The appliance of science
Uses of technology in social scientific theory and practice

Submitted for the Degree of Ph.D.
December 2004

MACC
Una McCormack

Department of Sociology
University of Surrey
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>IV</td>
</tr>
<tr>
<td>GLOSSARY</td>
<td>V</td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1 Research overview</td>
<td>1</td>
</tr>
<tr>
<td>2 Overview of chapters</td>
<td>4</td>
</tr>
<tr>
<td>3 Summary</td>
<td>7</td>
</tr>
<tr>
<td>2 THEORETICAL AND METHODOLOGICAL APPROACH</td>
<td>8</td>
</tr>
<tr>
<td>1 Epistemology in the social sciences</td>
<td>8</td>
</tr>
<tr>
<td>Introduction</td>
<td>8</td>
</tr>
<tr>
<td>The hermeneutic tradition in philosophy and social science</td>
<td>9</td>
</tr>
<tr>
<td>The sociology of knowledge</td>
<td>19</td>
</tr>
<tr>
<td>2 Hermeneutics, method, and the 'problem' of validity</td>
<td>24</td>
</tr>
<tr>
<td>Introduction</td>
<td>24</td>
</tr>
<tr>
<td>The qualitative/quantitative debate and the use of case study methodology</td>
<td>25</td>
</tr>
<tr>
<td>Validity and the hermeneutic process</td>
<td>30</td>
</tr>
<tr>
<td>3 TECHNOLOGY AND SOCIETY</td>
<td>39</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>39</td>
</tr>
<tr>
<td>2 Determinism in accounts of the relationship between technology and society</td>
<td>40</td>
</tr>
<tr>
<td>The relativist critique of determinism</td>
<td>40</td>
</tr>
<tr>
<td>Theories of the information society</td>
<td>42</td>
</tr>
<tr>
<td>3 Social influences on technology</td>
<td>56</td>
</tr>
<tr>
<td>Introduction</td>
<td>56</td>
</tr>
<tr>
<td>The social shaping of technology</td>
<td>58</td>
</tr>
<tr>
<td>The social construction of technology</td>
<td>61</td>
</tr>
<tr>
<td>The soft option</td>
<td>66</td>
</tr>
<tr>
<td>4 The development and reception of hypertext</td>
<td>70</td>
</tr>
<tr>
<td>Introduction</td>
<td>70</td>
</tr>
<tr>
<td>History of hypertext and the World Wide Web</td>
<td>71</td>
</tr>
<tr>
<td>Hype and hypertext: the reception of hypertext by the humanities and social science</td>
<td>73</td>
</tr>
<tr>
<td>5 Summary</td>
<td>83</td>
</tr>
<tr>
<td>4 USES OF TECHNOLOGY IN SOCIAL SURVEY RESEARCH</td>
<td>85</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>85</td>
</tr>
<tr>
<td>2 Government social survey research in the UK</td>
<td>86</td>
</tr>
<tr>
<td>Historical and organizational overview</td>
<td>86</td>
</tr>
</tbody>
</table>
Acknowledgements

At the University of Surrey, thank you to my academic supervisors Martin Bulmer and Geoff Cooper. Nigel Gilbert provided insights and assistance at a pivotal point in writing. At the DWP and DSS, thank you to my supervisors Steve Ellerd-Elliot and Jo Semmence. Andrew Fryer gave essential technical input and support.

Thank you to all friends and e-friends for their support and advice, especially: Tavia Chalcraft, Iain Coleman, Nicola Collie, Andria Laws, Alison Page, and Katharine Woods. Particularly grateful thanks to Paul Taylor, for friendship, support, and heroic proof-reading in three languages (twice).

Written with love and gratitude, for Matthew.
Glossary

ASD  *Analytical Service Division:* Former name for the IAD.

CAI  *Computer-assisted interviewing*

CAPI  *Computer-assisted personal interviewing:* A means of personal (face-to-face) interviewing in which responses are entered directly into a computer, usually a laptop.

CATI  *Computer-assisted telephone interviewing:* A method of telephone interviewing in which responses to answers are entered directly into a computer terminal.

FRS  *Family Resources Survey:* A large, complex multi-level survey which collects information on the incomes and circumstances of around 22-24,000 private households in Great Britain.

GUI  *Graphical user interface*

HCI  *Human-computer interaction:* A diverse field in computer science, drawing from various engineering fields such as computer graphics, operating systems, human factors, industrial engineering and cognitive psychology to investigate the design of interfaces between the user and the computer.

HTML  *Hypertext markup language:* An authoring language and distribution system developed for creating and sharing multimedia-enabled, integrated electronic documents over the Internet which, by means of hypertext linking, connects documents located anywhere in the world.

HTTP  *Hypertext transfer protocol:* The universal protocol devised by Tim Berners-Lee for moving files around the Internet; particularly suited to the Web's hypertext system.
IAD  Information and Analysis Directorate: The central core of analysts at the DWP. The section within IAD that manages the FRS is IAD Income Analysis 1.

ICTs  Information and communication technologies

NatCen  National Centre for Social Research: A non-commercial organization for designing, conducting and interpreting major social surveys, which holds the contract (with ONS) for implementing the FRS.

ONS  Office of National Statistics

OPCS  Office for Population, Censuses and Surveys

SAS  Statistical analysis software

SSK  Sociology of scientific knowledge

WWW  World Wide Web: The global web of linked files that can be retrieved using HTTP.
1 Introduction

1 Research overview

The research on which this thesis is based is a combination of a practical project and an analytical reflection upon that project. The project itself used hypertext technology (in this case, HTML)\(^1\) to develop questionnaire documentation for a large-scale government survey, the Family Resources Survey (FRS). The analytical work reflects upon the context in which the project emerged, the process of developing the application, the choices made during this process, and the reasons for these choices. The thesis as a whole is therefore a critical investigation of the implementation of a hypertext application.

The present study emerged from the introduction of the laptop computer into large-scale survey research, the consequent shift away from paper questionnaires to computer-assisted interviewing (CAI), and the ramifications of this for questionnaire documentation. Large-scale social survey research for public policy is increasingly being carried out by interviewers using laptop computers and CAPI (computer-assisted personal interviewing) programs (de Leeuw and Nicholls, 1996). Many national surveys are now CAPI-based, having initially been paper-and-pencil questionnaires, but the Family Resources Survey (FRS), with which this study is concerned, has been a CAPI-based survey since its inception in 1993. The FRS is a large and complex multi-level survey into the living standards and characteristics of around 22-24,000 households each year, and has been developed using the CAPI program BLAISE. The FRS is managed and sponsored by the Department for Work and Pensions (DWP) Information and Analysis Directorate Income Analysis 1 (IAD IA 1)\(^2\). The contract is currently held by the Office of National Statistics Social

---

\(^1\) The term *hypertext* describes a method of presenting information in which text, images, sounds, and actions become linked together in a complex, non-sequential web of associations that permits the user to browse through related topics, regardless of their presented order. HTML, or *hypertext markup language*, is an authoring language and distribution system developed for creating and sharing electronic documents over the Internet which, by means of hypertext, connects documents located anywhere in the world.

\(^2\) The DWP was formed by a merger of the former Department of Social Security (DSS) and parts of the former Department for Education and Employment after the general election in 2001. The department which managed the FRS within the DSS was ASD3E; following the Division's reorganization this was renamed ASD IA 1; following a later reorganization in November 2002, this became the Information and Analysis Directorate, so IAD IA 1. I refer to both the DWP and DSS where appropriate; for example, when discussing the inception of the FRS, I refer to the DSS.
Survey Division and the National Centre for Social Research (formerly Social and Community Planning Research: a non-commercial organization for designing, conducting and interpreting major social surveys), and has been since the first full survey year (1993-1994).

Traditionally in survey research, a key document for the survey researcher and user has been the paper questionnaire, which contained clearly set out information on question wordings, response categories, routing, checks made by interviewers, and so on. The paper questionnaire has now been replaced by the electronic questionnaire; for the FRS, written in BLAISE code. This is viewed by the interviewer on the laptop by a series of screens. These questions are not all necessarily visible, but are written into the program, to ensure that only the right people are asked the right questions. As a result, the paper questionnaire has been replaced by a relatively raw printout of the questions and the routing instructions in BLAISE code. The latest version of the FRS 'questionnaire', produced by one of the survey agencies, is over 1000 pages long, and is not regarded as particularly user-friendly. In order to investigate the origin and evolution of the application under development, I document the reasons for the introduction of laptops into the survey process, the benefits which this transition has brought and the problems which have emerged (as they are perceived in the literature in this field), particularly in relation to questionnaire documentation.

The research project contributed to providing full electronic documentation of the Family Resources Survey for the DWP intranet, and potentially in future for the World Wide Web³. It is therefore a case study of the implementation of a hypertext application which uses HTML to produce online documentation. The main work of the project was the development of a website containing documentation of the 1998-1999 Family Resources Survey questionnaire and associated information (e.g. metadata - i.e. data about data - related to variables on the questionnaire, and information on the structure of the questionnaire and the dataset). The documentation which was built during this project is presented on a CD-ROM included with this thesis, and the reader is encouraged to examine this. The CD-ROM consists of documentation in two distinct sections (there is also a help file which relates to the whole site):

³ The project was a CASE studentship funded by ESRC and the DWP.
• General documentation which holds for each survey year (e.g. background to the survey, structure of the database).

• Version-specific documentation for the 1998-1999 FRS. This consists in turn of documentation of three distinct areas: the questionnaire, the database, and the metadata. There are also three search facilities for using this information.

Throughout the text of this thesis, file names related to this CD-ROM are given in square brackets and blue text, directing the reader towards example pages from the site relevant to the points under discussion in the text. A summary of the contents of this CD-ROM is provided in Appendix 1. All material on the site is Crown copyright.

The project was also concerned with investigating the social nature of the innovation process, the choices made during this process, and the reasons for these choices. These included the expectations of the users; the institutional setting in which the project is embedded; the nature of the material to be presented; and the broader academic and professional community in which the material would be assessed. As part of my research, I was on site at the DWP one day a week for the bulk of the project. This was primarily for access to the documentation and to computing resources; in addition, however, I met frequently with members of the FRS team and with computing staff to discuss site content and structure; and, at the end of the project, I conducted interviews with users of the online documentation both from the FRS team and from other sections of IAD (the division in the DWP of which the FRS team is a part). For the duration of the project, I had full access to DWP facilities and was effectively a part-time staff member rather than a visitor.

Drawing upon the relativist and hermeneutic tradition within social science, I use my project as a case study in order to examine claims to 'better' and 'valid' knowledge in social science, and the rhetorical use of technology in these claims. I am particularly concerned with the rhetorical use of those technologies significant to my own project, i.e. CAI and hypertext. I begin by discussing accounts of the relationship between technology and society, and I review the literature on the development of hypertext. Theoretical accounts of hypertext have, in the main part, emerged from the humanities and qualitative social science, and I examine the rhetorical uses of hypertext in this literature. When I turn to look at quantitative social survey research -
specifically, the case of my project to produce online documentation for the FRS - I show how my use of HTML for the questionnaire documentation took place within the context of a general shift towards CAI in survey research, and also according to specific government policy initiatives concerning e-Strategy as it relates to government practice and publications. I describe the computerization of the survey research process, with particular reference to the departments concerned with the FRS, and I describe government e-Strategy initiatives related to the dissemination of government statistics.

Having described in detail my project to produce online documentation for the FRS, I reflect upon the nature of this intervention. I draw upon both hermeneutic and actor-network perspectives to consider my project in terms of it being the writing of a bureaucratic text. I look at the social process of my project in terms of its collaborative and anonymous authorship; discuss 'influences' on the text, including academic accounts of the information age, and the models of hypertext that appear in the literature; and I examine my own role as 'translator' of the project at key points. I reflect, in particular, upon Law's (2000) discussion of writing academic accounts of bureaucratic processes, and I argue that Law's emphasis on 'collusion' with bureaucratic actors is not the only way of writing such an academic narrative or giving an academic account of working within a bureaucratic context.

In this introduction, I have given an overview of the practical project which I conducted, and the theoretical and analytical concerns which inform the thesis. I now outline the contents of the chapters which follow.

2 Overview of chapters

Chapter 2, *Theoretical and methodological approach*, is organized into two sections. The first section outlines the theoretical orientation of the research, which is located in a relativist and hermeneutic tradition in social science. The theme of the enduring tension between natural and social scientific knowledge is introduced; I discuss the persistence of the idea that natural scientific knowledge is a 'special case', and the issue of whether the methods of natural science are appropriate for the study of social phenomena. (These issues are germane to later discussion of how hypertext has been received in the social sciences.) I provide an overview of the development of the
sociology of knowledge and the emergence of the sociology of scientific knowledge (SSK) as a sub-discipline, and the theoretical basis which SSK provides for the study of the social construction of technology. The second section of this chapter is concerned with issues related to methodology. Conceiving of the research as a case study, I reflect upon some of the assumptions which underlie case study methodology and, by extension, consider the qualitative/quantitative divide in social research. I shall suggest that a number of strategies have been used to make claims about the truth status of knowledge acquired through qualitative research; however, what these strategies have in common is a commitment to both a realist ontology and also a conception of social science as a teleological project aimed at the advancement of social science and the production of 'better' knowledge. I outline a hermeneutic methodology drawing on Gadamer's (1992) emphasis on the 'giving of accounts'; Vattimo's (1992) discussion of Gadamer also assists us in understanding Gadamer's conceptualization of hermeneutics as being a 'matter of proclaiming extra-methodic experiences of truth' (Vattimo, 1992: 115). I address directly the issue of validity in such an approach.

Chapter 3, Technology and society, is organized into three sections. I begin with a theoretical critique of technological determinism and its converse social determinism, arguing that these ostensibly conflicting accounts of the relationship between technology and society share a fundamental realist ontology. I move on to a more specific critique of the technological determinism which underlies contemporary theories of the 'information society'. Two main themes developed here are the tendency of accounts of technological change to describe societies in terms of shifts between epochs, and the way in which technologically deterministic accounts of society shift discussion of new technology to competing and ultimately irreconcilable utopian and dystopian visions of the future. I then turn to literature which emphasizes social influences on technology. This serves a number of purposes in my argument beyond a review and critique of this literature, since I draw upon this literature in various ways to inform my own account. I draw out the connections between some of the work conducted in this area and theoretical work done in the field of SSK. I draw a loose conceptual distinction between the social shaping and social construction of technology in order to organize my discussion. Finally, I bring together the themes emerging from earlier discussions to consider hypertext. An account of the emergence of hypertext and the Internet is given, and a critique of theories of hypertext
developed, showing how they have reflected the twin themes of epochal social change and utopian/dystopian visions of the future. The reception of hypertext by the humanities and qualitative social science is discussed; I review and critique examples of taxonomies of hypertext applications which appear in the literature and which became significant during the development of the FRS questionnaire documentation, as described in chapters 6 and 7.

Chapter 4, *Uses of technology in social survey research*, focuses on hypertext, the Web, and the Internet in the context of qualitative social survey research; more specifically, I focus on the case of the Family Resources Survey. This chapter therefore provides a description of the context in which the current project emerged, placing the FRS in context as the first CAPI-based UK government survey. I give an historical and organizational overview of survey research in UK government; I describe the computerization of the survey research process; I give an overview of UK government policy initiatives concerning e-Strategy as it relates to government practice and publications, and with particular reference to those departments concerned with the FRS.

Chapter 5, *The Family Resources Survey: a CAPI-based survey*, gives a detailed description of the FRS, with particular reference to the 1998-1999 survey, with which my project was concerned. I give information on the aims, content, and principal bodies which conduct the FRS, and provide an overview of the methodology, the types of data collected, the unit of analysis, and the users of the FRS. I give detailed descriptions of the questionnaire, the dataset, and the questionnaire documentation which was in existence prior to my own project being conducted.

Chapter 6, *Online documentation for the FRS: a case study of a hypertext application*, is a description of the development work I carried out to provide online documentation for the Family Resources Survey. I provide an overview of the development process, outlining the stages of the process. I give a detailed overview of the completed documentation, which covers general survey documentation, and version-specific (year-by-year) documentation of the questionnaire and the metadata. I give a detailed account of the development work: determining existing documentation and outlining design principles for the site; designing and testing pilot documentation; designing search facilities; designing and implementing full documentation; and acquiring and implementing user feedback.
Chapter 7, *Technology as text: narrative and authorship*, and chapter 8, *Translating technology*, reflect in greater detail upon the social process of the project to develop online documentation for the FRS. I draw upon both hermeneutic and actor-network perspectives to consider the development project as the writing of a bureaucratic text. The two chapters are organized around my loose conceptual distinction between the social shaping and social construction of technology: chapter 7 is concerned with issues related to the authorship of the site and with influences upon it; chapter 8 examines my role as 'translator' of the documentation website at various points in the development. Chapter 9 draws this thesis to a conclusion.

3 Summary

The purposes of this chapter have been to provide an overview of the thesis: of the practical project which I conducted and the analytical work I carry out in the remainder of this thesis; and to outline the chapters which follow. I turn now to give a full account of the theoretical orientation of the research.
2 Theoretical and methodological approach

1 Epistemology in the social sciences

Introduction

The social sciences no longer model themselves solely upon the causal and functional model of the natural sciences, and are open to the types of knowledge generated from qualitative research. Nonetheless, the distinction between *Geisteswissenschaften* and *Naturwissenschaften* - the social and the natural sciences - and the belief that the latter produces 'better' knowledge of the world often remain unexamined assumptions in contemporary academic discourse. The belief remains that the type of knowledge generated in the natural sciences is somehow 'better' than that produced in social science; that natural science is a 'special case' which generates superior knowledge of the world.

But where does this leave the knowledge generated by social scientific investigation? Are the natural sciences in fact the best model for the study of social phenomena? Is the knowledge produced in the natural sciences 'better' than that produced in the social sciences? Can the knowledge created in the social sciences be judged as equally 'valid'?

This study is concerned in part with the nature of knowledge in the social sciences, and the accounts social scientists use in order to legitimize the knowledge they produce as being 'better' or 'valid'. Of particular interest are the ways in which technology, implicitly positioned as the application of science, is drawn into these legitimizations in order to maintain the primacy of an objectivist account of knowledge in social science. This account I intend to critique. Given the contemporary application of computer technology to the research and analysis methods of social science, this research is thus concerned with the ways in which computer technologies, in particular CAI and hypertext, have been received by the social and human sciences, and how computer technologies have been used to provide support for particular programmes of research or theory within the social sciences. Again, these are accounts I intend to critique.

The distinction between the social and the natural sciences emerges with particular relevance for this study in the work of nineteenth-century idealist thinker...
Wilhelm Dilthey, who developed hermeneutics and attempted to apply Kantian method to the human sciences. A contemporary version - which attempts rather to sharpen the distinction between the 'two cultures' - appears in the work of Charles Taylor. As I discuss later in this chapter (pp. 12-14), the position taken on the natural/social distinction in this thesis is one which perceives the formulation of such as distinction as a rhetorical act, and which then asks - what purpose is served by it?

In the remainder of this section, I shall outline the theoretical orientation I have adopted for this research. I shall locate the current research within a hermeneutic and relativistic tradition in philosophy, what has been called the 'linguistic turn' in philosophy. To this end, I shall review a number of philosophical approaches towards epistemology (theories of knowledge) and outline their applicability for the present study. This overview thus provides the background against which I can introduce those relativistic traditions within sociology that provide the theoretical basis of the research which will be presented later in this chapter (the sociology of knowledge, and the social construction of science and technology).

**The hermeneutic tradition in philosophy and social science**

In this section, I shall discuss the character of the hermeneutic tradition in philosophy, and how this can be related to a study of bodies of knowledge. From Gadamer I shall take the idea of a sceptical, relativist hermeneutics as a model for conceptualizing bodies of knowledge. I shall touch briefly on the 'problem' of relativism, and then, drawing on Foucault and also from the new rhetorical tradition, I shall argue that a relativist hermeneutic orientation can have a critical edge.

The Weberian distinction between 'explaining' (*erklären*) and 'understanding' (*verstehen*) illustrates the difference between analytic and hermeneutic philosophy. Analytic philosophy and much social science see human beings and societies as belonging to a natural order - the purpose of philosophy and social science is to model the causes of social structure and process; moreover, a single method can serve for all sciences. In contrast, the hermeneutic tradition seeks to interpret or understand the social world. As Hollis (1994) writes:
'[the] central proposition [of hermeneutic social science] is that the social world must be understood from within, rather than explained from without. Instead of seeking the causes of behaviour, we are to seek the meaning of action' (pp. 16-17: my emphasis).

