a fieldwork project to gain experience of data gathering, data analysis and writing up, yet rarely introduce the concept of exploring existing data sources to answer a research question. Surprisingly, even when a historical perspective is explicit, earlier qualitative research data have not been considered as a source (Marwick, 1970). Approaches to re-using survey data, on the other hand, are well documented with guidance on both theoretical and practical approaches to the method.

The ways that qualitative data can be re-used have much in common with those familiar for the secondary analysis of survey data, yet there are different and perhaps more challenging intellectual and practical problems for the user to consider. Here we identify six approaches to re-using data. We discuss in relation both to theoretical issues recently raised about re-analysis, and to the actual experiences of researchers <sup>1</sup>. Issues relating to the practical and perceived barriers to secondary analysis follow this discussion.

### *Description*

The first use of qualitative data is descriptive, and the possibilities here are very wide, encompassing both the contemporary and historical attitudes and behaviour of individuals, groups and organisations, or societies. Thus when David Kynaston came to write the final volume – from 1945 until 2000 - of his authoritative series, *The City of London* (2001), he was able to draw on the 'City Lives' interviews archived in the National Life Story Collection. These were first used by Cathy Courtney and Paul Thompson for their book *City Lives* (1996). Kynaston describes the original interviews as 'invaluable', 'a treasure house' for his own volume: 'City men have never been great talkers... The interviewers proved extraordinarily effective at getting even normally taciturn subjects to speak candidly.' The interviews were especially vivid in showing attitudes, such as the importance of 'trust', the consequent perfunctariness of checks on behaviour, the beliefs in nepotism and in using insider knowledge, and the reluctance to 'go American' and introduce new practices (2002).

In the same way, any significant data created now, will in time become a potential historical resource. The oral testimonies of ordinary men and women can complement official and public sources such as newspapers or government reports. Such evidence can also be used to document individual lives – including those of significant researchers – in a biographical approach. Similarly, Sheridan (2000) notes how Mass Observation material has been used not only to provide historical evidence, but also to examine the role of the Mass Observation itself in the social, political and cultural milieu of the 1930s and 1940s. Researchers wishing to re-use material in these ways need to first evaluate the evidence, examine its provenance, and weigh up the veracity of the sources – in short, to be prepared to adopt the practice of social historians.

### *Comparative research, restudy or follow-up study*

Secondly, qualitative data can be compared with other data sources or be used to provide comparison over time.

The cumulative building up of knowledge in the social sciences has been incremental, resting on the foundations of earlier findings, and hence interpretation has often depended upon comparisons: with other contexts, other periods of time, other social groups, and other cultures. Comparison is most effective when there are sufficient data to enable convincing re-evaluations. It is fortunate that many social scientists grasped this relatively early. For example, in Britain the original returns of the population census were kept as public records, and have proved an invaluable basis for consultation in recent years. Similarly Sidney and Beatrice Webb (1984), on completing their pioneering study of British trade unionism, archived their field notes from their national sample of interviews, which remain the principal source of information on trade unionism in the late 19th century.

---

<sup>1</sup> In 1999 Qualidata conducted a national survey of academics and researchers in 1999 to ascertain what kind of data resources researchers wished for, receiving over 550 valid responses from a wide range of user communities. 92% wanted to see datasets in a digital and accessible format that could be used for both research and teaching purposes. The highest demand was for qualitative data on health, criminology and social policy (Corti and Thompson, 2000).

Early classic restudies include Seebohm Rowntree's (1901) three surveys of poverty in York, Llewellyn Smith's (1930-5) repeat of Charles Booth's (1891-1902) famous poverty survey in London, or the Lynds' studies of *Middletown* (1929, 1937). In anthropology perhaps the best-known instance is the controversial restudy and reinterpretation by Oscar Lewis (1963) of Redfield's (1930) research on the village of Tepotzlan in Mexico.

Comparison brings greater power to answer scientific questions, for example when a dataset can be combined with data beyond its own sample or geographical limitations. An example of a comparative study from the U.S. Murray Research Center cited by James and Sorensen (2000) is provided by Stewart and Healy (1989). These researchers combined five archived datasets to create a multi-cohort study using coded questionnaire data and open-ended material for developing constructs such as employment patterns, careers, role models, internal conflicts and world views. The pooled cohort data spanned some 40 years dating from World War I to the 1960s baby boom. Using this rich combined sample, the authors showed how both social pathways and personality development differed by cohort, influenced by events and social changes such as the impact of the Vietnam War or trends in female employment.

Equally, samples from original studies that have been preserved can be followed up. Sometimes this is because the original longitudinal study is still ongoing but can be redirected, as with Glen Elder's *Children of the Great Depression* (1974), based on both new fieldwork and a reorganisation of the earlier interviews and participant observation of the Berkeley and Oakland cohorts recorded from the 1920s onwards. Alternatively, sometimes the sample from a one-off study can be re-contacted, allowing the new investigators to re-design the outcome measures, so that a study that was not longitudinal becomes so (James and Sorensen 2000). Most often, however, the original data will simply be re-analysed. One such example, using data archived at the Murray Research Center, is Franz and McClelland's (1991) follow-up of Sears, Maccoby, and Levin's (1957) *Patterns Of Child-rearing Study*. In 1951 mothers of 5-year-old children in the Boston area were interviewed about their own and their husband's parenting practices. The children from this sample were re-interviewed at later stages into adulthood. The resulting longitudinal data spanning some fifty years was reanalysed with a focus on predictors of types of adjustment.

Follow-up studies typically require ethical approval. In addition, particularly in the health field, original investigators are often keen to become collaborators, rather than just being cited as the original data collectors (Corti and Wright 2002).

### ***Re-analysis or secondary analysis***

Reanalysing qualitative data, as with the secondary analysis of survey data, allows both for re-interpretations and also for new questions of the data. New themes can be studied. For example, childhood sexual abuse was not brought to public attention until the early 1980s, long after many of today's important longitudinal studies were launched (James and Sorensen 2000). Charlesworth and Fink (1999) draw upon research material from *The Last Refuge* (1957), Peter Townsend's study of institutional care, to illustrate the potential which his archived data holds for the analysis of such diverse topics as the power dynamics within institutions, the spatial organisation of the workplace and the relationship between research and policy.