The modern tradition in hermeneutics derives from Dilthey, who identified 'meaning' as 'the category which is peculiar to life and to the historical world' (quoted in Hollis, 1994). Dilthey maintained a sharp distinction between the social and natural sciences: human life, he argued, could only be understood by reference to categories which do not apply to knowledge of the physical world, categories such as 'purpose', 'value', and 'ideal', i.e. aspects of 'meaning'. As Turner (1999) notes, Dilthey's orientation stemmed from a broader belief within German phenomenology that accepting natural science as the only legitimate model of scientific activity reduces values and meaning to no more than observable behaviour; to Dilthey and his contemporaries, this 'travesty' was seen as 'merely a further step in the colonization and subordination of the life of the mind to the impulses of matter' (p. 131).

Charles Taylor (2002) uses Gadamer's hermeneutic philosophy to maintain a formal break between the social and natural sciences (a break which will not be accepted at face value within this thesis). Taylor writes of:

'[A] big watershed in our intellectual world. There are those who hope to anchor an account of human nature below the level of culture, such that cultural variation, where it is not trivial and negligible, can be explained from this more basic account... And then there are those who find this account of human life unconvincing, who see it as an evasion of the most important explananda in human life, which are to be found at this level of cultural difference. Suffice it to say that Gadamer is one of the major theorists in the second camp... It is this bit of Gadamer's argument that often strikes philosophers and social scientists as scandalous, and "relativist", abandoning all allegiance to the truth. This interpretation is then supported by those among Gadamer's defenders who are in a "postmodern" frame of mind. But this grievously misunderstands the argument. Gadamer is anything but a "relativist" in the usual sense of today's polemics' (pp. 129-30).
The formal break, the watershed, which Taylor describes functions in his argument to maintain a domain for different but valid knowledge for the human sciences; i.e. in setting boundaries around a space within which the natural science model is not applicable. Geertz (2001) offers a response to Taylor which highlights the ramifications of maintaining such a break:

'Taylor's resistance to the intrusion of the "natural science model" into the human sciences seems in fact to accept his opponents' view that there is such a model, unitary, well-defined, and historically immobile... the problem is to confine it to its proper sphere... This division of the realm, which reminded one of nothing so much as the way some nineteenth-century divines (and some pious physicists) attempted to "solve" the religion versus science issues - "you can have the mechanisms, we can keep the meanings" - is supposed to ensure that ideas will not trespass where they don't belong. What it in fact ensures is symmetrical complacency and the deflation of ideas' (p. 156).

At the same time, Taylor is concerned to defend Gadamer from both the charge of relativism, and from adoption by relativists in 'the usual sense'. I would argue that, in doing this, Taylor steps back from the full consequences of the relativism in Gadamer's hermeneutics and, to some extent, I would also argue, there is in his argument an implicit conflation of relativism with nihilism; this is a mischaracterization which I shall go on to critique in detail (see below pp. 30-33). I interpret Gadamer's hermeneutics as being a relativist philosophy, and it is this which provides the philosophical basis of this thesis, as I now go on to discuss.

Gadamer's (1989) formulation of hermeneutic philosophy maintains Dilthey's distinction between types of knowledge; Gadamer argues that the appropriate model to invoke when attempting to understand social action is that of interpreting a text, and that one is not concerned with seeking causes or framing laws. Yet Gadamer's orientation is significantly different from traditional hermeneutics in that it is avowedly sceptical. Where traditional hermeneutics sought to understand texts, actions, and utterances objectively, in the implicit belief that a 'best' interpretation could be reached, Gadamer emphasizes instead the limitations of our own horizons, the prejudices which we inevitably bring to bear when we attempt to understand a
different form of life. Understanding comes from a 'fusion of horizons'; a partial rapprochement between our own world-view, from which we cannot be disconnected, and the different world-view we are seeking to comprehend. As Outhwaite (1985) summarizes:

'Our prejudices are not an obstacle to knowledge so much as a condition of knowledge, since they make up the fundamental structure of our relationship with our historical tradition' (p. 26).

Skinner (1985) points to the close associations which Gadamer's hermeneutic philosophy has with Wittgenstein's later philosophy, in its insistence that the meaning of any utterance is in its use, and that the understanding of any meaningful episode - whether an action or an utterance - always involves placing it within its appropriate 'form of life'. Winch's (1958) application of Wittgenstein's philosophy to social science leads him to a radical denial of the applicability of experimental method to the understanding of the social world; he argues that we should not presume that reality is independent of thought or that understanding reality necessarily means to explain how it works causally. Instead, 'our idea of what belongs to the realm of reality is given for us in the concepts which we use' (p. 15). The criteria by which we judge what is true or what is real within a realm of knowledge are not located outside that realm, but are embodied within it. As Hollis (1994) summarizes:

'Science embodies the key to the reality of a world of particles; religion embodies the key to the reality of a spiritual world. It is scientific practice to seek causes and religious practice to seek meaning. These practices, each being particular to its own form of life, are not in competition, since reality has no external or universal key' (p. 156).

There are two distinct ideas running through Winch's philosophy of social science. The first is a questioning of appeals to external criteria upon which one can base judgements of what one should believe, or do, or know. The second idea is, I believe, a misstep on Winch's part arising, unnecessarily, from this relativization of knowledge in social science - his rejection of the applicability of experimental method to the social sciences. This is a restatement of the 'two cultures' dichotomy which I
wish to critique. Winch's account does have ramifications for the nature of our belief in experimental method, and I shall expand upon this and other issues related to validity and method in greater detail in the latter part of this chapter. For the moment, I wish to focus upon the first aspect of Winch's philosophy (with which I do concur) - the theoretical consequences of relativizing our conceptions of bodies of knowledge. This form of radical relativism has been savagely criticized as *laissez-faire* and even reactionary: if we undermine the external basis on which to make judgements as to the validity of knowledge or forms of life, then, it is suggested, we are no longer in a position to criticize them - philosophy and social science become nothing more than description; they no longer ask difficult questions.

A relativist philosophy is not necessarily a politically neutralized philosophy; indeed, the questions which hermeneutic enquiry poses about the genesis and deployment of knowledge can be the source of its critical edge. When confronted by a claim about truth, it asks: whose purposes does this claim serve? What social systems produce bodies of knowledge and how do they claim that this knowledge is special and truthful? Such an orientation is derived from, amongst others, Foucault (1984):

'[T]ruth isn't outside power, or lacking in power... Truth is a thing of this world: it is produced only by virtue of multiple forms of constraint... Each society has its regime of truth, its "general politics" of truth: that is, the types of discourse which it accepts and makes function as true; the mechanisms and instances which enable one to distinguish true and false statements, the means by which each is sanctioned; the techniques and procedures accorded value in the acquisition of truth; the status of those who are charged with saying what counts as true... [W]hat must now be taken into account in the intellectual is not the "bearer of universal values." Rather, it's the person occupying a specific position - but whose specificity is linked, in a society like ours, to the general functioning of an apparatus of truth' (pp. 72-73).
Hunter (1999), too, adopts a relativist epistemology and, in addition, brings to bear the analytical tools of rhetoric. Firstly, rhetorical analysis assists with a conceptualization of truth as plural and diverse:

'Unlike discourse studies, which can overdetermine individuals into "subjects" because discourse works in closed systemic worlds analysing the negotiations that occur along the ideology-subject axis, rhetoric comprehends a range of stances...' (p. 145; my emphasis).

Hunter also provides a means of conceptualizing the 'mechanism' by which these different stances are expressed and negotiated:

'Rhetoric is a field concerned with the recognition of and negotiation among differences so that actions can be taken... it takes all communication to be a form of persuasion' (p. 145).

This contemporary version of an ancient tradition is an understanding of knowledge as potentially emerging from diverse sources, and which pays attention to the strategies of persuasion employed in claims to the status of truth for knowledge. I shall return to this in greater detail at the end of this chapter, in my discussion of validity and the hermeneutic process, and the 'giving of accounts'.

To summarize: the theoretical orientation adopted in this research is hermeneutic in that it seeks to understand and interpret meanings of individual and social action; relativist in its view of the possibility of access to multiple truths (including its own analysis); and critical in its attention to the rhetorical strategies

---

4 Edwards et al. (1995) go so far as to argue the case for the ethics of relativism: 'The advantage of relativistic notions of reality as rhetoric is that we can take positions and argue... Realism is no more secure than relativism in making sure that the good guys win, nor even of defining who the good guys are - except according to some specific realist assumptions that place such issues outside of argument... Realism is the rhetoric of no rhetoric, marshalled in favour of one claim against another... Far from ruling out the possibility of justification of a particular view, relativism insists upon it' (pp. 35-39; my emphasis). It is not that anything goes, rather that nothing goes. Relativism need not imply interests, power, etc., a point I shall address in more detail towards the end of this chapter, in my discussion of readings of Wittgenstein, in particular Pleasants' (1999) critique of critical social theory. Moreover, such a focus on language can, of course, be deterministic and reductionist in turn. Edwards et al. (1995) come, I think, close to this in the quotation cited.

5 As well as providing this thesis with an analytical cue through her use of rhetorical analysis, Hunter's (1999) work on hypertext will be discussed later in chapter 3, pp. 81-82, and was a significant influence on the development of the FRS documentation website (see chapter 7, pp. 209-210).
used when claims to special knowledge are made, including its own. It is also 
*generative*; I shall now go on to explain what I mean by this crucial characteristic of 
my theoretical orientation. I shall draw substantially on Gadamer's (1989) conception 
of experience as *Erfahrung*; that is, a notion of experience as something you undergo, 
when one's subjectivity is overcome and one is drawn into an event (see Weinsheimer 
and Marshall, 1989: xiii-xiv)\(^6\). I shall also make a distinction between what I perceive 
to be 'closed' and 'open' systems of knowledge.

A hermeneutic analysis of knowledge seeks to understand and interpret bodies 
of knowledge; in its relativist form, such an analysis is aware not only of the socially 
contingent nature of knowledge, but also of its own contingency: the analysis *itself* is 
generative of knowledge. Thus, rather than *closing down* ways in which we can 
conceptualize knowledge, it is *open-ended*. Gadamer writes (and this should be read 
with the conception of *Erfahrung* in mind):

"[A] person who is called experienced has become so not only *through* 
experiences, but is also open *to* new experiences. The consummation of 
his [sic] experience, the perfection that we call "being experienced" 
does not consist in the fact someone already knows everything and 
knows better than anyone else. Rather, the experienced person proves 
to be, on the contrary, someone who is radically undogmatic; who, 
because of the many experiences [s]he has had and the knowledge 
[s]he has drawn from them, is particularly well-equipped to have new 
experiences and to learn from them. The dialectic of experience has its 
proper fulfilment not in definitive knowledge but in the openness to 
experience that is made possible by experience itself" (1989: 355; his 
emphases)\(^7\).

\(^6\) Gadamer contrasts *Erfahrung*, a notion of experience which he applauds, with a more pejorative sense 
of the word 'experience', using the word *Erlebnis* to describe a sense of 'experience' as something you 
have rather than something you undergo. This sense is connected with a subject and with the Kantian 
subjectivization of aesthetics which the earlier part of *Truth and Method* challenges. See Weinsheimer 

\(^7\) Wittgenstein was, I think, saying much the same, when he wrote: 'The real discovery is the one that 
makes me capable of stopping doing philosophy when I want to - the one that gives philosophy peace, 
so that it is no longer tormented by questions which bring *itself* into question. Instead, we now 
demonstrate a method, by examples; and the series of questions can be broken off. Problems are solved 
(difficulties are eliminated), not a *single* problem. "But then we will never come to the end of our job!" 
Of course not, because it has no end. Philosophy solves, or rather gets rid of, only philosophical 
problems; it does not set out thinking on a more solid basis. What I am attacking above all is the idea
An example of what I understand as 'closing down' ways in which we can conceptualize knowledge would be the typology offered by Habermas of three forms of knowledge (see Ritzer, 1996a: 291): analytic science, which is aimed at technical control; humanistic knowledge, which is concerned with understanding the world; and critical knowledge, which aims for human emancipation. Hunter (1999), and feminist critique more generally, point to how such typologies 'close off' (or, in effect, delegitimize) forms of knowledge; where, for example, does feminist epistemology fit into such a schema as that offered by Habermas? As Hunter writes (bearing in mind her conceptualization of rhetoric as comprehending a 'range of stances'):

'While rhetoric has in the past been a place for recognising and negotiating among difference so that material actions as practices can occur, it has also been a place where descriptions about abstract and therefore fixed knowledge give guidance about how to be the same as other people... ' (p. 147).

Hunter is here showing a mechanism whereby a rigid system of classifying knowledge can function to exclude; for example, Habermas' rigid schema of knowledge runs a serious risk of excluding, amongst others, the possibility of feminist epistemology; its descriptions (its rhetorical actions) can close down this area of experience and of difference as a legitimate source of understanding the world. Habermas' schema, I would argue, does not comprehend a 'range of stances' that includes, for example, feminism and, indeed, other 'stances'. Eliding certain forms of epistemology, e.g. some forms of radical feminism or queer theory, under the all-purpose descriptor of 'critical knowledge' might well be imposing affinities upon these theories when those who espouse them are seeking rather to emphasize difference.

To demonstrate what I mean by a generative analysis, I now want to explore the ramifications of an account given by Kuhn of the pivotal moment in the development of his thinking. I draw on a lengthy quotation from Steven Weinberg, Nobel Prize Winner for Physics, from an article written for The New York Review of Books in 1998. In this article, Weinberg critiques Thomas Kuhn's The Structure of
Scientific Revolutions (1963), arguing against Kuhn's account of the shifts from one scientific paradigm to another. I offer this extensive quotation because it shows not only, on Kuhn's part, a hermeneutic understanding of an epistemology different from Kuhn's own, but also a profound failure on the part of Weinberg to grasp this most salient point of Kuhn's philosophy:

"What went wrong? What in Kuhn's life led him to his radical skepticism, to his strange view of the progress of science? Certainly not ignorance - he evidently understood many episodes in the history of physical science as well as anyone ever has. I picked up a clue to Kuhn's thinking the last time I saw him, at a ceremony in Padua in 1992 celebrating the 400th anniversary of the first lecture Galileo delivered in the University of Padua. Kuhn told how in 1947 as a young physics instructor at Harvard, studying Aristotle's work in physics, he had been wondering:

"How could [Aristotle's] characteristic talent have deserted him so systematically when he turned to the study of motion and mechanics? Equally, if his talents had deserted him, why had his writings in physics been taken so seriously for so many centuries after his death?... Suddenly the fragments in my head sorted themselves out in a new way, and fell into place altogether. My jaw dropped with surprise, for all at once Aristotle seemed a very good physicist indeed, but of a sort I'd never dreamed possible."

I asked Kuhn what he had suddenly understood about Aristotle. He didn't answer my question, but wrote to me to tell me again how important this experience was to him:

"What was altered by my own first reading of [Aristotle's writings on physics] was my understanding, not my evaluation, of what they achieved. And what made that
change an epiphany was the transformation it immediately effected in my understanding (again, not my evaluation) of the nature of scientific achievement, most immediately the achievements of Galileo and Newton."

Later, I read Kuhn's explanation in a 1977 article that, without becoming an Aristotelian physicist, he had for a moment learned to think like one, to think of motion as a change in the quality of an object that is like many other changes in quality rather than a state that can be studied in isolation. This apparently showed Kuhn how it is possible to adopt the point of view of any scientist one studies' (Weinberg, 1998: my emphases).

There is something almost engaging about Weinberg's bafflement throughout - and his failure to see that Kuhn had answered his question. One can see in Kuhn's account of his encounter with Aristotle a perfect example of Gadamer's conception of Erfahrung. Moreover, the significance of Kuhn's encounter with Aristotle was not simply in grasping the detail of outdated science, or in 'a mysterious communion of souls' (Gadamer, 1989: 292), but was in the integration of that encounter with his own point of view - that is, in the creation of a common meaning. Gadamer (1989) writes: 'understanding is to be thought of less as a subjective act than as participating in an event of tradition' (p. 290) and expands: 'when we try to understand a text, we do not try to transpose ourselves into the author's mind, but we try to transpose ourselves into the perspective within which he has formed his views' (p. 292). Kuhn experienced the tradition of the past in conjunction with his own contemporary tradition, and from that encounter emerged The Structure of Scientific Revolutions: the past and the present, the two horizons, fused into a new horizon. This is the value and the richness of hermeneutic philosophy for the study of the world, one which, I would argue, cannot be dismissed with a rhetorical flick as hollow as 'What went wrong?'

In the following section I shall apply this analytical orientation to a review of the study of bodies of knowledge within sociology. I shall trace the emergence of the field of the sociology of scientific knowledge (SSK) and its apparently radically relativist conception of science. Nonetheless, I shall suggest that, to some extent, some flavours of SSK have maintained an essentialist conception of knowledge, by
attempting to cast knowledge of the social world not in opposition to that of the physical world, but as superior. I shall thus suggest that SSK has, to some extent, failed to treat itself rhetorically, and I shall offer a relativist analysis of the division between social and natural scientific knowledge. In the final section of this chapter, I shall return to issues related to method which have been touched upon briefly in this earlier section.

The sociology of knowledge

The interpretative, hermeneutic tradition in sociology has been most developed both theoretically and empirically in the field of the sociology of knowledge. Mulkay (1992) describes the sociology of knowledge as being concerned with 'how bodies of thought and knowledge are influenced by the social and cultural contexts in which they are produced' (p. 1). The perspective emerges from the work of Karl Mannheim (1936) on systems of thought such as conservatism; Ritzer (1996a) describes Mannheim as being 'almost single-handedly responsible' for the creation of the field of the sociology of knowledge.

Mannheim saw the sociology of knowledge as both a method of historical-sociological analysis, and also as a distinctive theoretical perspective. As a theory, it was concerned primarily with 'the ways in which social relationships... influence thought' (1936: 267) and, by analyzing these relationships head on, removing barriers to 'good' knowledge:

'The sociology of knowledge has set itself the task of solving the problem of the social conditioning of knowledge by boldly recognizing these relations and drawing them into the horizon of science itself and using them as checks on the conclusions of our research' (p. 265).

Mannheim acknowledges the philosophical implications of such an approach, that it might 'pass... into an epistemological inquiry concerned with the bearing of this interrelationship upon the problem of inquiry' (p. 267). Mannheim's own concern was to avoid a 'sterile form of relativism' (p. 264) and he devotes significant space in his essay on the sociology of knowledge to countering charges of epistemological
relativism (pp. 281-285). He rejects the idea that providing an analysis of the historical emergence and social context of a body of thought is necessarily a refutation of its validity (p. 283) or, conversely, that it tells us nothing concerning the value of the assertion (on the grounds that the manner in which a statement originates says nothing logically about its validity). He argues instead that:

'Mere factual demonstration and identification of the social position of the assertor... tells us nothing about the truth-value of his [sic] assertion. It implies only the suspicion that this assertion might represent merely a partial view... It would be incorrect to regard the sociology of knowledge as giving no more than a description of the actual conditions under which an assertion arises. Every complete and thorough sociological analysis delimits, in content as well as structure, the view to be analyzed. In other words, it attempts not merely to establish the existence of the relationship, but at the same time to particularize its scope and the extent of its validity... The sociology of knowledge... reaches a point where it also becomes a critique by redefining the scope and the limits of the perspective implicit in given assertions' (pp. 264-5).

It is worth noting that this formulation does not logically exclude the possibility that a particular assertion may be shown to be universal in its 'scope and limits'.

Since the 1970s, a sociology of knowledge perspective has been applied most fruitfully - and controversially - to the study of scientific knowledge. The sociology of scientific knowledge (SSK) is a term used to encompass a variety of theoretical approaches and research interests, all of which have in common the attempt to ask questions about the nature of scientific knowledge. Mulkay (1992) argues that while Mannheim continues to treat the physical sciences as being a special case in terms of its claims to truth and maintaining a distinction between the social and theoretical sciences (p. 16), his theoretical formulation opens up the possibility for the analysis of natural science as a body of thought in much the same way as a body of thought such as conservatism or liberalism.
Traditionally, the history and philosophy of science have been concerned with outlining the best social conditions for the generation of scientific knowledge, and describing how social influences such as prejudice and ambition led to scientific error (see Potter (1996) for a discussion of Mertonian sociology of science). SSK has a radically different view of the nature of scientific knowledge and its claims to truth. Bloor (1976) sought 'symmetry of explanation': rather than looking to social processes only to explain false belief, Bloor argued that the idea that true scientific knowledge resulted from a decontextualized human rationality discovering the causal features of material reality was fallacious; a full explanation of knowledge demands an understanding of material input and social processes.