Similarly, new angles can be applied, or new methods employed which may not have been possible at the time of the original data analysis. Sometimes new analytical tools can spotlight sections of data that were previously ignored. In this book, Åkerström, Jacobsson and Wästerfors argue for the significance of 'continuous analytic digging', allowing themselves the opportunity to revisit and reanalyze material, even if already written up. In general, the more in-depth the material, the less likely that has already been fully exploited. Finally, theoretical sampling, as proposed by Strauss and Corbin (1990), can also be applied to the data sample itself (e.g. Szabo and Strang 1999).

There are relatively few published examples of experiences of re-analysis of data in the UK. Fielding and Fielding (2000) present an insightful account of their work on revisiting Stan Cohen and Laurie Taylor's (1972) original analysis of long-term imprisonment of men in maximum security published as *Psychological Survival*. They argue for the ethical and practical advantages of re-analysis of existing data in their field:

secondary analysis has a particular role in qualitative research, addressing sensitive topics or hard to reach populations, because researchers can best respect subjects' sensitivities, and accommodate restricted access to research populations, by extracting the maximum from those studies which are able to negotiate these obstacles. Secondary analysis can protect the sensitivities of subjects and gatekeepers by ensuring they are not over-researched, and can position further enquiries so that they ask what is pertinent to the state of analytic development, building on, rather than simply repeating, previous enquiries (2000: 678).

In their findings, Fielding and Fielding further state how the reanalysis offered 'a means to extract further analysis purchase from research on a group seldom exposed to fieldwork' (2000: 688).

In the US, the Murray Research Center offers some excellent examples of re-analysis. These demonstrate, from a variety of disciplinary perspectives, the ways that qualitative data can be 'radically restructured' for new research. Creative approaches include seldom-used methods such as creating new prospective studies out of existing ones and using multiple data sets for multi-cohort designs. The availability of raw data, such as transcripts of in-depth interviews, observations, and responses to tests, are especially valuable in enabling the application of different perspectives and new scoring procedures to the original data (James and Sorensen (2000).

In the field of psychology, Elaine May (1988) reanalysed a study by Kelly on long-term personality development among married persons. While Kelly had originally mainly used psychological test assessment, May shifted her focus to the respondents own testimonies. In these they 'wrote about their lives, the decisions they made concerning their careers and children, the quality of their marriages, their family values, their sexual relationships, their physical and emotional health, and their major hopes and worries. In these open-ended responses, freed from Kelly's categories and concerns, they poured out their stories' (May 1988 p. 12). May combined these testimonies with other documentary sources to write about how the Cold War period affected aspects of family life, thereby challenging the assumptions of the happy housewife of the 1950s and suggesting some of the causes for disruption in family life over the following decades. James and Sorensen highlight how 'quotations from the Kelly data, sprinkled liberally throughout each chapter, illuminate and validate her claims in ways that closed-ended answers to the survey questions never could' (2000, para 46)

A second example from a U.S. based criminological study concerned the restructuring and recoding of data. Laub and Sampson (1998) applied event history analysis techniques, not yet formulated in the 1960s, to data from a longitudinal study, *Unravelling Juvenile Delinquency* (Gluek and Gluek 1968). The sources included case records, interviews with the respondents and their families, teachers, criminal justice officials, employers, and correspondence relating to family and school experiences, employment and the military service. Laub and Sampson re-coded a wide variety of materials into a longitudinal framework for studying, very successfully, criminal careers.

### ***Research design and methodological advancement***

Data can be used to help design a new study or develop a methodology or research tool by studying the sampling methods, data collection and fieldwork strategies and interview guides of earlier research. The Mass Observation Archive has been used to explore the process of researching, including methods at both the collecting and interpretation stage, and ethical issues (Sheridan 2000). Paul Thompson refers below to the stimulation of drawing on interview guides designed by earlier researchers in a similar field.

There is a growing emphasis on publishing the details of a study's methods in books and reports, but often the details offered are too brief. Of course earlier research may be equally valuable as a warning of potential dangers and pitfalls, of what to avoid as much as what to follow. Perhaps the most dramatic instance of this was the publication of Malinowski's fieldwork diaries, peppered with the racist and sexist assumptions which helped to shape his social observation (1967). Researchers' own research notes, fieldwork diaries, or memos can both offer much insight into the history and development of the research, and also help inform new thinking.

### ***Verification***

Archived data can be scrutinized with scientific rigour to support or challenge a set of findings or to appraise the method. Generally, qualitative researchers are not used to making their findings accountable, and some may feel vulnerable about others seeing their data, leaving themselves open to criticism. But the practice of opening data for inspection may lead to better and more transparent research. On the other hand, even where archived data are already available for research findings to be verified, as yet, there is no evidence of this actually happening

Hammersley (1997) identifies the benefits and weaknesses of using 'replication' to check findings. True scientific replication, he argues, is not possible as studies generally do not have equal social phenomena. Even restudies suffer from differences in time and the researchers' subjective perspectives. Well-documented and 'complete' data sets can help the scrutinizer to reconstruct the evidence by re-tracing the original analytic steps. However, it is unlikely that the research process could ever be made fully explicit, primarily because the path of qualitative analysis is never linear, and is almost certainly likely to involve a degree of trial and error in the pursuit of interesting lines of investigation. A researcher report is not simply a representation of data peppered with narrative and interpretation, but is in itself a social and literary construction (Kvale 1996).

Fielding (2000) suggests that qualitative software may help the process of verification. Retention of the original coding frames means that these can be reapplied by another investigator, providing us with a type of audit tool.

### ***Teaching and learning***

Older 'classic' studies in the social sciences and more contemporary focused sets of transcripts along with supporting documentation can provide valuable material for social science teaching, both in research methods and in substantive areas. Students can learn many fundamental aspects of qualitative research, and the theoretical and methodological strategies that helped to create chosen datasets, while also gaining first-hand experience of critically re-analysing and comparing data from well-known sources.

A good example of the learning potential of such datasets is provided by the Katherine Buildings study undertaken by Peter Townsend in the late 1950s, and now archived by Qualidata. This unpublished study of social change is itself a re-study. It focused on the inhabitants of Katharine Buildings in Stepney, London, established in 1885 as an experiment in improving working-class housing conditions. Townsend's study includes his analysis of the original ledgers and notes kept by the managers of the Buildings, archived at LSE, who included the young Beatrice Potter (Webb), and his own interviews with residents, some of whom were descendants of the original inhabitants. The questionnaires, with many open-ended questions, provide a unique picture of a working class community in transition. Using these materials in a learning environment, students can examine data collected using sociological and anthropological research methods, and can compare these with earlier historical data collected in the late nineteenth century. Training exercises can take up aspects of long-term social change, such as the kinship patterns recorded in both Townsend's interview material, questionnaires and notes, and Webb's detailed notes on the tenants (Qualidata 1999). Students can

also be encouraged to examine the methods used and the research outcomes, and to consider whether they may have approached the research differently.

Learning about the work of researchers who have made a significant impact in their field allows young researchers to take the best practice elements from this work and further develop them in their own research work (cf Zeitlyn 2000).

### **3. Difficulties in re-using data**

We have demonstrated earlier that the practice of secondary analysis of qualitative data is not a commonplace research activity. Why has there been a reluctance to draw on material created by other researchers? Is it that it is a problem of the implicit nature of qualitative data collection and analysis? Or is it a question of lack of time to get fully acquainted with research materials created by someone else? How constraining is informed consent? And what about scientific verification – is there an insecurity about the exposure of one's own research practice?

The recent, though sparse, literature points to a number of key concerns regarding the practice of re-using qualitative data. However, in discussing the issues directly with qualitative researchers, it appears that the views are by no means homogenous. In fact, when asked what, if any, barriers existed to further exploitation of data by a secondary analyst, responses varied from overt support for sharing one's own data to vehement displeasure at the thought of being asked to share a 'possession' considered to be of personal value (Corti et al 1995; Corti 2000; Fink 2000).

Here we identify the key issues that present themselves as difficulties in both re-using and sharing data, and discuss possible measures for overcoming them.

#### ***Ethical and consent considerations***

The prime anxiety of most researchers relates to questions of confidentiality and agreements made at the time of fieldwork. Such concerns about confidentiality have been discussed in detail in other papers (Corti, Day and Backhouse 2000), and a number of guidelines for researchers on the ethical and legal issues surrounding informed consent, confidentiality and copyright with respect to re-using data have now been published (e.g. Qualidata 1998; the ESRC 1999; and the Oral History Society 2002).

Archived data should always conform to ethical and legal guidelines with respect to the preservation of anonymity when this has been requested by informants or guaranteed to them. There are various ways of achieving this, ranging from fully anonymising the original data – which in some cases would seriously damage its value – to obtaining legal undertakings to protect informants' anonymity from any researcher allowed to consult the dataset.

Finally, tied up with ethical considerations is that of the copyright ownership, by the interviewee, of the words they have spoken. Under the 1998 UK Copyrights, Designs and Patents Acts, copying work, such as citing large extracts in a research publication, is considered to be an infringement unless it is within the context of 'fair dealing'. For other future researchers to be able to fully exploit their data, it is essential for investigators to obtain from informants in writing a transfer of copyright, or permission to archive for other future research use (Qualidata 1998).

#### ***Representation, coverage and context of the research and fieldwork***

In the process of analysing and coding data, researchers use their own personal knowledge and experiences as tools to make sense of the material, which may be indefinable and cannot be easily documented. A pertinent question we must consider then, is whether data can be effectively used by someone who has not been involved in the original study? How much of the jigsaw can be missing yet leave the puzzle still worth attempting?

Mauthner et al (1998) argue that 'data are the product of the reflexive relationship between the researcher and researched, constrained and informed by biographical, historical political, theoretical and epistemological contingencies' (1998: 742). The researcher's own reflexivity enhances the raw data gathered and stimulates the formulation of new hypotheses in the field. Additionally, in-depth interviewing or observation requires a full advance preparation, with the researcher becoming familiar with the context of a respondent's community, generation, work, personal and family life. Some researchers thus believe that qualitative data cannot be used sensibly without the accumulated background knowledge and tacit understanding that the original investigator had acquired – understanding commonly not written down but held in the researcher's head. Thus the original context can never be truly reconstructed. The complexity, quirks, or lack of adequate documentation of data may thus be major difficulties in re-analysis, particularly when no input from the original investigating team is possible (Corti and Wright 2002). The loss of the essential contextual experience of 'being there' and the lack of being able to engage in reflexive interpretation may then be viewed as a barrier to re-use. Even when revisiting one's own data, the problem of loss of context may apply. Mauthner et al (1998) highlight how their own 'ability to interpret their own data may also decline over time as memories wane; changes in personal situation and new knowledge that they have gained since the primary study may also influence their re-interpretation of the data'.

However, the loss of context in archived data should not be seen as an insurmountable barrier to re-use. Indeed, there are very common and accepted instances where research data is used in a 'second hand' sense by investigators themselves. For example, principal investigators writing up the final analysis may not have been directly engaged in fieldwork, having employed research staff to collect data. Similarly, for those working in research teams sharing one's own experiences of the fieldwork and its context are essential, but never total. In both instances, the analysers or authors must rely on fieldworkers and co-workers documenting detailed notes about the project and communicating them. However, documentation of the research process can help recover a degree of context, and whilst it cannot compete with 'being there', field notes, letters and memos documenting the research can serve to help aid the original fieldwork experience. Audio-visual recordings of interviews can also enhance the capacity to re-use data without having actually collected them.