By focusing on the social practice of science, SSK radically relativizes our understanding of science. Potter (1996) provides an overview of the various research strands in SSK. Collins (1985) has focused on scientific controversy as a means of studying how scientific knowledge is judged true or false and consensus achieved in scientific academic debate. Amongst others, Bloor (1976) and Pickering (1984) have studied how the content of scientific knowledge relates to various external interests, i.e. group membership and professional allegiances. Constructionist work by Latour and Woolgar (1979) and Knorr Cetina (1996) has involved detailed ethnographic study of the processes whereby scientists construct facts. What the various research orientations in SSK have in common, however, is an insistence that there is nothing epistemologically special about scientific work; that scientific practice produces knowledge in ways comparable to the production of other forms of specialist knowledge. Maynard and Schaeffer (2000) argue that, using the principle of symmetry, the production of social scientific knowledge can be studied in the same way; in their research, they observed practitioners of social science (just as SSK studies have observed practitioners in natural scientific laboratories), and took seriously their subjects' belief that they are carrying out a scientific practice. Mulkay (1992) sums up the basic approach of SSK:

'[What has]... been taken by sociologists to be a complete set of basic principles specifying proper conduct for scientists engaged in research should be seen as no more than a complex social repertoire which scientists use flexibly in the course of negotiating the meaning of their own and their colleagues' actions' (p. 119).
The practice of science (and social science) is thus opened to sociological study.

Work in the field of SSK has, perhaps unsurprisingly, been met with a certain amount of ambivalence by practising scientists (Wolpert, 1993). There is also an extensive philosophical debate to be had over the nature of 'reliable' knowledge and whether, because science produces knowledge which is generally reliable, it can be considered more valid (Ziman, 1978).

Having brought natural scientific claims to true and special knowledge under scrutiny, what does the SSK approach have to say about the relationship between natural and social scientific knowledge? Mulkay (1992) argues strongly that scientific knowledge has, fundamentally, a social foundation:

'Scientific knowledge, then necessarily offers an account of the physical world which is mediated through available cultural resources; and these resources are in no way definitive' (p. 60; my emphasis).

Mulkay's critique of the rigid distinction between the methods and concepts of the natural and the social sciences is based in part on the concern that, when the disjunction is allowed, non-scientists are considered to be lacking in technical competence to judge the adequacy of the claims of science; even though 'claims made by scientists in the wider social context... will often be ideological' (p. 9). However, the denial of a difference between social and natural scientific knowledge, in arguing that scientific knowledge is predicated on social action, can be seen to reaffirm and even prioritize the expertise and competence of the social scientist and, by extension, their access to special knowledge.

In attempting to conceptualize the forms of knowledge which are generated by the natural and social sciences, there is no need to maintain a simple opposition, or to exclude the possibility of multiple sources of knowledge. Mulkay's own reading of Mannheim offers the possibility of another, more hermeneutic analysis:

'[In Mannheim's final position]... both types of knowledge are seen as inherently limited and revisable, but he maintains the distinction between the two spheres by claiming that the limitations of constraints essential to each intellectual domain are quite different in character' (p. 16).
This reformulated version of Mulkay's is flexible and dynamic in its account of interaction between bodies of thought. It dissolves the dichotomy between natural and social scientific knowledge, and recasts any formulation of that split as a rhetorical action, asking what purpose is served by it. Gieryn (1999), who adopts a similar approach, characterizes his interest as being in 'boundary work' and gives a fuller account of it:

"Science" is a cultural space: it has no essential or universal qualities. Rather, its characteristics are selectively and inconsistently attributed, as boundaries between "scientific" space and other space are rhetorically constructed... Science is a symptom of the legitimate power to decide reality - its edges and contours disputed, moved all over the place, settled here and there as decisions about truthful and reliable claims are acted upon by judges, legislators, managers - and ordinary folk. Representations of science - where it is, where it is not - have less to do with the realities they supposedly depict, and more to do with the cultural realities they sustain' (p. xii).

As a result:

'People often take shortcuts when faced with practical decisions about how to allocate "epistemic authority", the legitimate power to define, describe, and explain bounded domains of reality' (p. 1).

I shall make use of the idea of 'boundary work' in the next section, when I discuss how case study methodology has been strategized by sociologists. In this section, I have provided an overview and analysis of those traditions in sociology which have most in common with hermeneutic philosophy: the sociology of knowledge and its offshoot, SSK. In the following section I return to the implications for methodology of my theoretical orientation.
2 Hermeneutics, method, and the 'problem' of validity

Introduction

I now turn to address directly some of the questions put at the start of this chapter: Are the natural sciences the best model for the study of social phenomena? Is the knowledge produced in the natural sciences 'better' than that produced in the social sciences? Can the knowledge created in the social sciences be judged as equally valid and how? These questions have been typically managed in sociology by careful demarcation of the boundaries between qualitative and quantitative research methods. In choosing to describe the work which I have undertaken in terms of 'the case study', I am apparently placing this work on the qualitative side of the qualitative/quantitative divide in social research. However, the theoretical orientation which I have outlined demands that a number of assumptions related to method be explored in greater depth. How have the boundaries of qualitative and quantitative methods been managed within the discipline of sociology? How have proponents of each attempted to legitimize the knowledge they generate based on these methods?

I shall reflect on these questions with particular reference to case study methodology which, besides being the terms in which I describe my own research, is, as Ragin writes, central to the debate over methodology: 'different conceptions of the term "case" are central to the enduring gulf between qualitative and quantitative social science' (1992: 3). I shall suggest that a number of strategies have been used to make claims about the truth status of knowledge acquired through qualitative research; however, what these strategies have in common is a commitment to both a realist ontology and also a conception of social science as a teleological project aimed at the advancement of social science and the production of 'better' knowledge.

Finally, I shall address directly the issue of validity in hermeneutic analysis. I shall argue that casting this as a 'problem' for hermeneutics misconceives the hermeneutic impulse (a word chosen in preference to the more goal-directed 'purpose'). I put forward instead a hermeneutic conception of methodology which draws on Gadamer's (1992) emphasis on the 'giving of accounts' and his reflections on the idea of the 'experience' of truth. Vattimo's (1992) discussion of Gadamer also assists us in understanding Gadamer's conceptualization of hermeneutics as being a
'matter of proclaiming extra-methodic experiences of truth' (Vattimo, 1992: 115). I conclude this chapter by outlining how I conceptualize the accounts I give of the research which I have conducted.

**The qualitative/quantitative debate and the use of case study methodology**

In this section, I shall consider in more detail the nature of the debate between proponents of qualitative and quantitative research in social science. Rather than making the case for either form of inquiry, I want to consider the 'boundary work' (Gieryn, 1999) which advocates of each carry out when making claims as to the effectiveness of their method in gaining 'best' or 'better' knowledge of the world. Gieryn's useful term allows an analysis of how the different claims made about method play an important part in demarcating the boundaries of social science from other disciplines. I shall argue that the case study, in particular, has played a number of strategic roles in attempts by sociologists to manage the boundaries of their discipline, which I shall go on to outline and critique. Moreover, despite analytical work done by, for example, Ragin (1992), I shall argue that the case study remains firmly harnessed to a realist project for the 'advancement' of the social sciences; as Orum, Feagin and Sjoberg (1991) write: 'the study of the single case or an array of several cases remains indispensable to the *progress* of the social sciences' (1991: 1; my emphasis). Relativist hermeneutics seriously questions any claims of 'progress' and 'advancement', a point to which I shall return in the next chapter, where I provide a more detailed critique of the notion of 'progress' in relation to accounts of the relationship between technology and society.

Maynard and Schaeffer (2000), in a discussion of how the respective proponents of qualitative and quantitative methods conceive of the relationship between their various methods, suggest it has been strategized by sociologists in one of two ways (pp. 331-2). Firstly, they discuss what they call *conditional complementarity*, i.e. that qualitative and quantitative methods are complementary modes of inquiry that can be integrated with one another, but conditional in the sense that qualitative research contributes in an exploratory fashion to support quantitative formulation, or else in supporting interpretation (p. 332). The relationship has been secondarily conceived of as one of *critical remediation*; i.e. that survey research needs
the remedial power of qualitative investigation to assist in restoring elementary or
cومونPLACE meaning, through offering understanding of respondents' lifeworlds
(p. 334). Maynard and Schaeffer themselves offer a third means of conceptualizing
the relationship between qualitative and quantitative methods, which they call *analytic
alternation* (p. 335). This emphasizes not just complementarity in terms of
interpretation of results, but in providing a reflexive check on the analytical
presuppositions of both modes of inquiry; asking more what the proponents of the
other method don't know (or what they miss or distort), rather than what they do know
(p. 335). Maynard and Schaeffer's analysis is satisfying, but could be usefully
expanded by a more explicit consideration of the boundary work done by sociologists
to define their discipline's territory, against not just the claims of the natural sciences
but also of historical analysis. I shall now go on to use the example of case study
methodology, which has been claimed by both qualitative and quantitative researchers
as an integral part of their programme of research, to open up this discussion.

Ragin (1992) writes that the centrality of the 'case' to the
qualitative/quantitative debate emerges in part from the unclear definition of what
constitutes a case: 'in quantitative research we use the terms "cases" and "units of
analysis" interchangeably without considering the problems that may come from
conflating data categories and theoretical categories... one researcher may use families
as data collection sites in a survey; another may write a book called "What is
Family?"
(1992: 1). Orum, Feagin and Sjoberg (1991) see the case study as
necessarily rooted in qualitative research, defining it as 'an in-depth, multi-faceted
investigation, using qualitative research methods, of a single social phenomenon... [It]
is conducted in great detail and often relies on the use of several data sources'
(1991: 2; my emphasis). All writers in the field seem to agree that, to distinguish case study
research in sociology from the 'historicist' analysis in which comparisons between
cultures and societies are avoided, there is 'an implicit notion that the objects of
investigation are similar enough and separate enough to permit treating them as
comparable instances of the same general phenomenon... [that] the principle of
repetition is often implicated in statements concerning the relation between the chosen
case and other cases' (Ragin, 1992: 1-2). As Wieviorka (1992) writes: 'the
complement of the case study is comparative analysis' (p. 170).

Case study methodology has thus been used in the literature to carve out a
space for the discipline of sociology in contrast to or in conjunction with the
hypothesico-deductive model of the natural sciences or the singular perspective of historical analysis. Ragin (1992) goes on to offer a detailed overview of how case studies have been commonly conceptualized in the literature:

Table 2.1 Conceptual map for answers to 'what is a case?'

<table>
<thead>
<tr>
<th>Understanding of cases</th>
<th>Case conceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specific</strong></td>
<td><strong>General</strong></td>
</tr>
<tr>
<td>As empirical units</td>
<td>1. Cases are found</td>
</tr>
<tr>
<td>As theoretical constructs</td>
<td>3. Cases are made</td>
</tr>
</tbody>
</table>

Adapted from Ragin (1992: 9).

To summarize: 1. **Cases are found**: cases are empirically bounded, but specific; they must be identified and established as cases in the course of research. The assessment of the empirical bounding of cases is an integral part of the research process. 2. **Cases are objects**: cases are empirically real and bounded, but there is no need to verify their existence or establish empirical boundaries in the course of research, because cases are general and conventionalized. Cases are generally based on existing definitions present in research literature. 3. **Cases are made**: cases are specific theoretical constructs which coalesce in the course of the research. Neither empirical nor given, they are gradually imposed on empirical evidence as they take shape in the course of the research. 4. **Cases are conventions**: cases are general theoretical constructs that structure ways of seeing social life and doing social science, and shape and constrain the practice of social science. Intellectual fashions change, and past work is selectively reconstructed (1992: 9-11).

Ragin goes on to argue (p. 8) that two dichotomies run through these conceptualizations of the case study. The first is identified as one between specificity and generality, i.e. are cases *specific* (e.g. 'the authoritarian personality') and developed in the course of research (e.g. through in-depth interviews or historical research), or are they *general* (e.g. individuals, families, cities, firms) and relatively
external to the conduct of research? Ragin suggests (p. 8) that this dichotomy in turn reflects the qualitative/quantitative divide in social science. Quantitative cases tend to exist as conventionalized, generic categories independent of the research effort, while qualitative cases tend to coalesce as specific categories in the course of research.

Ragin's second dichotomy is related to whether cases are seen as empirical units or theoretical constructs and this, he suggests, overlaps with the philosophical distinction between nominalism and realism. Realists, he argues, believe in cases 'out there' that can be more or less empirically verified as such; nominalists think cases are theoretical constructs that exist primarily to serve the interests of investigators: realists see cases as either given or empirically discoverable; nominalists see cases as the consequences of theories or constructions (1992: 8). In a departure from Ragin, I would suggest that the realist/nominalist dichotomy is less pronounced than he suggests, and that the nominalism employed in the literature is of a limited kind which is allied to attempts to use the case study to justify the discipline of sociology through contrasting or complementing its methodology with those of subjects whose boundaries potentially overlap with it.

There are three main ways in the literature in which case study methodology is deployed to distinguish social scientific research from either natural science or history. These are: contrasting the purpose of the method with that of natural science (i.e. maintaining a divide between natural and social science in order to mark out a distinctive space for the latter); recasting quantitative methods such as the large-scale survey as forms of case study (i.e. implying priority for a particular form of social scientific method); and claiming special efficacy for the comparative method (i.e. claiming that sociology is an 'improvement' on historical analysis). By combining these strategies, writers on case study methodology attempt to distinguish a space for social science which is, most often, a combination of indicating how the discipline can both complement and also surpass other, 'rival' disciplines such as history or psychology.

For example, Orum, Feagin and Sjoberg (1991) take as their starting point a divide between quantitative and qualitative methodology, placing surveys and experiments on one side, and field research (ethnography, participant observation), life histories, and social histories on the other (p. 2). From the outset, then, the utility and claims of quantitative methodology (and, by assumption, the natural science model in the social sciences) are taken as read; the question is straight away one of
outlining and justifying the particular claims of the qualitative method. The case study
is brought into play here, with emphasis on its depth of focus (the two 'principal
advantages' of case studies are detailed, close readings of contextualized phenomena
[p. 278]), and the breadth of technique: 'the study is conducted in great detail and
often relies on the use of several data sources' (p. 2). The 'lessons to be learned' from
the case study (p. 6) are that objects and concepts can be grounded in natural settings
and studied at close hand and, more interestingly, that it furnishes the 'dimensions of
time and history to the study of social life... [allowing] the examination of continuity
and change in lifeworld patterns' (p. 6). Here, the writers attempt a slightly more
daring assault on the previously unchallenged precepts of the hypothetico-deductive
model, noting that 'in the context of time and space, a census or a national survey is a
case study' (p. 48) and, moreover, that 'the preoccupation with surveys and/or
experimentation has led to a disregard of historical processes' (p. 52). Nonetheless,
this challenge has to remain of a limited nature, since the entire argument is dependent
on maintaining the integrity of the opposing model to formulate a simultaneously
contrasting and complementary model for the qualitative end of sociology.

Introducing the question of the ahistorical nature of quantitative method in the
social sciences does, however, leave the account with the claims of another discipline
to manage: that of historical analysis. Here, the account reaches back for the language
of the natural science discourse from which it had previously been attempting to
differentiate itself: the other lesson to be learnt from the case study is that 'it
encourages and facilitates... generalization' (p. 6). Cases 'are usually seen as an
instance of a broader phenomenon, as part of a larger set of parallel instances' (p. 2).
This is set in direct contrast to the efforts of 'the historian, who assumes [s/he] cannot
make these comparisons' (p. 2). A similar set of rhetorical manoeuvres is carried out
by Sjoberg et al. (1991) who position their approach as 'an alternative to the natural
science model on the one hand and the historicist approach on the other' (p. 28). After
critiquing the attempt of natural science to establish universal laws, and the denial of
the historian of the possibility of establishing cross-cultural generalizations, they align
themselves 'with scholars who are committed to the advancement of social science not
by emulating natural science but by formulating a methodology for social inquiry that
has an integrity of its own' (p. 28). The case study 'not only serves as a strategic
supplement to the natural science model, but is an essential feature of sociological
inquiry in its own right' (p. 28).
Having managed the competing claims of the twin rival disciplines and created a distinctive space for sociology which simultaneously adapts from and surpasses both, it is then possible for Feagin et al. to conclude that one can now see the emergence of 'two sociologies': one consisting primarily of journal articles which emphasizes the hypothetico-deductive model of the natural sciences; the other of book-length case studies which, they argue, draws on a long tradition which has 'made up most of the distinguished core of the sociological opus from the beginning of the nineteenth century to the present day' (1991: 270). With both the methodological and even the presentational tensions of the discipline tidily solved, one is now able to proceed along the seemingly inevitable path of 'the progress of the social sciences' to which they had already committed themselves at the outset (1991: 1).

In line with my theoretical standpoint, I have critiqued objectivist accounts of method, in particular the justifications used to claim validity and even superiority for the case study in obtaining knowledge of the social world (the case study being the terms in which I describe my own research). I now outline a hermeneutic conception of method, with particular reference to the notion of validity.

**Validity and the hermeneutic process**

The realist ontology upon which most qualitative and quantitative research methods remain predicated looks to external criteria to judge whether or not findings in such research are true or valid. But the relativist ontology upon which this research is based makes no similar claims or appeals. This is most often judged by realists to be an inherent flaw in relativist reasoning, leading to a circularity of understanding referred to as the 'hermeneutic circle'.

The realist account of the circularity of hermeneutic understanding can be summarized in this way (see, for example, Hollis, 1994: 240-1): the hermeneutic process is an attempt to understand forms of life or texts from within. However, because a text as a whole can only be understood by reference to elements within it, and these elements can be understood only in terms of reference to the whole, the process becomes a vicious circle. Hollis (1994) puts it in terms meant to refute Winch:
"Winch... maintains that social worlds make sense only from within and that, in general, "reality has no key"... "forms of life" include rules of classification governing what counts as real and rules of reasoning governing what counts as rational in belief and action... [n]either perception nor logic can break the circle, because both are internal to the form of life we are trying to penetrate. To put it simply, we need to know what is locally rational before we can know what is locally real; and we need to know what is locally real before we know what is locally rational' (p. 240).

Moreover, the hermeneutic circle, Hollis continues, 'threatens every attempt to justify one interpretation over another' (p. 241), and collapses under the weight of its own argument.

Critiques of relativist accounts of ontology misrepresent these relativist accounts in two ways. Firstly, relativists are cast as arguing a denial of 'truth' rather than emphasizing the existence of 'truths'. Secondly, and resulting from this mischaracterization, it is suggested that relativists argue that truth can be reduced to no more than one's point of view. This particular mischaracterization of the relativist account of truth, apart from a reliance on a Cartesian subject whose continued existence most relativists usually query, fails to be sociologically satisfying. I shall argue that, properly understood, the relativist sociological account can answer the challenge posed by realists concerning the hermeneutic circle.

This dual mischaracterization is closely connected to Gadamer's (1989) argument that hermeneutic theory has been 'far too dominated by the idea of a procedure, a method' (p. 290). Cast in the terms of realism, a relativist account of truth and method is bound to fail. After the relativist account challenges the notion of objective truth, the realist version of this account mischaracterizes this challenge as a denial of the more satisfying position of the possibility of 'truths'. And the outcome of this denial of 'truths' leads one straight back into the hermeneutic circle. But this is an unnecessary move, and is one that remains preoccupied with the questions of validity which are the focus of a realist ontology. Gadamer (1989) writes:

'[U]nderstanding is not merely a reproductive but always a productive activity as well. Perhaps it is not correct to refer to this productive
element in understanding as "better understanding". For this phrase is... a principle of criticism taken from the Enlightenment... Understanding is not, in fact, understanding better, either in the sense of superior knowledge of the subject because of clearer ideas or in the sense of fundamental superiority of conscious over unconscious production. It is enough to say that we understand in a different way, if we understand at all. Such a conception of understanding breaks right through the circle drawn by romantic hermeneutics. Since we are now concerned not with individuality and with what it thinks but with the truth of what is said, a text is not understood as a mere expression of life but is taken seriously in its claim to truth' (pp. 296-7: his emphasis).8

Consider this with reference to Kuhn's account of understanding Aristotle which I have already discussed (see above pp. 16-18): 'What was altered by my own first reading of [Aristotle's writings on physics] was my understanding, not my evaluation, of what they achieved. And what made that change an epiphany was the transformation it immediately effected in my understanding (again, not my evaluation) of the nature of scientific achievement, most immediately the achievements of Galileo and Newton' (see above, pp. 17-18; my emphases). Kuhn was not concerned with making an evaluation, i.e. with the validity, of Aristotelian physics, but with something else.

What is this impalpable 'something else'? Gadamer's concern is with the use of hermeneutics to proclaim the possibility of extra-methodic experiences of truth (see Vattimo, 1992). The purpose of this, Vattimo argues, is to 'counter the claims of modern scientism, which became exclusive only after Kant' (p. 115), and to put forward an idea of truth as belonging to a tradition which can be traced back to the Greeks. Gadamer (1989) describes the discounting of this notion of truth as a symptom of:

'the old quarrel between the poets and the philosophers in the garb of science. It is now said, not that poets tell lies, but that they are

8 In Wittgenstein's (2001) formulation: 'The understanding itself is a state which is the source of the correct use' (p. 49: §146).
incapable of saying anything true; they have only an aesthetic effect and, through their imaginative creations, they seek to stimulate the imagination and vitality of their hearers or readers' (p. 274).

It is provocative, perhaps, to invoke this Platonic opposition between the poets and the philosophers in the context of social scientific discourse, and my stance so far has been to reject such binary distinctions in the classification of thought and knowledge. Gadamer does not, I think (contra Taylor's reading), intend such a dichotomy to be representative of his ideas; he warns, for example, against the 'romantic refraction' inherent in post-Enlightenment thought, whereby a 'mirror image' (romanticism) exists which perpetuates an abstract contrast between myth and reason (1989: 273-4; see also below, chapter 3, p. 51, footnote 11 for further examination of this idea). To my mind, more satisfactory than mirrored opposites is a pluralist idea of 'truths' which conceives of 'family resemblances' between traditions; which looks for affinities between multiple understandings of truth, rather than forming boundaries between opposites. A relativist account operating within the context of the linguistic turn in philosophy might indeed imply some resemblance to 'the poets' - or at least can comprehend the truth they speak as truth.

The second mischaracterization commonly made of relativist ontology is similarly bound up with the concerns of realism, by reducing the notion of 'truths' to no more than the point of view of the individual, that enduring subject of Enlightenment philosophy. But, as with the denial of the possibility of truths, this can be avoided. Gadamer (1989) writes:

'Self-reflection and autobiography... are not primary and therefore not an adequate basis for the hermeneutical problem, because through them history is made private once more [i.e. it remains attached to a realist ontology]. In fact, history does not belong to us; we belong to it. Long before we understand ourselves through the process of self-examination, we understand ourselves in a self-evident way in the family, society, and state in which we live. The focus of subjectivity is a distorting mirror. The self-awareness of the individual is only a flickering in the closed circuits of historical life' (p. 276; my note).
The location of our elusive conception of truth is therefore the similarly intangible arena of our common understandings: 'The common behaviour of mankind is the system of reference by means of which we interpret an unknown language' (Wittgenstein, 2001: 70: §206). In Gadamer's terms: '[t]he task of hermeneutics is to clarify this miracle of understanding, which is not a mysterious communion of souls, but sharing in a common meaning' (1989: 292).

How, then, to clarify this? Firstly, this is a conception of plurality of truths which has a strongly sociological character, interested in the reasons for and ramifications of the common grounds of understanding between and across world-views, whether contemporary or historical. Secondly, we are once again alerted to Hunter's (1999) suggestion that all communication is a form of persuasion, with a rhetorical function. Wittgenstein (2001) also outlines the function of persuasion in reaching or teaching an understanding:

'What do I mean when I say "the pupil's capacity to learn may come to an end here"? Do I say this from my own experience? Of course not. (Even if I have had such experience.) Then what am I doing with that proposition? Well, I should like you to say: "Yes, it's true, you can imagine that too, that might happen too!" -But was I trying to draw someone's attention to the fact that he is capable of imagining that? -I wanted to put that picture before him, and his acceptance of the picture consists in his now being inclined to regard a given case differently: that is, to compare it with this rather than that set of pictures. I have changed his way of looking at things' (p. 49: §144; his emphasis).

Acceptance of another's truth thus becomes a matter of, firstly, recognizing the possibility of that truth ('it is enough to say that we understand in a different way, if we understand at all' [Gadamer, 1989: 297]), and then a matter of being persuaded by the account that they give; '[a]lthough it gladly calls itself "theory of science" (Wissenschaftstheorie), [philosophy...] still stands by the claim... to be a giving of accounts (Rechenschaftsgabe)' (Gadamer, 1992: 2). But how does the relativist choose between accounts? Is it possible for her to choose?

Pleasants (1999), in his reading of Wittgenstein, is concerned with critiquing the central 'idea' of social theory - that is, that social and political criticism needs a
foundational theory of both the individual and society - as a pseudo-scientific 'myth' based on a discredited positivist conception of social science. In particular, Pleasants contrasts - as mutually exclusive - two interpretations of Wittgenstein: one which treats Wittgenstein's philosophy as a radically new method for obtaining a more accurate picture of various phenomena than traditional philosophical perspectives; and another interpretation which accepts Wittgenstein's statement:

'[W]e may not advance any kind of theory. There must not be anything hypothetical in our considerations. We must do away with all explanation, and description alone must take its place'
(2001: 40; §109).

Pleasants (1999: 2) takes Wittgenstein's anti-theoretical injunctions seriously, and advocates 'a thoroughgoing scepticism towards the "explanatory power" and critical efficacy of the social theory that is critical social theory [represented in his argument by Habermas, Giddens, and Bhaskar]' (p. 10). Pleasants characterizes Wittgenstein's critique of traditional philosophy and his own critique of critical social theory as 'immanent' (p. 31); i.e. it seeks to demonstrate that a proponent's position is inadequate because of contradictions and absurdities internal to that position, rather than in contrast to one's own theory of 'the way things really are'. One of his overall aims is to 'expose the conflation of theoretical representation and social critique' (p. 80).

So far so good, particularly in the reading of Wittgenstein and the critique of critical social theory. But where does this leave the relativist who wishes to make a judgement; who wishes to offer what she believes is a more persuasive account? Pleasants does not deny her wish but, in line with his conception of immanent critique, does not offer a new agenda:

'I have not said, nor do I believe, that social and political scientists, sociologists, and other "social scientists" should restrict themselves only to describing/understanding what people do, according to people's own local criteria of relevance... However, I do contend that the presentation of a new "ontological picture" consisting of universally possessed tacit knowledge and a transcendental order of rules, does not
ipso facto constitute a critical perspective on, nor an emancipatory intervention in, social life itself... I suggest that this procedure is a form of emotivism, whereby the theorist presents their normative attitudes as "objective facts" of individual and social ontology' (p. 77).

The provision of an account need not necessarily mean suggesting one has a more accurate account; it can imply nothing more than one is offering a different account, as Gadamer writes: 'It is enough to say that we understand in a different way, if we understand at all' (see above, pp. 31-32; his emphasis). But neither does 'understanding' mean the same as 'agreement' or 'endorsement', i.e. one can believe that one has come to understand the meaning of another's stance without being persuaded of its truth. The relativist can and does make judgements. Herrnstein Smith (1997) advances this point onwards from Pleasants' 'immanent critique':

'contrary to the current and classic charge of self-contradiction, when non-objectivists do judge, act, and justify their actions, they do not stop being non-objectivists' (p. 3).

In part, this is done by recasting the objectivist terms of reference, as Herrnstein Smith continues:

'[I]f someone rejects the notion of validity in the classic (objectivist) sense, what follows is not that she thinks all theories (and so on) are equally valid but that she thinks no theory (and so on) is valid in the classic sense... Theories can be and are evaluated in other non-"objective" ways, e.g. applicability, connectability, etc.' (p. 77-78).

She also writes:
...the alternative to the dream of objective judgment is not pessimism, cynicism, or torpor, but attentiveness, responsiveness, and activity, both intellectual and pragmatic' (p. 23).9

There is, in addition, another aspect to Herrnstein Smith’s account, and it is no accident that the title of her book indicates that it is concerned with belief. This is closely connected to Wittgenstein’s (1975) account of certainty, and fundamental convictions:

'Certainty is as it were a tone of voice in which one declares how things are, but does not infer from the tone of voice that one is justified' (p. 6: §330; his emphasis).

Pleasants (1999: 115) teases out the implications of this: that one need not doubt the genuineness of the realist’s certainty, but that it is a perspective of the individual and should not be confused with knowledge claims. He continues:

'[t]he logic and rationality of a "fundamental conviction" consists in it not being open to argumentative refutation - that is what makes it a fundamental conviction' (p. 171; his emphasis).

Nonetheless:

'...the strange thing is that when I am quite certain of how the words are used, have no doubt about it, I can still give no grounds for my way of going on. If I tried I could give a thousand, but none as certain as the very thing they were supposed to be grounds for' (Wittgenstein, 1975: 39: §307; his emphasis).

It is not that the relativist refrains from judgements. But she understands the partiality of - the prejudices inherent in - those judgements; i.e. she understands the

---

9 The Greeks, of course, had a word for it: φρονησις (phronesis), i.e. practical wisdom or prudence, the application of sound judgement in everyday life, in contrast with the more theoretical inquiry leading to σοφία (sophia).
justifications upon which these beliefs are based, and that these beliefs are based upon justifications which - she believes - have no foundation. Belief is not conflated with objective knowledge.

'The difference between the concept of "knowing" and the concept of "being certain" isn't of great importance at all, except where "I know" is meant to mean: I can't be wrong' (Wittgenstein, 1975: 3: §8; his emphasis).

Much of what follows in this thesis is concerned with critiquing realist accounts of the relationship between technology and society, in examining the claims to truth which are explicit and implicit in these accounts, and in considering how these claims to truth based on realist accounts of technology operate within social scientific discourse.

The relativist perspective is brought to bear most immediately in the following chapter 3, in which I critique a particularly enduring form of realism within theoretical accounts of the relationship between technology and society - technological determinism and its 'mirror image' of social determinism. I critique the rhetorical strategies used in these accounts of the relationship between technology and society when I turn, at the end of chapter 3, to hypertext, where both realist accounts of technology and the rhetorical strategies used to maintain them are considered in the light of the use made of hypertext in academic discourse surrounding the use of technology in qualitative research and analysis in the human and social sciences. This is the point of connection between the general analysis of accounts of technology given at the start of chapter 3, and its relevance to my specific interests in knowledge claims in social science, as I have discussed here in chapter 2. I progress in chapter 4 to examining the uses of technology in quantitative social science, in the particular case of the FRS and my own project to develop online documentation.
3 Technology and society

1 Introduction

In this chapter, I review and critique those realist accounts of the relationship between technology and society which seek to model the causes of social structure and process in terms of technological capacity and change. After a general review and analysis of the literature related to information and communication technologies (ICTs), I look at the specific example of theories of hypertext arising from the humanities and qualitative social research. My analysis throughout is a relativist one which attempts to examine the fundamental convictions upon which these accounts are based, and to open up their claims to truth to scrutiny. In the following section, I begin with a theoretical critique of technological determinism and its converse, social determinism, arguing that these ostensibly conflicting accounts of the relationship between technology and society share a fundamental realist ontology. I then move on to a more specific critique of the technological determinism which underlies contemporary theories of the 'information society'. Two main themes developed here are the tendency of accounts of technological change to describe societies in terms of shifts between epochs, and the way in which technologically deterministic accounts of society shift discussion of new technology to competing and ultimately irreconcilable utopian and dystopian visions of the future.

Having critiqued theories which place technology as the prime feature of social structure and the prime cause of social change, I turn in the third section of this chapter to literature which emphasizes social influences on technology. This section serves a number of purposes in my argument beyond a review and critique of this literature, since I draw upon this literature in various ways to inform my own account. Firstly, I draw out the connections between some of the work conducted in this area and theoretical work done in the field of SSK (which I discussed in chapter 2). Secondly, I draw a loose conceptual distinction between the social shaping and social construction of technology in order to organize my discussion. Thirdly, I review and critique attempts made in the literature to resolve the tension between technological and social determinism. The 'resolution' which I find most satisfying results from a debate in the field of SSK which arose from the emergence of actor-network theory (ANT). I explicitly characterize this debate in terms of Gadamer's notion of
experience and the open-ended conception of knowledge generation which I explicated and endorsed in chapter 2.

In the final section of this chapter, the themes emerging from the earlier sections are brought to bear in a discussion of hypertext. Realist accounts of hypertext and the rhetorical strategies used to maintain them are reviewed and critiqued. An account of the emergence of hypertext and the Internet is given, and a critique of theories of hypertext developed, showing how the discussion of hypertext has reflected the twin themes of epochal social change and utopian/dystopian visions of the future. The reception of hypertext by the humanities and qualitative social science is discussed; I argue that hypertext has been constructed in this academic discourse as having particular characteristics in order to legitimize knowledge generated in social scientific research; this is connected back to discussions in chapter 2. I review and critique taxonomies of hypertext which appear in the literature and which were significant during the development of the FRS questionnaire, as I detail in chapters 6 and 7.

2 Determinism in accounts of the relationship between technology and society

The relativist critique of determinism

Attempts to describe the relationship between technology and society have taken a variety of forms. The critique in this chapter follows the principles outlined in chapter 2 in attempting to examine the fundamental convictions which underlie these accounts and enable the construction of explanatory models of social change which use technology in some way as their foundation. This chapter attempts to achieve the 'thoroughgoing scepticism towards the "explanatory power"... of social theory' that Pleasants (1999) advocates (see chapter 2, p. 35). Later sections develop my own accounts of how this relationship might be conceptualized beyond what I shall argue are overly simplistic deterministic models.

My initial focus, then, is on the most prevalent extrapolation from an essentialist ontology to the construction of explanatory models of social structure and process based on technology - technological determinism. Technological determinism attributes a causal role to technology in explaining social change; indeed, in its
stronger forms it asserts that technology is the most significant cause of social change. A very simple form of technological determinism might be the argument that the development of the steam engine enabled Britain to industrialize more quickly, hence allowing Britain to build an empire. More complex forms of technological determinism underpin, for example, such notions as 'the Industrial Revolution', attributing the creation of new forms of society to changes in productive technique.

The problem of such a model of social change is that it is overly simplistic. Social change is too complex a process to attribute to a single cause; in addition, such an account ignores how technology itself is open to social influence, for example, in the development process. Nonetheless, such accounts are common. MacKenzie (1999) suggests this is partly because many people experience technology in their everyday lives as something in which they have no involvement and over which they have no control, that technology becomes 'autonomous' (Winner, 1977). Moreover, a simple model of cause and effect allows the production of predictions of the social effects of new technology (MacKenzie, 1999: 39-40). Technological determinism is a deceptively appealing account of social change which gives the appearance of explanation and prediction. Thus, as new forms of information and communication technologies (the telephone, the personal computer, the Internet) have emerged, technologically deterministic accounts of their potential impact have also surfaced, for example, in notions of the emergence of the 'Information Society' (such as those that inform UK government policy initiatives on e-Strategy; see chapter 4, pp. 94-99). In the case of hypertext technologies, it has been argued that the shift to the electronic representation of information brought about by hypertext and the Internet marks a decisive break with the types of knowledge produced by previous modes of communication (Landow, 1992).

My immediate purpose, then, is to critique the technological determinism which underlies much contemporary theorizing on the relationship between technology and society. When I move on to develop my own account, I attempt to avoid technological determinism - and also the contrary temptation to privilege the social as a form of explanation. Woolgar and Pawluch (1985) critique such a manoeuvre in their analysis of contributions to the literature on social problems emerging from a 'social constructionist' perspective. Much of this literature, Woolgar and Pawluch argue, depends upon a 'selective relativism' (p. 214) in which the nature of a condition (e.g. deviant behaviour) is held constant, for the purposes of arguing
that it must be other phenomena (i.e. definitions of behaviours as sinful, immoral, criminal, etc.) which have changed. Woolgar and Pawluch call this 'ontological gerrymandering':

'The successful social problems explanation depends on making problematic the truth status of certain states of affairs selected for analysis and explanation, while backgrounding or minimizing the possibility that the same problems apply to assumptions on which the analysis depends... Some areas are placed as ripe for ontological doubt and others portrayed as (at least temporarily) immune to doubt' (p. 216).

In such accounts as Woolgar and Pawluch describe, a rhetorical manoeuvre takes place which apparently counters an essentialist explanation but, in offering a converse model, remains based upon the fundamental criteria which it ostensibly seeks to critique or reject. This theme of the construction of 'mirror images' which do not disturb the foundations upon which accounts rest is one which will re-emerge periodically throughout the rest of this chapter.

MacKenzie (1999) suggests that the case for technological determinism generally takes two forms. Firstly there is the idea that technological change follows a logic of its own independently of human will (MacKenzie, 1999: 39). I shall critique this in section 3 of this chapter, when I consider literature which emphasizes social influences on technology. The second form of technological determinism which MacKenzie describes is the belief that changing technology brings with it social change - i.e. that technology acts as the structural basis of an explanatory model of change in society. In the following section, I outline and critique theories of the impact of information and communication technologies (ICTs) on society.

Theories of the information society

Theories of the emergence of a new form of society based on changes in the technical organization of information have taken a number of forms over the past thirty years. In the mid 1970s, Daniel Bell (1976) argued that a new system was emerging: a 'post-
industrial society'; this was marked by the decline of the industrial and agricultural sectors of the economy and the growth of the service sector. The post-industrial society was distinguished by a heightened presence and significance of information, both quantitatively and qualitatively, i.e. that new forms of information, which he termed 'theoretical knowledge', were becoming more important. Bell's thesis contains a number of significant flaws, not the least being, Webster suggests (1995: 31), its unsustainability as anything other than an ideal type construct, its teleology, and its endorsement of a convergence theory of development (i.e. that all societies are set on the same kind of developmental journey, from pre-industrial, to industrial, to post-industrial).

During the 1980s, however, this hypothesis found new expression in the work of Piore and Sabel (1984) who suggested we were living through 'a second industrial divide' which was comparable to the one which brought about mass production at the end of the nineteenth century. This involved a shift to 'flexible specialization': a change from the repetitive forms of labour epitomized by Fordism and scientific management towards a form of labour which emphasized the skills of workers and greater variety in goods. Flexible specialization, it was argued, was particularly found in the sorts of small, high-tech firms which had begun to emerge in Silicon Valley from the late 1970s onwards. It is emerging information technologies which are seen as the major facilitators of this flexibility. A large literature devoted to the 'networked' or 'virtual' organization has emerged from this hypothesis, emphasizing the use of new ICTs such as video-conferencing in business, or novel forms of business collaboration, for example, in new product development (see the discussion and literature review in Dutton, 1999). Fukuyama (1999), highlighting again a purported split between what has gone before and what he sees as emerging, is more negative in his evaluation of this shift, arguing that the emergence of an information society, whilst leading to greater freedom of choice for consumers and increasing democracy, has created a "Great Disruption" in the social values that prevailed in the industrial age society of the mid-twentieth century' (p. 45).

In recent years, the most significant variation on the theme of the information society has been Castells' articulation of the 'informational mode of development' (Castells, 1996); i.e. a new form of economic organization in society predicated on advancements in information technology. New developments and applications of
information technology are causing a new form of social organization, the 'rise of the network society'. Castells argues:

'A new economy has emerged in the last two decades on a worldwide scale. I call it informational and global to identify its fundamental distinctive features and to emphasize their intertwining. It is informational because the productivity and competitiveness of units or agents in this economy... fundamentally depend upon their capacity to generate, process, and apply efficiently knowledge-based information. It is global because the core activities of production, consumption, and circulation... are organized on a global scale, either directly or through a network of linkages between economic agents... The Information Technology Revolution provides the indispensable, material basis for such a new economy... [T]he evolution of technology has indeed largely determined the productive capacity of society and standards of living, as well as social forms of economic organization. Yet... we are witnessing a point of historical discontinuity' (pp. 66-67).

Drawing explicitly on Bell (p. 14), Castells argues that societies can be characterized along two axes: the mode of production (capitalism, statism) on one axis, and the mode of development (pre-industrial, industrial and post-industrial or, in his terms, informational) on the other. Modes of development are the 'technological arrangements through which labor works on matter to generate product, ultimately determining the level and quality of surplus' (p. 16). In the agrarian mode, productivity arises from increases in labour and natural resources; in the industrial mode, it arises from the use of new energy sources. In the emerging informational mode of development, 'the source of productivity lies in the technology of knowledge generation, information processing, and symbol communication' (p. 17). While Castells acknowledges that knowledge and information are critical to all modes of development, 'what is specific to the information model of development is the action of knowledge itself as the main source of productivity' (p. 17). Moreover, Castells argues that technology and technical relationships:
'diffuse throughout the whole set of social relationships and social structures... [M]odes of development shape the entire realm of social behavior, of course including symbolic communication [i.e. cultures and collective identities]' (p. 18).

Since the informational mode of development is based on technology and knowledge, there is within it a particularly close link between culture and productive forces: 'It follows that we should expect the emergence of historically new forms of social interaction, social control, and social change' (p. 18).

Castells attempts to avoid the charge of technological determinism in two ways. The first of these is to allow a limited degree of social influence on technology by emphasizing the role of the state in technological development:

'[T]he role of the state, by either stalling, unleashing, or leading technological innovation, is a decisive factor in the overall process, as it expresses and organizes the social and cultural forces that dominate in a given space and time' (p. 13).

This is an impoverished account of the way in which social factors can influence technological innovation. It puts all social and cultural action within a given society under the aegis of the state, a view which ignores the other varied social relations in play in any given society and their influence on the development of scientific and technological knowledge (for example, the Catholic Church in early modern Europe). Moreover, 'the state' is an evolving concept; one which, for example, only emerged in its modern sense in Europe during the sixteenth century (Skinner, 1978). Castells applies it equally not only across historical periods in Europe, but also to modern Japan, the Soviet Union and, most problematically, China across the whole of its history in his discussion of the ways in which state activity can 'suffocate' the development of technology (p. 7).