A further issue concerning representation of the interview is the nature or method of transcription. Ways of transcribing interviews vary enormously between disciplines and individuals. As such, transcriptions are usually a subjective interpretation of the real-life original. While sociologists typically want to capture the words, conversation analysts and socio-linguistics are more concerned with documenting the para-linguistic features of speech, such as pauses, laughter, tears and so on. In terms of maximising the potential of a qualitative dataset for re-use then, the ideal scenario is to retain original audio recordings. Jacobsson, in this volume, demonstrates the problem of transcription very clearly in her description of revisiting her own collection. Her original interview transcriptions were highly edited, so that in order to pursue her new line of enquiry, she needed to go back and re-transcribe parts of the data – to consider the nuances of hesitancy, embarrassment, and defensiveness previously unobserved. This illustrates demonstrates the value of retaining audio material for archival collections of qualitative data.

Finally, the parts of a collection that find their way into an archive, either personal or public, may not represent the original collection in its entirety. Research are vulnerable to erosion or fragmentation, both from natural causes – like accidental loss or damage - or man-made reasons such as a policy of deliberate weeding or disposal for lack of sufficient storage space. Particularly sensitive interviews may have been destroyed, sometimes at the behest of the interviewee. In short, material is judged to be worthy of preservation by the originator as well as the archivist. It is therefore important for archivists to document, where possible, what data is missing and why.

Hammersley summarises these points concisely: 'the data collected by different researchers will be structured by various purposes and conceptions of what is relevant. As a result, users of archives are likely to find that some of the data or information required for their purposes is not available' (Hammersley, 1997, p. 139). Hence re-users of data will need to use their own judgement

in assessing the quality of the material. Nor should we forget that the practice of research in other disciplines, such as social history, is fully based on interpreting evidence created by other witnesses.

### ***Unfamiliarity with the method***

Because most qualitative researchers are unused to consulting data collected by other people, their concept of 'secondary analysis' is still typically associated with the 'number crunching' of survey datasets. By contrast the positive experiences of researchers who have gone back to their own data or have embarked upon new investigations on archived data testify how fruitful the effort can be (Fielding and Fielding 2000; Akerstrom and Jacobsson in this volume). We would therefore assume that with time the inhibition of unfamiliarity with re-analysis will become much less relevant than it is today.

### ***Lack of infrastructure for data-sharing***

One of the barriers to re-using data in the past, and indeed in many countries still, has been the lack of an infrastructure to enable access to the rich research data collected in the academic community. While in some disciplines professional networks enable high quality data to be shared on an informal basis, the preservation of these sources is not ensured, not does it allow equal access for all researchers. The growing establishment of national data-sharing policies is helping to secure data, provide access and provide support for re-use, although the stock of data archived still rests willing to share by the original investigators.

Documenting data to the high standard required to render it easily accessible, however, can be a huge task, particularly without specific funding to support this. Investigators who have constructed large-scale or perhaps long-running studies may be daunted by the prospect of transforming their data into a widely usable resource. If qualitative data are to be shared, the infrastructure needs to be in place to offer guidance, support and adequate funding from the start of the project to enable the documentation and archiving of its data.

The question of public investment in archiving is an issue that is often mooted by research investigators. A significant tension between investment in data preparation for archiving and new data collection and primary research, in a world in which scientific budgets are limited, must be recognised. Indeed where a prospective data policy exists, such as for the Economic and Social Research Council in the UK, and the National Science Foundation in the U.S., resources for archiving can compete with those for the new research. Many researchers see the need for a policy to determine relative levels of investment in data preparation and documentation according to different types of datasets and for different end uses. For example, for large-scale or longitudinal studies the not insignificant costs of good project housekeeping, high-quality data documentation, anonymisation, and sample maintenance are valued as good investments. With smaller studies, however, most of the needs for making future archiving possible can also be viewed as useful for the research itself, as a form of structuring and good housekeeping.

### ***Misinterpretation***

Concerns about misinterpretation of data may arise from fear of selective and opportunistic interpretation in re-analysis. Researchers conducting longitudinal programmes feel particularly vulnerable due to the potentially negative impact of bad publicity on sample attrition. Finally, there is a concern surrounding the mis-use of politically-sensitive occupational or environmental data by pressure groups, industry, or the press (Corti and Wright 2002). While journalistic misinterpretation is harder to control, most instances have in practice ensued from publication by the original researchers rather than from archiving. Straightforward misinterpretation by other academics is not often a problem. Much more typical, and usually after a considerable time lag, is a serious critique of the original study. But this is precisely how intellectual understanding advances. And paradoxically, at the

same time it may give the original study new attention and esteem, both because of the controversy generated, and because the criticism itself recognises the original as a worthwhile target.

### *Threat to intellectual property rights*

Some researchers have voiced concern over the loss of their control over data or their intellectual property rights when data become publicly archived. Many wish to complete working on the data before it can be offered to the wider community. In the medical field, there is a perception that investigators may not get their names on secondary research papers as collaborators if the datasets were accessed through an independent archive (Corti and Wright 2002). Keeping data in-house and restricting access is thus one way of preserving intellectual capital.

Other researchers consider their data as private property, and seem almost bonded to their own ethnographic fieldwork notebooks or interviews. Anthropologists may have built a career around studying one particular remote region, and the data generated over the course of a professional career will be seen as unique, as a stock of intellectual capital that he or she can exploit through their lifetime.

### **4. Personal reflections of re-using data: an account by Paul Thompson**

Having put forward a number of approaches to re-using existing qualitative data, and discussed the potential benefits of and difficulties in re-analysis, in this section we observe the reflections of an experienced researcher who has re-used data. Did he re-use data to answer substantive questions or for methodological purposes? How useful were the data sources he drew upon? Were they rich enough for him to pursue his line of enquiry? Following Thompson's testimony, we assess these issues.

I have worked in sociology for over thirty years, but my first research in the early 1960s was as a social historian, and I still keep some of the influence of that early experience. Historians were and are still essentially jackdaws, scavengers: they use other people's detritus rather than create their own data. At that time historians got no training of any kind in research methods, but two things were always clear. First, you should be willing to make the best use you could of whatever sources you could find, whether in public archives, or in business offices, or in private house attics or sheds. And second, you should search for unknown new sources, for finding a significant new source was the biggest scoop you could make.