There is a persistent tendency within sociology and social theory to describe social change in terms of shifts between epochs. Sociology's emergence as a discipline in the nineteenth century took place in an intellectual environment in which the evolution and (most usually) progress of societies was a given. Comte offered an initial formulation of the evolutionary nature of society in a three-stage model (see
The association of progress with industrialization and technology was implicit in this foundational account and reinforced in the 1880s by Tönnies' (1963) distinction between Gemeinschaft (community) and Gesellschaft (association). This was an attempt to describe Germany's transition from rural to industrial society, and was based on the hypothesis that pre-modern European societies were made up of communities of close networks of personal relationships. In the shift from agricultural communities to large, urban, industrial societies, an impersonal framework of laws and other formal regulations replaced these informal relationships.

While nineteenth-century accounts almost invariably associate a discourse of change with a discourse of progress, the twentieth-century experience of totalitarianism and the techniques of mass warfare brought deep pessimism about the inevitable betterment of society by means of technological advancement. (I shall, however, show below how now both optimistic and pessimistic visions of the purported new society are constructed on this basic assumption that society is shifting into a new era.) Nevertheless, the impulse to theorize societal change in terms of epochal shifts remains strong and is manifest in current theories of a transformation from the industrial to the post-industrial or information society, not to mention, in some versions, the battle over the transition (or not) from the modern to the postmodern era. Where such theories fail is in their view of historical change, which tends to be simplistically teleological, implying a straightforward trajectory which (invariably Western) society has followed since pre-industrial times. A model of change in society based on developments in communication media has been offered by writers such as Landow (1992) and Bolter (1991), in which hypertext is positioned as the next agent of epistemological change in a trajectory which began with oral communication and has passed through the printed book to the current position.

Critics of the idea of a shift to a new form of society based on the emergence of new ICTs emphasize continuities with the past. As Webster (1995) sums up:

'On the one side are subscribers to the notion of an "information society", while on the other are those who insist that we have only had the "informatisation" of established relationships... On the one hand there are those who subscribe to the notion that in recent times we have seen emerge "information societies" which are marked by their
differences from hitherto existing societies... On the other hand there are scholars who, while happy to concede that information has taken on a special significance in the modern era, insist that the central feature of the present is its continuities with the past' (pp. 4-5; his emphasis).

Those who emphasize continuity with the past frequently offer examples which emphasize the long histories of apparently new technological developments, and also similarities between the reception of earlier technologies and 'new' ICTs. For example, Winston (1998), in what Golding (2000) calls 'a book-length assault on technological determinism' (p. 171), traces the history of the Internet back to the mid 1800s and the 'first wired network' - the telegraph system. He argues:

'In order to provide a context for outlining the development of the Internet we need to go back to the beginning, to the start of electronic communications, to show how central the building of networks has been to their success and how much the current networking of computers conforms to these historical patterns' (p. 243).

Winston goes on to develop a model of what he calls 'the "law" of the suppression of radical potential', i.e. a process whereby 'general social constraints coalesce to limit the potential of [technologies] radically to disrupt pre-existing social formations' (p. 11) or what Golding calls 'the solidity and endurance of social and economic formations in the face of technical novelty' (2000: 171). In the case of the Internet, Winston emphasizes the incremental nature of its inception rather than it being a radical invention:

'The Internet emerges in the US in the 1970s as a species of spin-off from a (largely still classified) national security project rather than any sort of "discrete" invention' (p. 325).

From the late 1980s, control of the Internet's backbone was handed over from the US government-funded National Science Foundation to private telecommunications giants Sprint, Ameritech and Pacific Bell (a transition completed in 1995). Winston suggests that:
'Those who seriously believed they were in a brave new world of free and democratic communications were simply ignoring the reality of their situation... a straightforwardly classic expression of the suppression of radical potential whereby the new technology is distributed among the established players to minimise the threat to their business' (p. 334).

Winston carefully sets the word 'law' within quotation marks, and insists that 'although the phenomenon under discussion can be found in the histories of all telecommunications technologies it is not so regular as always to manifest itself in the same form with equal force at the same point of development' (p. 12). Nonetheless, it is difficult to see how he can avoid precisely the charges of inevitability and determinism which he himself is laying at the door of those who speak of radical change.

Historical studies which emphasize the similarities between the reception of new technologies in the past seek to show how the claims made for them then often mirror claims made for ICTs now. For example, Marvin (1988), in her study of the social impact of the telephone, cites a newspaper report enthusiastically outlining a vision of a society changed by the recent arrival of a new technology:

"[N]othing less than a new organization of society - a state of things in which every individual, however secluded, will have at call every other individual in the community, to the saving of no end of society and business complications, of needless goings to and fro, of disappointments, delays, and a countless host of those great and little evils and annoyances which go so far under present conditions to make life laborious and unsatisfactory" (p. 65).

This does not refer to the Internet, or to email, but to the telephone, and it appeared in the *Scientific American* in January 1880. These sorts of claims are instantly recognizable to those tracking the reception of and debates surrounding the impact of new communication technologies 120 years later. Marvin describes how it was widely discussed how fortunes could be made through the use of these technologies; how women's use of new technology was singled out for particular
discussion; how the new communications techniques were brought to the fore in debates over fighting crime and 'an association between sensational crime and the new electric media was strong in popular and expert literature' (p. 92).

The use of history as a rhetorical device to counter technological determinism is a deliberate tactic on the part of some writers. For example, Robins and Webster (1999) write:

'Without history, the new technologies become an unstoppable force which, though incomprehensible to [ordinary people], is understood sufficiently for them to realise that they must change their whole lives' (p. 74).

They argue that treating technology ahistorically creates:

'a general sense of acquiescence to innovation. We believe this happens because technology, without discernible origins, is something that ordinary people cannot understand' (p. 74).

The somewhat arrogant overtones of this formulation - in which 'ordinary people' are seen as being at a disadvantage when attempting to make sense of the world (presumably because of a lack of familiarity with social theory) - sit uneasily with Robins and Webster's appropriation and celebration of Luddism as a motif for popular responses to technological change throughout their argument\textsuperscript{10}. In their discussion of the continuities between our current experience of new technology and those of the past, Robins and Webster argue that Luddism was a response to 'the unfolding logic of the Enclosures movement... [and t]he logic of enclosure was the logic of the new capitalist order' (p. 7). Their refusal to accept the notion of an information society is thus an explicitly political stance. Robins and Webster argue that new ICTs do not justify the utopian (or, indeed, dystopian) rhetoric which surrounds them, and suggest that these technologies often work to reproduce 'conservative' social practices:

\textsuperscript{10} A body of literature which does take seriously how people conceptualize technologies and how they are used in unexpected ways - how they have 'interpretative flexibility' - is discussed later in this chapter, under Social influences on technology.
[W]hat is unfolding now is the continuation of what was set in motion in the early-nineteenth century: what we now call the global information economy is, we argue, the most recent expression of the capitalist mobilization of society... [T]here is much about the "information revolution" that is just business as usual (if the technologies are new, the social visions that they generate tend to be surprisingly conservative)' (pp. 6-7).

Golding (2000) also associates the impact of new technologies with the continuation of modernity, where 'modernity' carries many of the markers of Ritzer's (1996b) 'McDonaldization' thesis, e.g. increased rationalization of social life and the extension of capitalist organization. Golding writes:

In assessing the impact, both recent and immanent, of these technologies, we find, above all, the abiding fault lines of modernity' (p. 179).

These fault lines he sees as being associated with three broad trends: convergence, i.e. the merging of ICTs as big entertainment corporations from different spheres (TV, video, film, Internet etc.) engage in takeovers and mergers; the deregulation of state intervention into communications industries; and differentiation, i.e. a translation of income inequalities into ICT stratification - the process of social exclusion of the poorer segments of society from new ICTs (p. 179). He goes on to invoke a specific debate within social theory, between those who are suggesting society has moved into a postmodern era and those who, whilst acknowledging flaws in modernity, still perceive value in its 'project':

In part this [Golding's approach] is an insistence on the endurance of modernity and the intellectual and political baggage that comes with it. In part it is a plea for a stay of execution of the core tools and methods of the sociological imagination, and a reminder that the basis of prediction lies in examining social dynamics rather than technological innovation' (p. 166; my emphasis).
However, by invoking the concept of modernity, theorists of informatization reinforce the idea of the possibility of a decisive break in society by providing the rhetoric to support the opposing case for continuity. Robins and Webster's complaint of the tendency in information society theories to herald the arrival of a new epoch (1999: 75) rings false: the discourse of radical change invokes the discourse of continuity; but both are 'grand narratives' of historical epochs which carry with them implicit notions of inevitabilism, progress and historicism - precisely those deterministic notions which these theorists attempt to criticize. It is possible to offer a plausible account of the complex relationship of technology and society without falling into the trap of placing technological advances on one side of the continuity/change divide, or invoking the notion, by agreement or disagreement, of trajectories along which society proceeds (or declines).

Castells' second tactic to counter accusations of technological determinism is to deny its importance as an issue:

'The dilemma of technological determinism is probably a false problem, since technology is society, and society cannot be understood or represented without its technological tools' (p. 5).

Underlying such an argument and, also, the theoretical perspectives of Bell, Piore and Sabel, and Fukuyama is the idea that there is a technical realm within society which, in important ways, remains unaffected by the realm of values and beliefs - yet is the foundation on which society is built. (I shall return at the end of the next section to actor-network theory, which offers a similar account, but which I shall argue has a different philosophical basis from those accounts which I now go on to

---

11 This is reminiscent of Gadamer's account of the romantic movement as mirroring the Enlightenment: 'the romantic reversal of the Enlightenment's criteria of value actually perpetuates the abstract contrast between myth and reason... all criticism of the Enlightenment now proceeds via this romantic mirror image of the Enlightenment' (1989: 273-4). This mechanism of constructing a 'mirror image' critique which leaves unchallenged certain fundamental criteria can be seen in operation in the technological vs. social determinism debate discussed above (p. 42); also in Ragin's use of a limited nominalism which can be encompassed in his taxonomy of case studies (see chapter 2, p. 28); and, of course, in those technologically produced utopian and dystopian visions of the future which are under scrutiny at this point in this chapter. The debates which arise on top of these foundations of mirroring opposites frequently take on the aspect of circularity. Le Guin, on the utopia, writes: 'It is sad that so many stories that might offer a true vision settle for patriotic or religious platitude, technological miracle working, or wishful thinking, the writers not trying to imagine truth. The fashionably noir dystopia merely reverses the platitudes and uses acid instead of saccharine, while still evading engagement with human suffering and with genuine possibility' (2004: 219).
This idea is not novel, but the emerging ICTs have given it new currency. Webster characterizes these theories as having 'the same emphasis on the transformational, indeed foundational, characteristics of changes in techniques of production throughout history and, most recently, in the role of information and knowledge' (1995: 196). Webster describes such thinking, particularly from Castells, as a form of Althusserian Scientific Marxism which draws a distinction between relations of production (classes) and forces of production (techniques). This creates a fundamental dichotomy between the realm of values and social organization, and that of technique and technology.

This idea has not gone uncontested, from within both Marxist theory and sociological perspectives. Gouldner (1980) suggests there are 'Two Marxisms': Scientific Marxism, as described above; and a tradition of Critical Marxism, present in writers such as Theodor Adorno, Herbert Marcuse, and, in the UK, Raymond Williams and E.P. Thompson. These writers are marked by:

'a characteristic refusal to privilege technique in examination of social change, either by regarding it as the primum mobile of change or by presenting it as something set apart from the social world' (Webster, 1995: 198).

Within these very varied critical accounts, technology and technique are generally perceived as being part of a whole arrangement of relationships under capitalism, which have to be understood in their historical context, and in ways which mean that social values are present in the process of technological development itself. Various sociologists have argued that both science and technology are not developed independent of social pressures and values (e.g. Latour and Woolgar, 1979; MacKenzie and Wajcman, 1999a; MacKenzie and Wajcman, 1985); I shall discuss this literature on the 'social shaping of technology' in greater detail below.

Robins and Webster (1999: 68-73) provide a robust theoretical argument against the discourse of the emergence of an information society, describing the problems of removing technology from its social context and treating it as an isolated phenomenon. Once desocialized, technology is also seen as neutral, a tool to be used either appropriately or not, depending on the motives of a society. This, they suggest, can have potentially deleterious political effects: 'if technology is socially neutral and
leaves policy choices to the public, then on what reasonable grounds can it be suspected?' they ask rhetorically (p. 69).

Shifting the focus from theoretical critique to empirical data, Golding (2000) collects evidence from empirical research to suggest that there are four fallacies associated with the information age thesis. The first fallacy relates to identity. He argues that much writing on identity in the information age, particularly from cyberculture studies (e.g. Jones, 1995 and Turkle, 1997) contains what he calls the 'fallacy of the postmodern subject' (p. 172); i.e. that stable identities are being eroded. He counters this by pointing to the resilience of expressions of national identities (Billig, 1995), but does not take account of the possibility that nationalism might in part be a response to the erosion of stable identities. The second fallacy relates to inequality: it is claimed that ICTs will lead to the end of deprivation and need (p. 174); against this he sets ONS statistics which suggest 'a settling pattern of high users and excluded non-users which will provide a digital underpinning to structures of material inequality that are more likely to become self-replicating than abating' (p. 175). The third fallacy, he suggests, relates to power: a fallacy of interactivity. It is suggested that new ICTs will create more democratic forms of political organization and stronger communities. Golding argues that instead individualization, unequal access, and disenfranchisement could as easily be the result of Net politics (p. 176). The fourth fallacy concerns change: that there is a fundamental shift in the organization of society related to the compression of time and space: Cairncross' (1998) 'death of distance' - 'probably the single most important force shaping society in the first half of the next century' (p. 1). Golding draws on studies of travel statistics and the slow take-up of homeworking to argue that this theoretical account is fallacious.

Conceptualizing new technology as the main driver of social change fails to address the substantive issue of how technologies are constructed and implemented. Instead it provokes a sterile debate between believers in utopian and dystopian visions of the potential of the 'information society'. For example, contemporary debate over the impact of the Internet and digital communication technologies is underscored by great ambivalence about what to make of these innovations. Women are portrayed as either disadvantaged by technologies which reward a 'masculine' sense of mastery (Turkle, 1988), or as appropriating such technologies to their own ends (Spender, 1995). The fortunes made by dot-com millionaires are discussed in the same breath as
the bursting of the net bubble. In contemporary debates over crime and new technology, 'the Internet' seems almost to be synonymous with pornography, or else to provide the key to successful surveillance of criminal activity. Whilst failing to capture the dystopian element which plays such an important role in the construction of this deterministically based discussion of new technology, Robins and Webster (1999) excoriate:

'this wishful marketing discourse, with its magical vision of new technologies as the solution to our social ills - promoting participatory politics, material comfort, improved pedagogy, better communications, restored community, and whatever else you may think of' (p. 5).

An analysis of the way in which the Internet has been conceptualized as impacting on political action provides a good example. Rheingold (1994) suggests that there are two visions of the Net running through contemporary discourse (pp. 14-15). Enthusiasts see the Net as the new 'agora', i.e. the Athenian marketplace where citizens met to talk and debate. In this vision, the Internet has the potential to revitalize democracy, and enables people to form communities across gender, class, race and national boundaries. Rheingold's own description of his experiences of the WELL online community focuses primarily on the positive aspects of cyberspace: he describes numerous occasions in which community members provide each other with information and emotional support. For example, when the young son of one member became seriously ill, the community rallied round with practical information and advice on his condition from a doctor in the community, and good wishes and emotional support from other community members. Rheingold also describes to great effect the impact on the WELL community of the death of one of its more eccentric members.

Conversely, Rheingold suggests, pessimists see the emergence of a 'panopticon', a term devised by Bentham and popularized by Foucault to describe an environment in which people act as if they were under surveillance all the time. Samarajiva (1996) suggests that electronic environments by their nature are more open to surveillance than physical environments, since tracking and storage procedures can be built into their design:
'Being relatively more malleable than physical environments, electronic environments are more conducive to dynamic and continuous exercise of control through... technical features...

[Electronic environments can be designed to enable pervasive and transparent surveillance through the tracking of usage patterns and long-term storage of such information' (p. 133).

Whilst Samarajiva suggests that such surveillance can be evaded by users either unaware of it or actively subverting it, other commentators remain more pessimistic. Shields (1996) describes the potentially deleterious effects of new technologies on our existing communities:

'The neglect of face to face communities has also raised fears about the decline of the public sphere into a virtual world controlled by telecommunications corporations where only the privileged have access and the body is disdained as an embarrassing and imperfect support for minds infatuated with virtual, representational bodies' (p. 1).

It appears, therefore, that we can conceptualize the impact of the Internet on political action in two ways: as a means of breaking down community barriers and revitalizing social interaction, or else as a means for furthering the interest of government and big business at the expense of an already moribund public sphere. However, as Mansell and Silverstone (1996) somewhat dryly point out:

'Simplistic utopian or dystopian visions of the future provide us neither with an understanding of how these changes come about nor with an understanding of the longer-term implications' (p. 3).

The rhetoric of optimism or pessimism which permeates contemporary discussion about new ICTs carries with it an implicit assumption that technology

\[12\] Similar arguments (with considerably less hyperbole) can be seen in the literature discussing possible effects of the Web on survey research (see chapter 4, pp. 106-107); the language of 'transformation' appears in UK government policy initiatives on the information age (see chapter 4, pp. 94-99).
impacts upon rather than is influenced by the social realm; technology is thus withdrawn from the 'realm of values and beliefs'. This section has provided both theoretical and empirical objections to such a conception of technology. In the following section, I examine literature which is concerned with social influences on technology, and outline a relativist perspective which attempts to avoid privileging the social, as critiqued by Woolgar and Pawluch (see above, pp. 41-42).

3 Social influences on technology

Introduction

As discussed above (p. 42), MacKenzie's (1999) second form of technological determinism is characterized as the idea that technological change follows a logic of its own independently of human will (p. 39). Robins and Webster (1999) argue that such a conception of technology constructs a particular view of history, one in which history is seen as the process of technological advance, and which carries an underlying inevitabilism (p. 69). This perception, they believe, construes technology as 'a hidden hand in development apart from the social issues of power and control' (p. 70). The view of historical change offered in these theories is a simplistically teleological one, bound to notions of inevitable technological (and hence societal) progress. Williams (1999) outlines the ramifications of such an approach:

'Policy-makers and the public have often taken the course of technological progress for granted - as if technology developed according to some predetermined technical rationality - and assumed that the content and direction of technological innovation were not amenable to social analysis and explanation. Such a view limits the scope of social-scientific enquiry to monitoring the "impacts" of technological change upon economic and social life' (p. 41).

In this section, I shall examine the literature which has attempted to answer this form of technological determinism by studying social influences on technology. Responses to such a conception of technology have taken two main forms (see
Mackay, 1995). Firstly, studies of the social shaping of technology have called attention to the many social influences on the process of innovation. Secondly, a theoretical perspective emerging (broadly) from cultural theory emphasizes the way in which technology can be used in unexpected ways. In this section I shall review the literature emerging from these two orientations in turn.

However, as I stated in the introduction to this chapter, this section is intended to be more than a review and critique of this literature, since I draw upon it in various ways to inform my own account. I begin below by tracing the influence of SSK on social studies of technology, to connect this literature back to theories outlined in chapter 2. I then draw the conceptual distinction to which I have already alluded between the social shaping and social construction of technology, and examine work done which falls under each. At the end of this section, I review and critique attempts made in the literature to resolve the tension between technological and social determinism, and offer a 'resolution' which draws upon Gadamer's conception of experience.

Mulkay (1992) describes how the new conception of science which emerges from SSK demands a re-evaluation of the relationship between science and technology:

'From the standard view of science this relationship is relatively unproblematic... effective technology is seen as a simple by-product of objective knowledge. But if we stress the socially and culturally contingent character of scientific knowledge, we must be prepared to question the widespread assumption that modern technology is on the whole a derivative of basic scientific research and/or move towards an analysis of the social meaning of technology' (p. 121).

What is significant here is that the social character of technology can now be seen in two ways: firstly, technologies are the product of socially contingent knowledge - they have been socially shaped. Secondly, since technologies exist within social and cultural practices, they have social meanings - they are socially constructed. I go on now to review in turn literature concerned with the social shaping and then the social construction of technology. (Note that this is a distinction which I
am drawing primarily to assist in the organization and analysis of the material which follows, rather than my implying that such a distinction necessarily exists.)

**The social shaping of technology**

One example of a significant body of work emerging from the 'social shaping of technology' perspective has come from research carried out in various disciplines and establishments under the auspices of the ESRC-funded Programme on Information and Communication Technologies (PICT; see Dutton [1999: v] for greater detail on PICT research). This body of work challenges technological determinism in discussions about the relationship between technology and society. Edge (1995) argues that such approaches imply:

'a linear model of the innovation process which treats technology as a "black box", and is preoccupied with the "social impacts" of a largely pre-determined technological trajectory' (p. 1).