My thesis was about the labour movement in London, and I was also writing about architectural history as a sideline. I certainly had my own high moments of discovery. I still have vivid memories of an old socialist bringing out precious minute books from his garden shed in a North London suburb; or of cold winter days in the grimy unlit cellar of a Birmingham glassworks, reading packets of letters from the 1850s, unopened for a century, in the fading daylight. One certainly had to be prepared to suffer in the cause, of course much less drastically than an anthropologist, but I do remember eating squashed flies in a cut bread sandwich offered to me by a Yorkshire vicar whose son was keeping a chicken farm on the glebe (the parish priest's land) next to the vicarage—and it paid off, because within an hour he had lent me another precious bundle of forgotten correspondence.

As it turned out, one of the best sets of already archived sources for my labour movement thesis turned out to be in the library of the London School of Economics. The core was the deposit by LSE's founders, Sidney and Beatrice Webb, pioneers of British social science, of their own papers covering their lifetime's personal activity, politics and research work. Beatrice Webb had the foresight to keep her handwritten notes of the large number of interviews which she carried out with union leaders from the 1880s for their classic pioneering study of British trade unionism (Webb 1894, 1920). These notes provided me with unique information, by far the best source on the earlier development of trade unionism in the capital. I did not realize the full significance of this experience at the time, or recognize how the systematic quality of this material and the thoroughness of its scope were

consequences of its being research data: for me, it was simply a very good source. But I did not forget Beatrice Webb's example, and it was one of the early influences which set me on the path to Qualidata.

### ***First Experiences of Archiving and Re-using Interviews: Early Twentieth Century Family Relationships***

Most of my research in the last thirty years has focussed on family relationships, how they change, and how they relate to the economy and the community. Soon after I came to Essex, initially as a social historian, Peter Townsend, the first professor of sociology at Essex, with whom I taught a course on social policy and social change, suggested to me that it might be worthwhile to interview older people for my research. And to help convince me, he showed me bunches of extracts from the interviews he had recorded himself for his first major book, *The Family Life of Old People* (Townsend 1957). I remember in particular the vivid material on changing funeral customs. So here was a second experience of re-use. And it is good to know that those same interviews, along with the whole data from Peter Townsend's lifetime of research, very likely the most in-depth documentation which will ever be collected of the conditions and experience of old age and poverty in Britain, a unique and unrepeatable set of qualitative research material, have now been archived by Qualidata in the National Archive of Social Policy and Social Change at the University of Essex.

Initially, however, the influence of these interviews led me towards my first important experience of creating my own research material through interviewing. Commissioned to write a social history of Britain in the early 20th century, I wanted to write about the themes which a sociologist would have then highlighted for a portrait of contemporary Britain, including gender issues, childhood, youth culture, the informal culture of work and family leisure, sexuality and marriage. As a result of receiving a generous grant from the new Social Science Research Council in 1970, our research team were able to carry out some 450 interviews right across Britain with a quota sample of men and women born between the 1870s and 1906.

These interviews provided the basis for a large part of my book *The Edwardians* (Thompson 1975). In practice it proved impossible to use the whole set, because there was much too much information to absorb. Nor was there appropriate software to help us: we sorted out the transcripts into themes using scissors and staples. I used the interviews not only in detail for cross-analysis, for example of childrearing practises such as punishment by parents, but also as a set of in-depth family portraits to illustrate the whole range of the social spectrum, from a landowning family at one extreme to a Welsh West Indian boarding house or an impoverished London labourer's home at the other; and two final portraits of women, a London socialist and a struggling Stoke pottery worker, as examples of the contrasting dynamics of social change, conscious and unconscious.

We developed our interviewing method for this project initially from sociological influences. Right from the start we always used an interview guide, but equally importantly, we used it with a light touch, flexibly, giving as much space as possible to the informant to talk freely. It was only later on that we learnt that a comparable practice using the memories of the old—but without samples or interview guides, and then mainly focussing on "great men"—had evolved in the USA under the title "oral history". But by the later 1970s there had sprung up not only an international oral history movement, but also a parallel movement among life story sociologists launched by Daniel Bertaux from France. This first project for *The Edwardians* was closer to oral history, in that we only collected detailed evidence from informants about their lives up to 1918, which was the period I needed for my book. In retrospect, I feel that this was a serious mistake, because the potential value of the interviews, for example for studying social mobility or intergenerational influences, would have been far greater if they had also covered the missing fifty years up to the present. It would have been better to have done fewer but fuller interviews. All my later projects have used full life story interviews, and I believe that this is the form not only most likely to reveal insights to the original researcher, but also for subsequent users.

It was however clear to us from very early on that we were collecting a unique set of interviews of great potential value to others, and so we set up an informal archive in the department. The interviews were read over the next twenty years by large numbers of both students and outside researchers, and the result was a far larger number of publications by others than we could have ever achieved on our own—at least five times the output from the research. These include, for example, Standish Meacham's, *A Life Apart*, David Crouch and Colin Wards' *The Allotment* on working class culture; substantial parts of Charles More's, *Skill and the English Working Class* and Michael Child's, *Labour's Apprentices* on work; and John Gillis, *For Better, For Worse* on marriage, as well as articles on class by Patrick Joyce Richard Trainor, on social mobility by David Vincent, and on women and the family by Ellen Ross.

Although in recent years the most important users have been other researchers, I have also had the experience of re-using my own material from a new perspective for a later project. This was the research on the experience of ageing, funded as part of the Economic and Social Research Council's Ageing Initiative which I carried out in the late 1980s, leading to the publication of *I Don't Feel Old* (Thompson, Itzen & Abendstern 2000). We also collected a substantial new dataset for this, interviewing up to three generations in a national sample of "The Hundred Families" (Thompson and Bertaux 1993), a very rich source which, to my surprise, has not attracted re-users in the way that the set for *The Edwardians* does, although providing just the same type of information for the period since 1920. However, combining information from the earlier dataset with the new evidence undoubtedly gave us added strengths.

A key issue for the new project was how relationships between grandparents and grandchildren were changing over the longer timespan. So the first advantage of the older material was that it went back twenty or more years further, with recollections of grandparents by informants born as far back as the 1880s. This meant that I was able to write a whole chapter in *I Don't Feel Old*, "At the Edge of Living Memory", on the experience of ageing in an earlier generation, through the point of view of grandchildren. A second advantage was more accidental. When I was researching for *The Edwardians*, I had not been interested in grandparenting as such, so that the interview guide did not include any questions about this topic. But the consequence was that the material on grandparents which did appear in the interviews was spontaneous, because people just wanted to tell us about them. With the later research, by contrast, we deliberately drew them out. So we could compare how far the two forms of interview brought in similar material—and reassuringly, it was indeed broadly similar. We were also able to evaluate from the earlier interviews, I think more clearly because they were more spontaneous, for how many people was a grandparent really significant?

Being able to estimate the earlier significance of grandparents was one important gain from having an alternative dataset to re-use. We were also able to draw a good picture of older people's occupations and other sources of income in the late 19th century, a point on which contemporary statistics are not at all reliable. And we gained a lot of information too about mutual help, exchanges of help each way between the generations, the extent of influence of grandparents, and the varying ways in which grandchildren helped them in later life. None of this would have been possible if we had not kept the material, and we undoubtedly learnt a lot from it in new ways.

### ***Commentary on Paul Thompson's experience***

Like, Jacobsen's account of recycling data in this volume, Thompson's own experiences of attempting reanalysis of one's own data offer us firsthand insight into some of the real issues encountered.

In his early days as a young researcher, Thompson was heavily influenced by famous social research pioneers, such as the Webbs and Peter Townsend, both of whom had collected highly detailed notes from fieldwork. But his first experience of re-using research materials was for historical and

descriptive purposes, to provide information on the development of the London labour and socialist movement.

His desire to consult existing sources of data later on in his career arose from this first-hand experience of using archives as a social historian, and from the belief that rich data, such as the life story material he began to collect, should be made available to others.

Some years down the track, now as an experienced researcher, Thompson hoped to study social mobility and intergenerational influences using a set of interviews he had already collected himself. This hope was thwarted due to the fact that he did not have access to critical information - the latter half of the interviewees lives that he could have collected at the time of interviewing. This oversight caused him to review his own data collection strategy for subsequent studies, and advocated for all future studies, the use of full life story interviews. Here we see evidence of a fundamental development of methodology.

The second occurrence of re-use reported by Thompson concerned a new interest in a topic not specifically covered in the original set of interviews. Mentions of grandparenting, however, turned out to be richly sprinkled amongst the interviews. As a consequence the data could be re-analysed with this new focus in mind.

### **5. Qualitative data - its longevity, and what to preserve to ensure maximum re-use**

Creating a national stock of qualitative research resources requires that the collections acquired are suitable for informed use and meet the demand from researchers. National qualitative data archives such as Qualidata and the Murray Research Center have acquisitions policies to ensure that all materials deposited meet certain criteria: that data are documented to a minimum standard, are in appropriate formats, are complete, and that confidentiality, data protection and copyright issues have been addressed. Priorities must also be assigned to data, so that the inflow of data meets the resources available for processing. Potential studies are thus always evaluated from a long-range perspective to predict their future value. Priorities focus on:

- the historical value of the study
- data complementary to existing data holdings
- data that have further analytic potential than the original investigation, ie have not been exhaustively analyzed
- data based on large-scale national samples
- data which are longitudinal in design
- the possibility of further follow-up of the sample
- mixed methods data
- studies that include a wide range of measures

#### ***Data format***

The format and mark-up of data also determine the usefulness of a collection: for example, whether a study is available digitally, and how the text has been described and marked-up. Recommended strategies for the preparation, storage and dissemination of qualitative data are published elsewhere (Corti 2002). There is a debate about the long-term value of coded data, mainly because the coding process is subjective, often geared towards specific themes, and therefore may not be applicable to the secondary analyst's topic of investigation. For larger studies, however, there is a stronger case for retaining coded data, in order to aid searching within voluminous bodies of text. Indeed, the Edwardians Online project (Barker 2002) has followed the structure of the existing coding to provide navigation through the huge bulk of text – some 50,000 pages of interview transcript.

#### ***Conclusion***

In this chapter we set out to demonstrate how existing sources of qualitative data can be re-used. Firstly, this is because secondary analysis makes more effective use of material which is costly to collect; secondly, it enables further exploration of the data from a new perspective; thirdly, it enables comparative research to be carried out in a number of contexts (e.g. geographically, over time, cross-culturally); and last, it allows for verification of the original study.

In many ways these methods parallel those that are used and documented for the secondary analysis of survey data: comparative research, replication or restudy; asking new questions of old data; the strengthening of scientific inquiry through the open discussion of methods; help in new research designs; and providing resources for training in research and substantive learning.

We have shown that there are important gains to be made from re-analysis. At the start of a research project, it can be invaluable in providing a sense of the topics which can be successfully covered in interviewing, and therefore make the pilot stage of the new project both more effective and also much swifter. At a later stage a comparable interview set may also provide a crucial wider sample base for testing the interpretations which are emerging. Finally, by making research data available to re-analysis by others, the investigator may multiply the outcomes from this initial research through the publications of others from the same material.

Equally there are methodological and practical difficulties in re-using data, which include understanding the coverage and context of the research; ethical and consent considerations; unfamiliarity with the method; and the general lack of suitable data available.

Over the last five years we have witnessed a new culture of the secondary use of qualitative data, which has been largely borne out of data-sharing policies, such as in the UK. This new culture needs to be nurtured by acquiring relevant data and documenting and presenting it in user-friendly ways. Qualitative data services can help fulfil this role by encouraging a culture of sharing in research practice and enabling support; developing appropriate collection priorities, creating digital resources for teaching and research, and by offering support and outreach activities such as training. It is also significant from the experiences of two well established centres, the Mass Observation Archive in Britain and the Murray Research Center in the US, and also by the major ethnological archives in Scandinavia, that a particularly effective model is to combine archiving with in-house research on the data held: this generates both relevant acquisitions and a high level of use by researchers. Finally, looking into the future, innovative on-line data access and analysis tools are very likely to both encourage and facilitate the re-use of qualitative data.