Edge and others suggest a different view of technology, which asks questions of the origin and evolution of the technology itself, pays attention to the flexibility of the innovation process, and focuses on the choices made during this process. The emphasis is thus on the *socially embedded nature of technological development* and the *social factors which can shape the innovation process*. This approach argues that technological change cannot be fully understood solely by reference to individual inventions, and that there is a need to examine how broader social and economic forces affect the nature of technological problems and solutions. Edge (1995) suggests various ways in which social factors may act in the 'shaping' process. Social factors may influence selection between available technological possibilities; they may permit only one area of 'possible' technological development to be explored, to the extent that it becomes difficult to talk of 'alternatives'; they may operate by creating a particular environment (e.g. market) or intellectual climate where only certain technological configurations succeed; they may shape technological development by the specific embodiment of social models into the technology.
Empirical studies within this field have therefore focused on understanding better the social processes at work during the technological innovation process; for example, focusing either on providing an overview of an entire field of technology, or else on attempting to model general processes in innovation and technological development. For example, a large body of PICT research into the social shaping of tele-access has been gathered in Dutton's (1999) book Society on the Line. The very broad range of themes covered by this research has included: how information networks affect businesses (enabling the emergence of virtual organizations and assisting collaborative new product development); the nature of home-based telework and changes in the workplace; women's access to ICT and related work; how ICTs are used in the home; government policies on use of ICTs in policy implementation; use of ICTs in education, and so on.

As Mackay (1995) points out, from this perspective it is possible now to see technology as an implicitly political social phenomenon. Such an approach bears a close relationship to the 'political economy' approach of Herbert Schiller which, whilst acknowledging the increased importance of information technologies in our current era, also stresses their centrality to ongoing developments, and argues that communications technologies are foundational elements of established and familiar capitalist endeavour. For example, a 'political economy' of the Internet would look behind the information presented on websites to look at the structural features: e.g. patterns of ownership, or sources of advertising revenue, arguing that such factors constrain what information is presented. An area of interest might be the ways in which organizations can pay to increase the likelihood that their site will be listed first in a particular search engine, or the ways in which certain groups in society are excluded from access to the Net. Underpinning this perspective is an assumption that even with all the additional information and new, virtuoso technologies, the priorities and pressures of capitalism remain the same (Slevin, 2000).

Hughes (1999), in his work on Edison and the development of electric light, emphasizes the importance of incremental innovation in the development of new technology. Hughes argues that, rather than being in flashes of inspiration, Edison's real genius lay in being able to modify existing technologies, applying them to new areas, and recognizing the need for financial support and technical infrastructure to make his inventions successful; i.e. as much in a grasp of structural features as in technological skill.
Arthur (1999) offers another economist’s perspective on the innovation process, arguing that the path of adoption improves the performance of those technologies which are adopted early. Technologies are ‘path-dependent’, i.e. past events continue to exercise influence. Early-adopted technologies, adopted for whatever reason, can become permanently superior compared to those rival technologies which lose out, and which become neglected and therefore permanently inferior. The success of a technology is determined not just by its intrinsic characteristics, but also by the history of its adoption. An example of such a process might be the way personal computing has come to be dominated by the technical combination of IBM PC architecture, Microsoft’s Windows operating system, and Intel processors. In the early 1980s a dominant operating system was called CP/M; it was available on a variety of hardware platforms and had considerable market penetration. DOS was a late competitor to CP/M, but came to dominate by leveraging IBM’s powerful reach in the business equipment market. The current most popular version of Microsoft Windows (98/Millennium) is a direct enhancement of that original DOS product. Whilst CP/M was considered by users to be in many ways technologically superior at the time, it fell by the wayside while DOS was enhanced immeasurably. Moreover, many of the key drawbacks in Windows 98 (e.g. some forms of crash protection) are a direct result of architectural deficiencies in that original DOS - deficiencies that were not present in most versions of CP/M.

Caruzzi (1999) shifts emphasis to offer a cultural rather than an economic perspective of the development of the personal computer. Rather than emphasizing the changes in microchip technology, he describes the social influences which made the idea of the personal computer an acceptable one: these included a radical counterculture which wanted to free computing from commercial and military control; and a male-dominated hobby culture, which treated computing less seriously and ‘popularized’ computing as an everyday activity. It is argued that IBM’s continued commitment to mainframe computing and its adherence to the traditional view of computing as having no application for personal use was, in part, responsible for the company’s slow uptake of personal computing enabling, for example, Microsoft to emerge as the major player in the software industry.

The literature reviewed above examines the social influences on technological development. In the following section, I move on to examine other cultural accounts.
of technology which have emerged from a variety of fields, and which are concerned less with technological production than with technological consumption.

The social construction of technology

In this section, I review diverse work which can be broadly described as being concerned with the social construction of technology, i.e. with the meanings that can be made of technology. (Again, this is a provisional distinction to assist in organizing my discussion and analysis.) I shall draw on this idea further when I come to examine the meanings which have been constructed around both hypertext (in the final section of this chapter) and CAI (in chapter 4). I intend, however, to draw on this literature whilst simultaneously attempting to avoid privileging social explanations. In some of the literature I examine below, I shall argue that authors have indeed fallen into the 'mirrored' position of an overly socially deterministic account of technology. I look at the particular example of Turkle's (1984, 1988) and Herring's (1994, 2000) descriptions of women's computer use in which a doubly essentialist account - of both the computer and of social identity - comes into operation despite an ostensible focus on computer use. I critique the account of feminine identity in both writers and connect my critique back to Woolgar and Pawluch's description of 'ontological gerrymandering'. This leads into my discussion in the following section of 'soft' technologically deterministic accounts, and how a hermeneutic interpretation of accounts of technology might be made.

Media and cultural theorists such as Stuart Hall and David Morley have emphasized how design and development processes may encode preferred forms of development which are reinforced through marketing and that, in this semiological sense, one might propose that a technology is a text. The issue becomes less one of production than of reception or consumption, putting a greater emphasis on the role of the decoder of the text; that is, the user. Combined with Hall's notion of the polysemic nature of texts; i.e. that texts always have several possible readings, the ways in which users may (or may not) appropriate technologies for ends other than those intended by their creators comes to the fore. Rooted in media and cultural theory, this orientation has, in large part, concerned itself with the ways in which television programming is
received and appropriated by viewers (Fiske, 1987), rejecting preferred authorial interpretations (see, for example, Jenkins [1992] on media fandom). But the point can be generalized. For example, Turkle (1997) argues that literature on computer use has traditionally emphasized the pathological nature of computer use, and points out how the word is used for other pathological attachments (e.g. drug use). Psychological studies have emphasized forms of technophobia (Brosnan, 1998) and addiction (Shotton, 1989; Orleans and Walters, 1996). Turkle discusses elsewhere (Turkle, 1992) the ways in which images of the computer do not match the multiple ways in which computers are used in practice. As Mackay (1995) summarizes:

'Technologies facilitate, they do not determine, and they may be used in a variety of ways... The subjective, social appropriation of a technology is thus a crucial force in the social shaping of technology - one which cannot be "read off" from either the physical technology or the social forces behind its development' (p. 45).

The point can be generalized beyond technological forms to understand the notion of 'the user' more widely. A parallel theoretical development draws on work done in the field of the sociology of scientific knowledge to examine the social construction of technology. Bijker et al. (1987) developed the idea of the 'interpretative flexibility' of technology. This is described by Kline and Pinch (1999) as follows:

'Different social groups associate different meanings with artifacts leading to interpretative flexibility appearing over the artifact. The same artifact can mean different things to different social groups of users... [users] play a role in the development of a technological artifact [and...] share a meaning of the artifact. This meaning can then be used...'

---

13 This trend in media and cultural theory, which emerged largely as a critique of the pessimistic vision of 'the culture industry' inherited from the Frankfurt School, aims to examine the assumptions underlying dismissals of popular culture and the 'high/low' cultural divide. It is not solely concerned with technological forms; see, for example, Radway (1984) on romance readers and, indeed, Parker (1996) on embroidery and the 'arts/crafts' divide. Both writers have concerned themselves with the devaluing of female cultural forms; this kind of celebratory tone is evident in literature concerned with women's use of the Internet, examined later in this section.
to explain particular developmental paths. Typical groups might include engineers, advertisers, consumers, and so on' (p. 113).

As MacKenzie and Wajcman (1999a) somewhat dryly point out, this emphasis on users and social influences is not news to those engineers involved closely in designing new technologies:

'In this shift in thinking, social scientists have in one sense merely been catching up with engineers. Successful practising engineers have always known that their work is as much economic, organizational, even political, as it is "technical". They know that a design will fail if it is too expensive, if it is unattractive to employers and customers, if its 'fit' to the structure of an organization is too poor, or if it falls foul of powerful political forces' (p. xv).

For example, the field of human-computer interaction (HCI) has been of significance in computer science since the 1960s, with roots in many different engineering fields such as computer graphics, operating systems, human factors, industrial engineering and cognitive psychology. HCI work has been concerned with the design of interfaces between the user and the computer on the broadest level, with interest in a variety of aspects of this interaction, such as input/output devices, computer graphics, graphical user interfaces (GUIs), ergonomics, and so on. The field of HCI works on the basic insight that the design of an interface can have a significant amount of impact on its successful use. The interest in HCI work from social studies has grown rapidly with the emergence of the World Wide Web (Silver, 2000). Ethnographers of the Web have become interested in the design and construction of websites, the types of information presented, and the way their design affects the types of interaction made possible (Silver, 2000). As Silver notes (p. 28), writers on hypertext have similarly drawn from a wide range of disciplinary backgrounds to consider the ways in which hypertext alters the production and reception of text. Edward Barrett, in two edited volumes (1988; 1992), draws together perspectives from diverse disciplines within both computing science and social studies to emphasize both the social and technical aspects of hypertext. Articles in these volumes from an HCI perspective emphasize that the design of hypertext applications
should not subvert user expectations (e.g. Shirk, 1988). I return to a specific examination of the hypertext literature in the following section.

The emergence of cyberculture studies in the early 1990s brought a new focus on use of the Internet. Silver (2000) describes how cyberculture studies became focused on investigating the 'twin pillars' of collective communities and online identities. These two areas are respectively epitomized by two classic texts in the field: Howard Rheingold's (1994) study *The Virtual Community* and Sherry Turkle's *Life on the Screen* (Turkle, 1997). Silver goes on to suggest that a new field of 'critical cyberculture studies' is beginning to emerge, which seeks to offer 'more complex, more problematized findings' (p. 24) about using the Internet, with four broad areas of interest: the social, cultural and economic interactions that take place online; examining the stories told about such interactions; the various social, political and economic issues that encourage or thwart group or individual access to such interactions; and the deliberate, accidental and alternative technological decision and design processes which form the interface between the network and users (pp. 24-25).

Women's use of computers has attracted particular attention in the literature, and I shall now examine some of these to show how an emphasis on use of technology can nonetheless form part of a deterministic account. Drawing on Gilligan's (1982) work on the different ways in which men and women view their social worlds, writers such as Turkle (1988) have emphasized the ways in which women become 'reticent' towards computer technology 'because the computer becomes a personal and cultural symbol of what a woman is not' (p. 41). She argues elsewhere (Turkle, 1984) that there are two styles of mastering technology: one is an orderly, rational and systematic approach aimed at achieving precisely defined goals; the other is concerned more with the aesthetics of the final result than a precise blueprint, and skills are learnt through trial and error. The second style, Turkle suggests, is associated more with girls and women, and is traditionally unrewarded in the acquisition of IT skills, where mastery of specific programming tools is more valued. The difference that Turkle suggests is based on Gilligan's argument that men perceive their social world as a 'hierarchy' of autonomous positions whilst women perceive a 'web' of interconnections between people. As Turkle goes on to note, computers may become more attractive to women when perceived as supporting communication through networks (Spender, 1995).
Herring (1994, 2000), also drawing on Gilligan, examined computer-mediated communication and suggested that there are two different value systems at play: an 'anarchic/agonistic' male system that privileges freedom from censorship and debate, and a female 'politeness ethic' distinguished by 'expressions of support and appreciation', a dislike of flaming, and a tendency to equivocate when offering opinion. The 'masculine' anarchic/agonistic style dominates CMC, Herring suggests, 'with the result that women with a politeness ethic must create and defend women-centred spaces on-line in order to carry out the kinds of discourse they value'. Clerc (1996), drawing on this work in a study of online fandom, suggests that women fans prefer mailing lists to newsgroups since they seem 'safer, more civil places... to post and promote a greater sense of community. These features appeal particularly to women' (pp. 42-43).

I would argue that these are doubly romanticized discourses of women on the Net, which are based on a univocal conception of female communication and which project onto the Internet an idealized vision of its potential as an unconstrained sphere of more authentic community: the technological 'agora' eulogized by Rheingold (see above, p. 54). However, in considering claims such as these, one could be drawn very quickly into an opposing 'constructivist' argument which denies the essentialist emphasis on women's difference. To my mind, a more satisfying critique is offered by Fuss (1989) who argues that:

'[E]ssentialism is neither good nor bad, progressive nor reactionary, beneficial nor dangerous. The question we should be asking is not "is this text essentialist (and therefore 'bad'?)) but rather, "if this text is essentialist, what motivates its deployment?"' (p. xi; author's emphasis).

Here, Fuss is asking a kind of question about the rhetorical purposes of truth claims which is similar to those questions that I posed in chapter 2 (pp. 13-14). Fuss continues:

'[T]he deadlock created by the long-standing controversy over the issue of human essences (essential femininity, essential blackness, essential gayness...) has, on the one hand, encouraged more careful attention to cultural and historical specificities where perhaps we have hitherto
been too quick to universalize them but, on the other hand, foreclosed
more ambitious investigations of specificity and difference by fostering
a certain paranoia around the perceived threat of essentialism' (p. 1).

An emphasis on essential differences can easily be turned into a
marginalization of women's experience\(^{14}\); however, a rigid adherence to the opposite
stance reminds one of the ontological sleight of hand which Woolgar and Pawluch
critiqued (see pp. 41-42 above), in which an argument based on apparent relativism
holds certain elements of its argument constant and unexamined whilst varying others.
In the following section, I critique a number of theoretical approaches which have
attempted to tread a 'third way' between either technological or social determinism. I
then turn to consider the debate that emerged in social studies of science and
technology in the wake of the development of actor-network theory (ANT) which,
like Fuss' account of human nature (in her example, gender and, in the case of ANT,
technological capability) requires a scepticism about the possibility of coming to a
conclusion about what determines nature or culture.

*The soft option*

Theories of the relationship between technology and society which emphasize the way
in which technology is socially influenced are most frequently criticized on the
grounds that they become deterministic in reverse - technology is socially determined.
To counter this, writers frequently adopt a theoretical approach which can be
described as 'soft' technological determinism, an approach which conceptualizes
technology and society as being mutually constitutive, what MacKenzie and Wajcman
(1999b) call 'the reciprocal relationship between artifacts and social groups' (p. 22).

Hughes (1994) has coined the phrase 'technological momentum' to describe a
phenomenon which he noted in a variety of empirical studies of the emergence of

\(^{14}\) And can also have deleterious consequences. Instructive here is the analysis offered by Kosofsky
Sedgwick (1990), in her account of the dangers of denying essentialist accounts of gay origins: 'Every
step of this constructivist nature/culture argument holds danger: it is so difficult to intervene in the
seemingly natural trajectory that begins with identifying a place of cultural malleability; continues by
inventing an ethical or therapeutic mandate for cultural manipulation; and ends in the overarching,
hygienic Western fantasy of a world without any more homosexuals in it' (p. 42).
technological systems, that younger, developing systems tend to be more open to sociocultural influences, while older, more mature systems prove to be more independent of outside influence and therefore more deterministic. He explicitly views this idea as a way out of what he sees as an impasse between technological determinism and social constructionism:

'A technological system can be both a cause and an effect; it can shape or be shaped by society. As they grow larger and more complex, systems tend to be more shaping of society and less shaped by it. Therefore, the momentum of technological systems is a concept that can be located somewhere between the poles of technical determinism and social constructivism. The social constructivists have a key to understanding the behavior of young systems; technical determinists come into their own with the mature ones. Technological momentum, however, provides a more flexible mode of interpretation and one that is in accord with the history of large systems' (p. 112).

Golding (2000) develops a short-hand taxonomy of technologies:

'We can conceive of two forms of technological innovation. Technology One allows existing social action and process to occur more speedily, more efficiently, or conveniently (though equally possibly, with negative consequences, such as pollution or risk). Technology Two enables wholly new forms of activity previously impracticable or even inconceivable' (p. 171).

Developments in biotechnology, he suggests, might constitute a 'Technology Two'; they may 'presage real change in what human action and activity might obtain and pursue' (p. 172). But, he argues, '[i]n essence, many new ICTs are more obviously Technology One than Technology Two' (p. 171). There is an attractiveness to these approaches offered by both Hughes and Golding, since they appear to offer a 'third way' between determinism and relativism. My own dissatisfaction lies less, perhaps, in the determinism in both accounts than in their shift from specifics to generalities; i.e. from examinations of the circumstances of particular cases to the development of
large-scale explanatory models for technological innovation: in Hughes, the
description of a kind of law which claims to hold true for all 'technological systems';
in Golding, a taxonomy which provides a neat empirical model for the construction
and investigation of hypotheses.

I want now to consider what I believe is a different case of soft technological
determinism, by looking at debates surrounding work done by Bruno Latour and
colleagues under the name of actor-network theory (henceforth ANT). ANT expands
from Bloor's principle of symmetry (i.e. that the same types of cause explain both true
and false beliefs) by adopting a semiotic approach which leads to a more radical
symmetrism. ANT treats all dichotomies, not just that between true and false,
symmetrically. A consequence of this is that all 'actants' (whether natural, human, or
technological) are treated with equal significance in ANT studies of science and
technology. Moreover, ANT sees such dichotomies (e.g. technology vs. the social) as
constituted in the world on an ongoing basis (allowing us to see, e.g., 'nature' and
'society' not as causes but as consequences; as outcomes rather than determinants).

Latour's approach has been controversial within the field of social studies of
science. The controversy has been documented in chapters by Collins and Yearley,
Woolgar, and Callon and Latour in Pickering (1992), but the broad thrust is as
follows. Collins and Yearley argue that, under the disguise of increased radicalism,
ANT's granting of agency to things is in fact a reversal into technological determinism
(p. 312). Callon and Latour respond by arguing that Collins and Yearley have fallen
prey to a determinism which privileges the social (p. 348) and hence their (Collins and
Yearley's) own status as social scientists (p. 358).

In the final analysis, the most interesting aspect of this controversy must be the
debate itself. With both sides firmly convinced of their superior radical credentials
and each providing solid critiques of the other approach, one is hard pressed to judge -
and indeed left with the conclusion that such a choice is unnecessary. I shall explain
my point further by focusing not solely on the points of difference between the
approaches, but also on the similarities.

The 'family resemblance' between the two approaches lies in their sensitivity
to and even suspicion towards claims of more accurate knowledge of the world - and
the claims of both have persuasiveness. ANT, in its reconfiguration of the 'nature' of
the 'actor' highlights an anthropocentrism in Collins and Yearley's version of SSK, a
'human essentialism' in the discourse which has not previously been scrutinized. In
Woolgar and Pawluch's terms (see above, pp. 41-42), Collins and Yearley have fallen into a selective relativism by holding 'human' nature constant in their account, in opposition to the non-human. But the price of releasing this particular constant in the formulation offered by ANT is, as Collins and Yearley point out, that other variables are held constant.

The debate, if conceived in this way, rather than as one in which either side must necessarily emerge as the standard-bearer of 'truth', having the more accurate picture of the world\textsuperscript{15}, takes on the aspect of a hermeneutic experience - the debate and the analysis themselves become generative of knowledge (see my discussion of this in relation to Gadamer; chapter 2, p. 15). ANT demands a reappraisal of certain tenets of Collins and Yearley's version of SSK which, in its turn, already contains a reciprocal critique of ANT. The 'answer' to questions regarding the 'nature' of technology or society or of 'actants' thus remains open. actor-network theory has opened up an interesting and contemporary source of inquiry that offers a challenge to the tenets of SSK, which, given the strong programme's commitment to reflexivity, can only be welcome\textsuperscript{16}. (I shall bring together an ANT perspective and a hermeneutic emphasis on the text in my analysis of my own project in chapters 7 and 8.)

In the final section of this chapter, I shall bring together the themes which have emerged in these earlier sections to consider the uses made in academic accounts of a particular technology: hypertext.

\textsuperscript{15} To clarify: I am leaving these questions open, rather than implying that they are not answerable. See my discussion of relativism in Mannheim's account of the sociology of knowledge (chapter 2, pp. 19-20) which does not logically exclude the possibility that a particular assertion may be shown to be universal in its 'scope and limits'.

\textsuperscript{16} And has implications for political debates such as animal rights and environmentalism; for example, work Latour has done with primatologist Shirley Strum (Strum and Latour, 1999). Drawing on studies of primate (here, baboon) societies, Strum and Latour point out that primatologists increasingly view primate societies in ways similar to sociologists looking at human societies (i.e. baboons actively and continuously negotiate their relative roles, and from these processes social structure emerges). They argue that in order to differentiate between primate and human societies one needs to take into account the technological, since this is what enables the human use of materials and resources that makes our societies operate beyond the face-to-face interactions which characterize primate societies. Descriptions of human society, they suggest, necessarily involve reference to technology.
4 The development and reception of hypertext

Introduction

As Barrett (1992) has pointed out, hypertext technologies (such as HTML), when used across the Internet or other computer networks, are technical means of communication: his neologism 'sociomedia' is an intentional combination of both the social and technical aspects of hypertext technologies:

'hypertext is an embodiment in a machine of the social construction of knowledge in the human domain of thought and language' (p. 9).

Hypertext is the point of intersection between discussions related to the social impact of new technology and, also, the epistemological impact of different modes of representation on information.

This section will focus on the emergence and reception of hypertext, and draws on the conceptual distinction made in the previous section between the social shaping and the social construction of technology. My discussion of the social shaping of hypertext will review the history of its development, and the convergence of advances in communications and GUI technologies which gave birth to the World Wide Web. When I turn to the social construction of hypertext, I examine theoretical accounts and taxonomies of hypertext which have emerged primarily from the humanities and in qualitative social science. I review and critique this literature. This section therefore functions both as literature review and also to introduce theoretical accounts and taxonomies of hypertext which I drew upon extensively in my own development work to assist in design decisions and to formulate academic questions (specifically in the first stage of the project; see chapter 7, pp. 206-207).
History of hypertext and the World Wide Web

The term hypertext\(^{17}\) describes a method of presenting information in which text, images, sounds, and actions become linked together in a complex, non-sequential web of associations that permits the user to browse through related topics, regardless of their presented order. These links are often established both by the author of a hypertext document and by the user, depending on the intent of the hypertext document.

The term hypertext itself was coined in 1965 by Ted Nelson to describe documents, as presented by a computer, that express the non-linear structure of ideas, as opposed to the linear format of books, films, and speech. However, the concept of a mechanical web of information linked by association rather than selection predates Nelson by two decades. In July 1945, Vannevar Bush, a professor at MIT who had been associated with the development of the computer, outlined his vision of a machine that would allow access to the sum of human knowledge. This microfilm/audio recording device, which he called a memex, would allow '[s]election by association rather than by indexing' which, he argued, more closely mirrored the workings of the human mind:

> "When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. It can only be in one place...
> The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thought, in accordance with some intricate web of trails carried by the cells of the brain" (1945: 105).

The memex would thus provide a mechanical means of mirroring the associative selection and indexing patterns of the human mind:

\(^{17}\) Although the two terms 'hypertext' and 'hypermedia' can refer to, respectively, text-based and multimedia-based hyperlinked information, the term 'hypertext' will be used in its broadest sense in the discussion that follows.
[The memex] affords an immediate step, however, to associative indexing, the basic idea of which is a provision whereby any item may be caused at will to select immediately and automatically another. This is the essential feature of the memex. The process of tying two items together is the important thing... When the user is building a trail... [i]t is exactly as though the physical items had been gathered together to form a new book. It is more than this, for any item can be joined to numerous trails' (p. 106).

Although hypertext-based software applications were developed and became increasingly popular during the 1970s and 1980s, it was not until the early 1990s that something approaching Bush's vision emerged. This depended on two distinct but related developments: firstly, the expansion of the Internet and, secondly, the creation of an easy-to-use graphical user interface (GUI).

The Internet started in the late 1960s as a US military experiment in the construction of robust computer networks that could withstand nuclear attack. This network, called the ARPANET18, was a great technical success, but limited in scope. In 1983, the military felt the technology was stable and mature enough to be useful, and separated the ARPANET into two halves: MILNET (an operational network used by the military), and the other half was left for research by universities. In 1984, since connection to the ARPANET was dependent on being part of a military-funded research project, the US National Science Foundation started building a successor to the ARPANET, called NSFNET. During the 1980s, the ARPANET's role as a backbone linking other networks was gradually taken over by NSFNET, and the ARPANET was decommissioned in 1990 (Slevin, 2000).

NSFNET's initial exclusion of commercial traffic encouraged the growth of competitive private backbone networks, and it was not until the National Science Foundation's decision in the late 1980s to privatize key parts of its network operation that the Internet was opened to these commercial networks. In 1995, NSFNET was shut down completely, and most internet traffic is now carried by commercial networks (Slevin, 2000). However, use of the Internet in the late 1980s remained limited: the software to access and navigate one's way around the information

---
18 Advanced Research Projects network of the Department of Defense.
available was arcane, the information itself disorganized (Musciano and Kennedy, 1998).

The real explosion of the Internet came in the early 1990s when a user-friendly means of accessing and navigating one's way around the network finally became available (Musciano and Kennedy, 1998). Physicists at CERN (Europe's advanced atomic particle accelerator), notably Tim Berners-Lee, released an authoring language and distribution system which they had developed for creating and sharing multimedia-enabled, integrated electronic documents over the Internet called Hypertext Markup Language (HTML). The World Wide Web, using HTML documents, presented to the user unified text, pictures and sound that previously had appeared as fragmented items. Significantly, the Web allowed hypertext linking, so that documents located anywhere in the world could be connected together (Cailliau, 2000). Vannevar Bush's dream of a mechanical web which linked information by association seemed to have become a reality (Winston, 1998).

Hype and hypertext: the reception of hypertext by the humanities and social science

The fascination and enthusiasm with which hypertext has been greeted by the humanities lies in the apparent solutions it offers to the analytical constraints of formal computer applications, by virtue of being a communications - and hence social - medium. Where earlier software had dealt in organization, hierarchy, and categorization, hypertext appears to offer users the chance to organize texts in a less formalized, more associative fashion, contextualize information, and emphasize the relationships between bodies of knowledge. It is no surprise, therefore, that hypertext applications have been enthusiastically received by qualitative social researchers, and by literary and textual theorists, for whom textual analysis is at the heart of the research process. In this section, I discuss the accounts of hypertext which have been given by, firstly, qualitative researchers and, secondly, literary and textual theorists.

19 It is worth considering that the World Wide Web was based on existing formats: HTML was a simple application of an existing standard called SGML; pictures were handled with existing formats (GIF); and it used the existing Internet protocol 'TCP' for its communication, adding a very simple layer (HTTP) on top to support the HTML application. HTML unified these disparate source materials upon presenting them to the user. HTML was thus an incremental innovation based on existing technologies, rather than a significant departure from what had gone before (see Hughes [1999] on the incremental nature of innovation), rather than a radically new technological advance.
While these are distinct disciplinary starting points, the accounts of hypertext which have emerged follow patterns which I have described more generally in section 2 of this chapter. Discussion in the field of qualitative research has predominantly involved consideration of either optimistic or pessimistic predictions of the potential effects of the technology on the quality of data. Literary theorists have, in general, produced teleological accounts of the history of writing and knowledge in which hypertext is positioned as a significant agent of change. My intention in this section is not to conflate these distinct fields but to consider the assumptions brought to their accounts of hypertext. (I look at an example from quantitative survey research when I discuss computerization and the Web in relation to the FRS in chapter 4.)

Hypertext, it is suggested, at last offers a technical means of scrutinizing and ordering information in a way which emphasizes the situated nature of knowledge. Barrett (1992: 9) describes the appeal of hypertext for the humanities as based on its apparent support of all the functions that define social construction: the collection and classification of texts; the review and deconstruction of texts; the exchanges of texts among peers; the empowerment of the individual through the ability to create marginalia. Moreover, hypertext applications can do these things quickly, with varying degrees of thoroughness, and with documentation of all interactions with a text or another individual. However, one is left with the impression, from this account, of this being happy chance: as if, having conceived of social constructionism, theorists are now in the lucky position of being able to point to a technology which apparently offers support of this theory. There is little awareness of the possibility of reverse causation here: that the implementation of the technology is influenced by users' theoretical leanings and expectations. It is interesting also, rhetorically, that a legitimization based on technology is felt to be needed; perhaps this rhetorical manoeuvre also acts as an assertion of disciplinary expertise in a technical domain - an attempt at boundary crossing and even colonization?

The discussion of the potential benefits of hypertext applications for qualitative social research has taken place within the context of a broader debate about the use of computer applications in qualitative research and, as such, follows a recognizable pattern of debate in the face of new technology: a optimistic tone which seeks to embrace the potential of the new tool, countered by a sceptical fear characterized by Hunter (1999) as accompanying the inception of all technologies, that 'wisdom will be superseded by a patina of information' (p. 94). Fielding and Lee
(1998) are enthusiastic about the potential of computer software in assisting the research process, and sum up the main advantages in terms of managing better data of multi-stranded, multi-sourced forms. They are, however, wary that 'technologies can interpose themselves in ways which serve to distance the researcher from the data' (p. 1). Fischer (1994: 199) notes that computers can encourage the use of research procedures because they are easy rather than because they are appropriate.

Weaver and Atkinson (1994: 10-11) combine a pessimism about the general use of analytical software with a proselytizing enthusiasm for hypertext. Speaking generally about software for qualitative analysis, they suggest that such applications do not reflect 'messy' reality, and argue that:

'[T]he uncritical adoption and implementation of microcomputer software - or indeed the wholesale endorsement of the general approach - may commit the researcher to an implicit and uncritical adoption of particular analytic strategies' (p. 1).

Moreover, they go on to suggest that the introduction of computers into qualitative research may be tied up with issues related to the legitimization of data: it is bound up with attempts to 'clean up' the tradition, making it more systematic, standardized, and generally making its knowledge claims more acceptable to the scientific community (p. 16).

Nevertheless, they remain very optimistic about the possibilities of hypertext software, which they see as the most superior analytical software available to qualitative researchers. Other forms of software, such as coding software, lose their contextual information, whereas hypertext, they suggest, is more flexible and dynamic, and encourages reflexive modes of thinking. Echoing Vannevar Bush, they draw a specific link between the way knowledge is constructed in a hypertext database and the way it is organized in the human mind, concluding that unlike traditional print media, where information is ordered in a linear, unidimensional way, hypertext systems complement the human thinking process. They conclude:

'Our research suggests that, because the ideas and trails of a researcher themselves become 'objects' in the same way as data, hypertext encourages thinking that is much more reflexive than that encouraged
by other strategies. Researchers are encouraged to analyze and question their own idea, and the emerging construction of knowledge, in the same way that they do their data. The issue of reflexivity is bound up with the broader question of how different transformations of data make researchers think differently about the text' (p. 160).

It is left to Barry (1998) to provide the contrary - and expected - pessimistic view, that hypertext's lack of structure and too much flexibility could lead to cognitive overdrive, a 'patina of information' superseding wisdom. But Barry's note of caution does not recast the terms of the debate: hypertext technology, it is assumed, produces information which is multi-linear and non-hierarchical: the technology is imbued with an essential nature.

Moreover, since hypertext is both a technology and a communications medium concerned with the organization of knowledge, hypertext is given the capability to fix the nature of the knowledge which it represents, and is thus positioned as a driving force in social change. George Landow, one of the most influential writers on the convergence between critical and literary theory and hypertext, explicitly aligns the emergence of hypertext with a form of theory which emphasizes a decentralized, de-authored vision of the text:

'[M]any... who write on hypertext or literary theory argue that we must abandon conceptual systems founded on ideas of center, margin, hierarchy, and linearity and replace them with ones of multilinearity, nodes, links, and networks' (1992: 2).

The assumption is thus made that hypertext has the potential significantly to alter our culture and society; that this new technology can, in some way, act as a means of social change. Hypertext is positioned as the next agent of epistemological change in a trajectory which began with oral communication and has passed through the printed book to the current position. Landow sums up:

'Electronic text processing marks the next major step in information technology after the development of the printed book. It promises (or threatens) to produce effects on our culture, particularly on our
literature, education, criticism and scholarship, just as radical as those produced by Gutenberg's movable type' (p. 19; my emphasis).

The emphasis draws attention to the way, once again, that debates about the social impact of new technologies are cast in terms of optimism and pessimism - this imbues these technologies with a fixed nature, rather than emphasizing the extent to which social influences can impact upon both the development and the application of new technology. Similarly, Bolter (1991) offers a teleological overview of the history of writing from the papyrus roll to hypertext that owes a great deal to McLuhan, and sees a profound shift taking place in our concept of the text, which moves between optimism and pessimism, and assumes electronic modes of representation will have specific effects:

'[The] shift to the computer will make writing more flexible, but it will also threaten the definitions of good writing and careful reading that have been fostered by the technique of printing... [T]he printing press encouraged us to think of a written text as an unchanging artifact, a monument to its author and its age... [it] also tended to magnify the distance between the author and the reader... Electronic writing emphasizes the impermanence and changeability of text, and it tends to reduce the distance between author and reader' (pp. 2-3).

His central point is that hypertext and the shift to electronic publishing necessarily imply a very different form from the printed book and challenge some of our basic assumptions about the organization and presentation of text:

'a hypertext is like a printed book that the author has attacked with a pair of scissors... the difference is that the electronic hypertext does not... dissolve into a disordered bundle, as the book would, because the author defines a scheme of connections to indicate relationships' (p. 24).

Bolter's enthusiasm is tempered by a recognition of the potential hazards of hypertext which calls to mind the points made by Barry (1998). Bolter notes (p. 67)
that the impermanence of the electronic image discourages attention to detail. In a somewhat prescient remark (bearing in mind that the World Wide Web was only invented towards the end of 1990, and was not widely used or known), Bolter notes that:

'there is an inevitable degeneration in the quality of typography and graphics in the new electronic writing space, because the computer encourages the democratic feeling among its users that they can be their own designers' (p. 66).


- **Intertextuality.** The overt intertextuality of Web texts sets them apart from traditional texts: 'unlike printed texts, the virtual digital text offers the opportunity to connect various virtual texts with specific "links" that allow readers to move from one to another' (p. 182).

- **Non-linearity.** Traditionally, texts have been characterized as having a beginning and an end. Hypertext challenges linearity in two ways. Firstly, 'the fundamental proposition that there needs to be a beginning and end is problematized' (p. 185). Secondly, hypertext is overtly intertextual, 'constantly inviting the reader to move to another textual node' (p. 186).

- **The reader as writer.** Rather than the traditional role of the reader as an 'audience', the non-linear, intertextual text asks the reader to become more active in the reading process. Readers choose or 'write' their own way through hypertext documents.

- **The multimedia WWW.** Web authors can use a variety of representational strategies: documents can appear in text, video, sound or multiple formats.

- **The global WWW.** The Web spans the entire globe: 'this explosion of connectivity... has not only reshaped the way humans can communicate over long distances but also expanded the access to texts that can be used to produce a site on the WWW' (p. 189).
• **The disappearing hypertext.** Since Web texts exist primarily on computer systems and have (generally) no tangible form, they can easily disappear: either through links being removed, files being deleted, or content being changed. Web texts are 'ephemeral and impermanent' (p. 191).

They conclude:

'In summary, the WWW text has a set of specific characteristics *predicated by its hypertextuality* (1999: 192; my emphasis).

While it is arguable that hypertext, as constructed in this body of literature, is less technologically deterministic because it allows for flexibility of use, it is my contention that these analyses are substantially dependent upon imbuing hypertext technologies with a specific, fixed nature. Such a conception of hypertext is significantly flawed. By focusing primarily on the impact hypertext can have on the construction of knowledge rather than on, for example, the designer's role in constructing applications, these analyses essentialize the contextualizing capabilities of hypertext technology. Moreover, hypertext applications are constructed in this literature as providing a means of accessing and creating information which, because of its contextualized nature is - paradoxically - seen as somehow more authentic. The humanities, chronically mindful of the truth-claims of scientific discourse which emphasize objectivity, have turned to hypertext to construct a defence of context - but it is a defence in which technology is seen as a legitimating force, and which reinforces a basic opposition in which we are invited to see the associative and the contextualized as somehow more 'authentic' than the hierarchical and the objective. Moreover, by granting hypertext technology this essential nature, the door is left open to technological determinism: hypertext is given the ability to fix the nature of knowledge; it is removed from the social sphere and granted special status as an agent of social and epistemological change.

An example of an existing website and its associated community which displays almost all of the criteria set out by Mitra and Cohen is the WikiWikiWeb [http://c2.com/cgi/wiki?WikiWikiWeb]. The Wiki is an online community of professional software engineers, consultants, and academics and students in the area of computer science. The community aims to disseminate knowledge and expertise in the specialist
area of pattern programming. The Wiki website has been built (using Perl CGI
scripts) in such a way that community members are able and, according to the
practices of the community, actively encouraged to submit and alter existing text. In
addition, community members expect that their own submissions are open to
alteration from others. Members alter the content of other people’s contributions, and
the aim is the production of mutually agreed upon documentation. The site is, in
effect, in a process of continual peer review; moreover, this is not just in terms of its
content, but also in terms of its operation.

One could argue that the Wild demonstrates admirably the value of Mitra and
Cohen’s taxonomy: it is intertextual, non-linear, the user can both read and write (in
real time), the text is ‘ephemeral and impermanent’. Nonetheless, other examples of
online environments for the dissemination of expert knowledge are organized quite
differently from the idealized version of hypertext offered by Mitra and Cohen. Many
online journals (for example, amongst many, see Signatures
[http://www.ucc.ac.uk/signatures/] or Intensities [www.cult-media.com]), substantially
resemble their paper-based predecessors, with calls for submission, formalized peer
review processes, and regular publication of issues. An editorial which appeared in
Science sets out the case for such an approach in electronic publishing of academic
journals:

‘The current practice of peer-reviewed journals ensures that published
results have been carefully scrutinised and provide a level of assurance
of the quality of these results on which future research can be based.
We must face the challenge of providing new mechanisms to ensure the
same level of quality control in electronic publishing without
sacrificing the advantages of rapid dissemination’ (Winograd and Zare,
1995: 625; my emphasis).

When set alongside the example of the WikiWikiWeb, it becomes apparent
that the divergence of these journals from Mitra and Cohen’s classification is not
because of constraining factors in the available technology. Moreover, bearing in
mind the importance attached in the quotation above on the provision of new
mechanisms (which, presumably, given the emphasis on peer review, are at least in
part social mechanisms), one is led to ask whether the significant factor is not so much in features of hypertext technology itself but, rather, in its deployment.

Taking this as her starting point, Hunter (1999) has developed a significant critique of technological determinism with particular reference to hypertext. She challenges the assumption that hypertext applications have an inherent nature by emphasizing instead how hypertext applications are socially shaped and constructed. She argues:

'\[N\]o technique is enclosing, isolating and reductive, or exploratory, contextualising and flexible, in itself; nor is either authenticity or self-reflexiveness in itself enabling. Communicative texts from all disciplines need a rhetorical analysis of stance, which will position the techniques and strategies historically, politically and socially. Such an analysis situates the textuality, and in so doing situates the knowledge'
(p. 6).

Hunter draws extensively on her own experiences of developing and using hypertext for primarily literary and historical applications. She specifically seeks to counter the notion that hypertext is somehow implicitly flexible, relative, and non-hierarchical (pp. 110-11), arguing that hypertext applications are often based on information chosen by their designer and ordered in a structured and often hierarchical fashion. To this end, she has developed a classification of hypertext applications using case studies of projects as examples.

Hunter suggests that hypertext projects can be grouped under four approaches: topic driven hypermedia texts, central text hypermedia, multi-document hypermedia, and hypermedia nests (p. 113).

- **Topic-driven hypermedia texts** are best seen as providing a multi-dimensional filing system for existing information which is already highly categorized and hierarchical.
- **Central text hypermedia** provides an informational shell around a central text, a text which may be a person/writer or a literary artefact.
• Multi-document hypermedia presents a multiple document archive, using hypertext predominantly as a presentation device.

• Hypermedia nests contain a small number of closely related, multiply linked texts.

This contrasts with the taxonomy developed by Mitra and Cohen by focusing on the application (which necessarily implies a designer or designers) rather than on seeking to classify distinctive features of the technology itself.

While Hunter's purpose is to develop an account of hypertext technology which is freed from the technological determinism apparent in such accounts offered by Landow, Bolter, etc., I would suggest that in her own development of a classificatory system, her argument begins to move from a 'rhetorical analysis of stance' into a form of social determinism. Her focus shifts from an attempt to 'position the technique', i.e. to investigate the use to which technology is put discursively in order to support the knowledge claims of the social sciences (what I have called the construction of technology), and into classifying types of technology according to decisions made by the developer or designer during the process of production. Hunter's schema, while indeed avoiding the technological determinism she seeks to critique, reverts to an unacknowledged social determinism and to an unexamined essentialism which sits uneasily with her emphasis on stance and situatedness. In Woolgar and Pawluch's terms, her account holds constant, as a determining, causal factor, the purposes of a technology's designer, in order to demonstrate that the features of a technology can change.