#### NOTE

In 1999 Qualidata conducted a national survey of academics and researchers to ascertain what kind of data resources researchers wished for, receiving over 550 valid responses from a wide range of user communities. 92 per cent wanted to see datasets in a digital and accessible format that could be used for both research and teaching purposes. The highest demand was for qualitative data on health, criminology and social policy (Corti and Thompson, 2000).

#### REFERENCES

Allwood, J., Björnberg, M., Grönqvist, L., Ahlsén, E. and Ottesjö, C. (2000, December), 'The Spoken Language Corpus at the Department of Linguistics, Göteborg University', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [Online Journal], 1(3). Available at: <http://qualitative-research.net/fqs/fqs-eng.htm>

Barker, E. (2002), *The Edwardians Online Pilot Resource*, Qualidata, UK Data Archive.

- Booth, C. (1891-1902), *Life and Labour of the People in London*. London: Williams and Norgate; Macmillan
- Brunswick, A. F. (1980), 'Health stability and change: A study of urban black youth', *American Journal of Public Health*, 70, 504-513.
- Burgess, R.G. (Ed.) (1982), *Field Research: a Sourcebook and Field Manual*. London: Unwin Hyman.
- Charlesworth, J. and Fink, J. (2001), 'Historians and Social Science Research Data: the Peter Townsend Collection', *History Workshop Journal*, 51: 206 – 219.
- Cohen, S. (1971), *Folk Devils and Moral Panics, The Creation of the Mods and Rockers*. Oxford: Basil Blackwell.
- Cohen, S. and Taylor, L. (1972), *Psychological Survival: The Effects of Long-term Imprisonment*, London: Allen Lane.
- Columbia University (2002), Oral History Research Office, [On-line]  
<http://www.columbia.edu/cu/lweb/indiv/oral/>
- Corti, L. (2000, December), 'Progress and Problems of Preserving and Providing Access to Qualitative Data for Social Research - The International Picture of an Emerging Culture', *Forum Qualitative Social Research* [Online Journal], 1(3).
- Corti, L. (2002), *Qualitative Data Processing Guidelines*. Qualidata UK Data Archive.
- Corti, L. and Ahmad, N. (2000), 'Digitising and Providing Access to Social-Medical Case Records: The Case of George Brown's Works', *Forum Qualitative Social Research* [On-line Journal], 1(3).
- Corti, L., Day, A. and Backhouse, G. (2000), 'Confidentiality and Informed Consent: Issues for consideration in the preservation of and provision of access to qualitative data archives', *Forum Qualitative Social Research* [On-line Journal], 1(3).
- Corti, L., Foster, J. and Thompson, P. (1995), 'Archiving qualitative research data', *Social Research Update*, 10, Department of Sociology, University of Surrey. **Pages?**
- Corti, L. and Thompson, P. (2000), *Annual Report of Qualidata to the ESRC*. University of Essex.
- Corti, L. and Wright, M. (2002), *Consultants' Report to the Medical Research Council on the MRC Population Data Archiving and Access Project*, UK Data Archive, University of Essex.
- Courtney, C., and Thompson, P. (1996), *City Lives*. London: Methuen.
- Dein, A. and Perks, R. (eds) (1993), *Lives in Steel: Audio compilation*, London: British Library
- Elder, G.H. (1974), *Children of the Great Depression: Social Change in Life Experience*. Chicago: University of Chicago Press.
- ESRC (1999), *Guidelines on Copyright and Confidentiality: Legal issues for social science researchers*. ESRC: Swindon
- ESRC (2002), *ESRC Datasets Policy*. Swindon: ESRC.
- Fielding, N. and Fielding, J. (2000), 'Resistance and adaptation to criminal identity: using secondary analysis to evaluate classic studies of crime and deviance', *Sociology*, 34 (4): 671-689.

- Fink, A.S. (2000), 'The Role of the Researcher in Qualitative Research', *Forum: Qualitative Social Research* [On-line Journal], 1(3).
- Franz, C., McClelland, D., & Weinberger, J. (1991), 'Childhood antecedents of conventional social accomplishment in midlife adults: A 36-year prospective study', *Journal of Personality and Social Psychology*, 60(4), 586-595.
- Glueck, S., and Glueck, E. (1968), *Delinquents and nondelinquents in perspective*. Cambridge, MA: Harvard University Press.
- Goldthorpe, J., Lockwood, D., Bechhofer, F. and Platt, J. (1968), *The Affluent Worker: Industrial Attitudes and Behaviour*. Cambridge: Cambridge University Press.
- Hammersley, M. (1997), 'Qualitative data archiving: some reflections on its prospects and problems', *Sociology*, 31(1): 131-42.
- James, J. and Sorensen, A. (2000), 'Archiving Longitudinal Data for Future Research. Why Qualitative Data Add to a Study's Usefulness', *Forum Qualitative Social Research* [On-line Journal], 1(3).
- Kuula, A. (2000), 'Making Qualitative Research Material Reusable: Case in Finland' [28 paragraphs]. *Forum: Qualitative Social Research* [On-line Journal], 1(3).
- Kvale, S. (1996), *Interviews—An Introduction to Qualitative Research Interviewing*. Thousand Oaks: Sage.
- Kynaston, D. (2001), *The City of London, IV: A Club No More, 1945-2000*. London: Pimlico.
- Kynaston, D. (2002), 'Three taps on the nose', *National Life Story Collection Newsletter*, 8: 1-2
- Laub, J., & Sampson, R. (1998). Integrating Quantitative and Qualitative Data. In J. Giele & G. Elder Jr. (Eds.), *Methods of Life Course Research* (pp.213-230). Thousand Oaks, CA: Sage.
- Leh, A. (2000) 'Problems of Archiving Oral History Interviews. The Example of the Archive German Memory', *Forum: Qualitative Social Research* [Online Journal], 1(3).
- Lewis, O. (1963), *Tepoztlan Restudied: Life in a Mexican Village*. Urbana: University of Illinois Press
- Llewellyn Smith, H. (1930-5), *The New Survey of London Life and Labour*. London: P. S. King.
- Luder, M., Neukom, M. and Thomann, B. (2000), 'The JAKOB database: psychodynamic psychotherapy research at the University of Zurich', *Forum: Qualitative Social Research* [Online Journal], 1(3).
- Lux, Z. (2000), 'Computerized Support for Research and Publication in Contemporary History', *Forum: Qualitative Social Research* [Online Journal], 1(3)
- Lynd, R. S. and H. (1929), *Middletown: a Study in American Culture*. New York: Harcourt, Brace.
- Lynd, R. S. and H. (1937), *Middletown in Transition: a Study in Cultural Conflicts* New York: Harcourt, Brace.
- Malinowski, B. (1967), *A Diary in the Strictest Sense*. London: Routledge and Kegan Paul

- Marsden, D. and Jackson, B.(1962), *Education and the Working Class*.London: Routledge and Kegan Paul.
- Marwick, A. (1970), *The Nature of History*. London: Macmillan.
- Mauthner, N., Parry, O. and Backett-Milburn, K. (1998), 'The data are out there, or are they? Implications for archiving and revisiting qualitative data', *Sociology*, 32(4), 733-45.
- May, E. T. (1988), *Homeward bound: American families in the cold war era*. New York: Basic Books.
- Mruck, K., Corti, L., Kluge, S. & Opitz, D. (Eds) (2000), 'Text. Archive. Re-Analysis', *Forum: Qualitative Social Research* [Online Journal], 1(3), Available at: <http://qualitative-research.net/fqs/fqs-eng.htm>
- Muhr, T. (2000), 'Increasing the Reusability of Qualitative Data with XML', *Forum: Qualitative Social Research* [Online Journal], 1(3).
- Newson, J. and Newson, E. (1963), *Infant Care in an Urban Community*. London: G. Allen & Unwin.
- Newson, J. and Newson, E. (1968), *Four Years Old in an Urban Community*. London: G. Allen & Unwin.
- Newson, J. and Newson, E. (1976), *Seven Years Old in the Home Environment*, London: G. Allen & Unwin.
- Oral History Society (2002), *Oral History Society Ethical Guidelines*. Oral History Society: London
- Platt, J. (1996), *A History of Sociological Research Methods in America, 1920-1960*. Cambridge: Cambridge University Press.
- Plummer, K. (1983), *Documents of Life: an Introduction to the Problems and Literature of a Humanistic Method*. London: Unwin Hyman.
- Qualidata (1998), *Legal and Ethical Issues in Interviewing*, Qualidata Website [www.qualidata.essex.ac.uk](http://www.qualidata.essex.ac.uk)
- Qualidata (1999), *Teaching Pack, The Last Refuge: A qualitative research study by Peter Townsend*. University of Essex.
- Redfield, R. (1930), *Tepoztlan: A Mexican Village; a Study of Folklife*, Chicago: University of Chicago Press.
- Rex, J. and Moore, R. (1967), *Race, Community and Conflict*. Oxford: Oxford University Press.
- Rowntree, B.S. (1901), *Poverty, A Study of Town Life*. London.
- Sears, R., Maccoby, E., and Levin, H. (1957), *Patterns of Child Rearing*. Stanford, CA: Stanford University Press.
- Sheridan, D. (2000), 'Reviewing Mass-Observation: The Archive and its Researchers Thirty Years on', *Forum: Qualitative Social Research* [Online Journal], 1(3).
- Silverman, D. (2000), *Doing Qualitative Research: a Practical Handbook*. London: Sage.

- Stacey, M. (1974), 'The Myth of Community Studies', in Bell C & Newby H (eds) *The Sociology of Community*. London: Frank Cass: 13-26.
- Stewart, A. and Healy, J. (1989) 'Linking individual development and social changes' *American Psychologist*, 44, 1, 30-42.
- Strauss, A. L., and Corbin, J. (1990), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage
- Supple, C. and Perks, R. (eds) (1997), 'Voices of the Holocaust: cross-curricular resource pack, NLSC, London: British Library.
- Szabo, V. & Strang, V.R. (1997), Secondary analysis of qualitative data. *Advances in Nursing Science*, 20(2): 66-74.
- Terman, L. M. (1954) 'Scientists and non-scientists in a group of 800 gifted men', *Psychological Monographs: General and Applied*, 68(7), (Whole No. 378): 1-44.
- Thompson, P. (1975), *The Edwardians: the Remaking of British Society*. London: Weidenfeld and Nicolson.
- Thompson, P. (1991) *Pilot study of archiving qualitative data: Report to ESRC*, Department of Sociology, University of Essex.
- Thompson, P. (2000) *The Voice of the Past*. (Third revised edition; first edition 1978.) Oxford: Oxford University Press.
- Thompson, P., and Bertaux, D. (eds) (1993), *Between Generations*. Oxford: Oxford University Press.
- Thompson, P. and Corti, L. (1998) 'Are you sitting on your qualitative data? Qualidata's mission', *Social Research Methodology: Theory and Practice*, 1, 1: 85-90.
- Thompson, P., Itzin, C. and Abendstern, M. (1990) *I Don't Feel Old: The Experience of Ageing*. Oxford: Oxford University Press.
- Townsend, P. (1957), *The Family Life of Old People*. London: Routledge.
- Townsend, P.(1962), *The Last Refuge, A Survey of Residential Institutions and Homes for the Aged in England and Wales*. London: Routledge.
- Townsend, P. (1979), *Poverty in the United Kingdom, a Survey of Household Resources and Standards of Living*. London: Penguin Books.
- UK Data Archive (2002a), *UK Data Archive Preservation Policy*. UK Data Archive, University of Essex.
- Webb, S. and Webb, B. (1894, revised edition 1920), *History of Trade Unionism*. London: Longmans, Green.
- Zeitlyn, D. (2000, December) 'Archiving Anthropology' *Forum: Qualitative Social Research* [Online Journal], 1(3).