The clue to the presence of this 'mirror image' version of hypertext lies, indeed, in Hunter's construction of such a classificatory system at all. The purpose of taxonomies is to describe phenomena in the world about us, to offer a 'best' description of how the world is. Having been organized in such a fashion - whether through a simple model such as Mitra and Cohen's description of the distinctive features of a technology, or a more complex description which relies on prioritizing human intervention in the design process - the world or, in this particular case, technology, can be manipulated. Taxonomies provide the variables needed for the construction of hypotheses and the investigation of explanatory systems of the world.
In this section, I have reviewed and critiqued theoretical accounts of hypertext which have emerged from the humanities and in qualitative social science. In chapter 4, I shall move on to quantitative survey research and issues surrounding hypertext and the Web with particular reference to the FRS. However, the material which I have discussed in this section is significant for my project beyond a review: the academic accounts of hypertext which I have described were influential in the research questions which I was formulating during the development of the FRS website, particularly in the first stage of the project. At this stage, I was concerned with examining how the taxonomies of hypertext worked in practice, and whether or not they applied in the case of the FRS documentation website. I shall discuss this in greater detail in chapter 7.

5 Summary

In this chapter I have reviewed several bodies of literature, and also introduced a number of themes and concepts that will be significant for the description of my own project and analysis that will follow.

The overarching purpose of the critique in this chapter was to examine theoretical accounts of the relationship between technology and society from a relativist standpoint. I contrasted technological determinist and social determinist positions; in particular, I reviewed theories of the information society and critiqued them, in particular, for the tendency to describe societies in terms of shifts between epochs, and for the way technologically deterministic accounts shift discussion to irreconcilable visions of the future; I applied a similar critique to the theoretical accounts of hypertext that have emerged from the humanities and qualitative social research. In addition, I have reviewed literature that examined social influences on technology, organizing my discussion around a conceptual distinction between the social shaping and the social construction of technology. I reviewed perspectives from cultural and media studies which treat technologies as texts, and I introduced actor-network theory as a potentially fruitful means of analyzing technology.

As well as a review and critique of literature, however, this chapter has served to introduce several themes and concepts which will continue to have relevance throughout the remainder of this thesis.
I shall return to theories of the information society in the following chapter, when I discuss how they relate to UK government e-Strategy, with particular reference to the online dissemination of survey material. Furthermore, in chapter 7, I shall examine how information society narratives were an important influence on the decision to present FRS documentation online - a decision which was taken before I began work on the project, and which bounded the technical scope of the project in certain ways.

As I noted above, the academic accounts and taxonomies of hypertext which I reviewed and critiqued in this chapter were of great significance throughout my project, particularly at the outset. During that first stage of the project, I attempted to formulate academic questions which would enable me to examine those accounts which imbued hypertext with an 'essential' (non-linear and associative) nature, and I tested the appropriateness of the various models of hypertext offered in the literature as I was building the site. I eventually abandoned this approach, but the literature on hypertext which I reviewed in this chapter was a significant influence upon the site, as I discuss in detail in chapter 7.

The conceptual distinction which I have drawn between the social shaping and the social construction forms the basis for my organization of the two chapters (7 and 8) which are concerned with my analysis of the social process of the project to develop online documentation for the FRS. Throughout these chapters, I shall continue to map the social shaping and the social construction of technology onto the production and reception of texts; I shall also, in these analytical chapters, draw further upon an actor-network perspective.

In the next chapter, I turn to issues related to those technologies used in quantitative social survey research that are particularly relevant to the FRS, i.e. CAI and the Internet. I shall discuss the historical and contemporary organization of government statistics and the increased computerization of the survey research process in order to give the background within which the FRS and my project to provide online documentation are located. Chapter 5 provides a detailed description of the FRS, and chapter 6 a detailed description of my project. I should note at this point that in order to describe fully the FRS, my project and the background to both, I shall be altering the tone of my discussion and, to some extent, suspending my critical voice. As I noted above, chapters 7 and 8 will provide a critical analysis of the social process of my project to provide online documentation for the FRS.
4 Uses of technology in social survey research

1 Introduction

In this chapter, I describe the context in which the project to provide online documentation for the FRS is embedded. The present study emerges from the introduction of the laptop computer into large-scale survey research, the consequent shift away from paper questionnaires to computer-assisted interviewing (CAI), and the ramifications of this for documentation. In this chapter, then, I draw together a number of organizational and technological issues to give a detailed description of the wider background within which the documentation website can be located.

I begin in section 2 by giving information on the historical and organizational setting of the FRS, providing an overview of UK government social survey research in the post-war period. I describe the organization and administration of government statistics with particular reference to the bodies relevant to the FRS (i.e. the GSS, the ONS, the Information and Analysis Directorate at the DWP). I then discuss those technologies which play a part in the production of the FRS, and were significant in the emergence of the current project (i.e. the Web and CAI). One of the earliest design decisions made for the FRS questionnaire documentation was that it should be accessible via the Web (see chapter 6, p. 145), and this decision was, in part, related to current policy initiatives to increase the amount of government activity conducted online. At the end of section 2, then, I give an overview of the emergence and current state of government e-Strategy, particularly in relation to the dissemination of social survey research output. I relate these policy initiatives directly to the theories of the information society which I discussed in chapter 3.

In section 3, I review literature concerned with the introduction of computer technology into the social survey research process. This puts the FRS into context as the first large-scale survey conducted by UK government to be CAPI-based from its inception. I give an historical overview of the introduction of laptop computers into the survey research process, and review the body of literature concerned with examining the effects of this technological intervention into the survey process. My emphasis here is on the particular concern of my own project, i.e. the literature concerned with documenting electronic questionnaires. At the end of this chapter, I
shall have given the technological and organizational context for the FRS. I describe the FRS itself in detail in chapter 5.

2 Government social survey research in the UK

Historical and organizational overview

The Family Resources Survey is a large and complex multi-level survey into the living standards and characteristics of around 22-24,000 households carried out each year. It is the first major UK government social survey to be CAPI-based from its inception. The FRS is managed by the Department for Work and Pensions (DWP) Information and Analysis Directorate Income Analysis 1 (IAD IA 1). The contract for conducting it is currently held by the Office of National Statistics Social Survey Division and the National Centre for Social Research (NatCen), and has been since the first full survey year (1993-1994). In this section I shall set the FRS in its historical and organizational context within UK government social survey research. I begin with a short history of social survey research in the post-war period, then provide an overview of the nature and extent of survey research at the time of writing. My emphasis is on the two main government bodies involved in producing the FRS: ONS Social Survey Division and the Information and Analysis Directorate (formerly the Analytical Services Division) at the DWP; I also provide some information on the third player in the FRS: NatCen. My account draws upon internal and online histories and organizational accounts prepared by these various bodies (e.g. Nissel, 1987; Moss, 1990; National Statistics, 2001; online material is referenced in footnotes).

Government social surveys in the UK in the post-war period have been conducted under the aegis of three different departments, in three periods: from 1941 to 1970, when such surveys were conducted by the Government Social Survey; from 1970 onwards, when the Government Social Survey was merged with the General Register Office to form the Office for Population, Censuses and Surveys (OPCS); and from 1996 onwards when the Central Statistical Office merged with OPCS to form the Office of National Statistics (ONS).

20 The FRS was initially conducted in Great Britain only, although from 2002-2003 it includes Northern Ireland, making it a UK-wide survey.
The Government Social Survey (GSS) emerged from the Wartime Social Survey; Nissel (1987), while not clarifying the way in which utility is measured, states that this survey 'came into being as a wartime expedient, but then survived into the post-war period because it had shown that it was too useful to do without' (1987: 86). Initial surveys were concerned with the effects of wartime rationing on the availability of foodstuffs and on the health of the population; these latter expanded the activities of Social Survey into other health surveys. The Wartime Food Survey, which later was renamed the National Food Survey (which is now part of the Expenditure and Food Survey), is the oldest regularly conducted government survey. Towards the end of the war, Social Survey became involved in providing surveys to support the implementation of the post-war welfare state set out in the 1942 Beveridge Report and the 1944 Education Act. The 1944 Survey of Sickness was the first continuous survey begun by Social Survey; at this time there were also surveys of public attitudes towards the growth of state education (National Statistics, 2001). The expansion of the welfare state throughout the late 1940s led to an increased reliance upon and demand for the types of information which Social Survey could produce. In addition, the prevailing Keynesian economic ideology motivated the production of statistics geared towards assisting economic management (Thomas, 1996). During the post-war period, Social Survey developed methods for measuring the cost of living which would later evolve into the Family Expenditure Survey (the FES).

Despite threats of abolition in its early years, Social Survey was eventually established as a stable research unit staffed by Social Survey Officers, and working under a single director, Louis Moss (Nissel, 1987; Moss, 1990). At first, Social Survey consisted of a handful of researchers and roughly 50 interviewers. Early surveys were narrow in focus and answered clearly defined research questions; they were not conducted nationally, but focused on specific geographical areas answering specific questions. However, the staff underwent rapid professionalization: by the end of the decade, recruitment tests were in place for interviewers (National Statistics, 2001). Data analysis and presentation methods were still primitive by contemporary standards. Data were coded by clerical staff and processed mechanically, hand-punched onto cards, sorted, and tabulated. Reports were typically 30 pages long, with the questionnaire and field instructions attached; charts were hand-drawn, if included (National Statistics, 2001).
Cutbacks in the Social Survey budget at the start of the 1950s did not stop the trend of growth in the Division, in line with the prevailing mood of economic and social planning. In 1953, the Household Expenditure Survey was started; this was designed to provide up-to-date weights for the Retail Price Index. Social Survey played a large part in developing this survey, which was relaunched in 1957 as the Family Expenditure Survey. The International Passenger Survey (IPS) was started in 1961; this survey is still running today. There was also growth in ad hoc surveys throughout this period (National Statistics, 2001).

The 1960s was a period of further expansion for Social Survey, and it was briefly a government department in its own right. This period of growth was encouraged by the report of the 1963 Heyworth Committee, which advocated increased co-ordination between social survey activities, and also by the Labour administration which came to power in 1964, with a raft of new legislation. This led to a dramatic expansion in social research. By 1965, the Division had received a 40% increase in funding, and a graduated pay scale had been introduced for interviewers, further professionalizing social research. In 1967, the Government Social Survey Department was created, reporting to the Treasury, and with Louis Moss at its head; the Department doubled its staff between 1961 and 1967 (National Statistics, 2001).

In 1970, Social Survey merged with the General Register Office (GRO) to form the Office for Population, Censuses and Surveys (OPCS). (The GRO, now a part of ONS, is responsible for ensuring the registration of all births, marriages and deaths in England and Wales, and for maintaining a central archive dating back to 1837.) The purpose of this merger was to use census and survey results to complement each other, since censuses were believed not to be the best way to investigate, for example, characteristics of households or to probe sensitive subjects (Nissel, 1987). The first major multi-purpose survey undertaken by OPCS was the General Household Survey (begun 1970), which provides data on economic and social characteristics of the British population, particularly in the periods between censuses. The Labour Force Survey was begun (1973) to fulfil a commitment to other EEC member states to provide a survey which could be harmonized and synchronized across the community. Thomas (1996) points to the increase in longitudinal studies from the early 1970s onwards as part of this shift towards linking between surveys and administrative data. A growing pool of social science graduates continued the professionalization of social survey research (National Statistics, 2001).
The 1990s saw a growth in the number of surveys conducted, and the launch of several big, continuous surveys, in contrast to the pattern of repeat or serial ad hoc surveys of the 1980s (National Statistics, 2001). The Family Resources Survey was begun in 1992, with the first full survey conducted in 1993-1994. Another major change in this period was in the competitive tendering of surveys; competitive tenders grew from 3% of turnover in 1991/92 to 36% in 1995/96. In April 1996, The Office for National Statistics (ONS) was formed when the Central Statistical Office merged with the Office for Population, Censuses and Surveys. By the end of the 1990s, the Social Survey Division had 270 staff. ONS is the government department that provides statistical and registration services and is responsible for producing a wide range of economic and social statistics used by policy-makers across government, and also for the implementation of the Census. The Family Resources Survey is conducted and funded by the DWP, with the fieldwork contracted to the ONS and NatCen jointly.

The other main change in survey research conducted by Social Survey was the increasing use, from the late 1980s onwards, of computers in the survey process. The first small-scale test of portable computers for collecting data in face-to-face interviews was conducted in 1987, and a computer-assisted telephone interviewing (CATI) system was tested later that year (National Statistics, 2001). A major trial, which also included transmitting data back to headquarters, was carried out on the Labour Force Survey (LFS) in 1988; development on the system led to its adoption on the LFS in 1992 (National Statistics, 2001). The shift from paper-and-pencil questionnaires to computer-assisted personal interviewing (CAPI) was implemented gradually throughout the 1990s, and the use of portable computers by interviewers increased (National Statistics, 2001). The Labour Force Survey also developed further CATI systems for use by office-based interviewers. By the end of the 1990s, all household surveys used CAPI or CATI systems. The Family Resources Survey was the first survey to be CAPI-based from the outset. I shall go on to discuss the introduction of computers into survey research and the literature associated with this in greater detail in the following section; I first want to conclude this section with more information on the organization and administration of statistics within government, with particular reference to those bodies connected to the FRS.
Statistics in government

There are several professional groups within government engaged in social research and statistical work. These include: economists, statisticians, social researchers, and other specialists such as operational research specialists. These professional groups are associated with decentralized organizations which are spread across government departments; see Table 4.1.

Table 4.1 Professional groups in government

<table>
<thead>
<tr>
<th>Professional role</th>
<th>Organization</th>
<th>No. of staff across departments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economists</td>
<td>Government Economic Service (GES)</td>
<td>900 across 30 departments *</td>
</tr>
<tr>
<td>Statisticians</td>
<td>Government Statistical Service (GSS)</td>
<td>Over 1000 across 30+ departments **</td>
</tr>
<tr>
<td>Operational researchers</td>
<td>Government Operational Research Service (GORS)</td>
<td>175 across 10 departments †</td>
</tr>
<tr>
<td>Social researchers</td>
<td>Government Social Research (GSR)</td>
<td>Around 1000 across 18 departments ‡</td>
</tr>
</tbody>
</table>

* See http://www.ges.gov.uk/
** See http://www.statistics.gov.uk/recruitment/GSS/default.asp
† See http://ds.dial.pipex.com/teem/avenue/vo31/public/index.html
‡ See http://www.gsr.gov.uk/careers/departments.asp

Data from the report of the Rhind Commission on the Social Sciences (2003: 124) puts figures as of 2002 at over 850 GSR research officers and, as of February 2003, at 963 GSS statisticians and 850 GES economists. Compared with the figures I have given above, drawn (November 2004) from the various professional group websites, this suggests an increase in numbers which is part of the continuing and steady increase in numbers of such professionals which the Rhind report identifies (p. 124).

The section in DWP which runs the FRS is staffed primarily by statisticians, who fall under the aegis of the Government Statistical Service (GSS). The GSS is a decentralized body spread across most government departments which is made up of about 7000 civil servants working at ONS and in more than 30 UK government departments and the devolved administrations. The primary function of government
The ONS administers a personnel framework which sets the standards for recruitment, qualifications, and training. The Head of the GSS is the National Statistician and Director of the Office for National Statistics, Len Cook. Each government department or agency with a significant GSS presence, i.e. which produces or uses GSS statistics, has a Head of Profession for Statistics, with managerial authority for GSS staff. These Heads of Profession, the Chief Statisticians in the devolved administrations (not Northern Ireland), and all the staff that they manage, make up the GSS. Staff are recruited at a graduate level from a variety of numerate disciplines (e.g. mathematics, economics, sciences, business studies, psychology, geography), into Statistical Officer (STO) posts or, with more experience, Assistant Statistician (AS) posts. AS and STOs are expected to progress to Statistician within six years of appointment (this is a Grade 7 civil service post). During my own period working on the FRS, most of its team members were STO level.

The Department for Work and Pensions (DWP) is the department within which the FRS is conducted. The DWP (formerly the DSS) is one of the largest government departments and one of the biggest employers in Britain, employing over 100,000 staff across Britain to deliver benefits and services, with an expenditure of over £100 billion. Social research and data analysis at the DWP informs policy development, monitoring and evaluation, with research focused on six main areas: work and welfare; pensions and older workers; children, poverty and housing; health, work and disability; European Social Fund programmes; and public service reform.

---

23 Information drawn from: http://www.statistics.gov.uk/about/ns/statistician.asp.
26 Both the division within which the FRS team operates and the government department of which that division is a part underwent several changes of name during the period of this research (1999-2002). The Department for Work and Pensions (DWP) was formed by a merger of the former Department of Social Security (DSS) and parts of the former Department for Education and Employment in June 2001 (shortly after the general election in that year). The department which managed the FRS within the DSS was ASD2E; following the Division's reorganization this was renamed ASD IA 1; following a later reorganization in November 2002, this became the Information and Analysis Directorate, so IAD IA 1. In this section, I provide information on the organization of research at the time of writing. However, at various points I shall refer to both the DWP and DSS where appropriate; for example, when discussing the inception of the FRS, I refer to the DSS.
Research methodology includes large-scale surveys, in-depth interviews, experiments, ethnographic methods, and analysis of administrative data.

The Information and Analysis Directorate (IAD) is the central core of analysts at the DWP, whose primary purpose is to provide analytical information in order to advise ministers, DWP colleagues, and the public. The IAD also manages the DWP’s research budget and is responsible for research projects being carried out by universities, survey organizations and other research institutions on behalf of the Department. Research staff members across the various divisions of the IAD are in general professional economists and statisticians; they are supported by specialist computing teams. The divisions within IAD are listed in Table 4.2 (information derived from internal documentation), with details on the size and organization of each division.

Income Analysis is a division within the IAD which comprises of about 20 staff organized into 6 sections. The general responsibilities of the Income Analysis division are to provide statistics and analysis on incomes and the take-up of income-related benefits, and to provide analysis of the take-up of child support, family benefits, income mobility, and benefit distribution. The divisions within IAD are listed in Table 4.3 (information is derived from internal documentation).

Income Analysis 1 (IAD IA 1) is the section responsible for the Family Resources Survey. IAD IA 1 is responsible for: firstly, the management of the FRS (e.g. liaison with ONS and NatCen; co-ordination with other government surveys; liaising with users); and, secondly, the development of the FRS (e.g. devising questionnaire content; developing the database through imputation and tabulation)\(^\text{29}\). During the period of time that I was carrying out the project, the FRS team consisted of between 3-5 statistical officers, headed by a project manager.

Table 4.2 Information and Analysis Directorate divisions

<table>
<thead>
<tr>
<th>Division</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Centre **</td>
<td>Around 150 statistical, IT specialist and administrative staff organized into 10 teams.</td>
</tr>
<tr>
<td>Operational Research ***</td>
<td>Around 50 staff organized into 10 sections.</td>
</tr>
<tr>
<td>Income Analysis</td>
<td>Around 20 staff organized into 6 sections.</td>
</tr>
<tr>
<td>Benefit Forecasting and Modelling</td>
<td>Around 20 staff organized into 5 sections.</td>
</tr>
<tr>
<td>Social Research</td>
<td>Around 40 staff organized into 10 sections.</td>
</tr>
<tr>
<td>Working Age Strategy Directorate</td>
<td>Around 50 staff organized Into 15 sections.</td>
</tr>
<tr>
<td>Pensions Strategy and Planning Directorate</td>
<td>Around 12 staff organized into 4 sections.</td>
</tr>
</tbody>
</table>

* Information dates from the end of period of research (August 2003).
** Located in Newcastle-upon-Tyne; all other divisions based in central London.
*** Note the relationship with the professional groupings described above (p. 91).

Table 4.3 Income Analysis division sections

<table>
<thead>
<tr>
<th>Section</th>
<th>Responsibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA1</td>
<td>Management and development of the Family Resources Survey.</td>
</tr>
<tr>
<td>IA2</td>
<td>Statistics on take-up of income-related benefits (e.g. Income Support, Housing Benefit); developing and maintaining grossing system for the FRS.</td>
</tr>
<tr>
<td>IA3</td>
<td>Households Below Average Income statistics.</td>
</tr>
<tr>
<td>IA4</td>
<td>Statistics on individual incomes of women and men.</td>
</tr>
<tr>
<td>IA5</td>
<td>Analysis and briefing on distribution of benefits by income; distributional aspects of pension changes.</td>
</tr>
</tbody>
</table>

In summary: three principal bodies are involved in the management and production of the Family Resources Survey. The contract for implementing the survey is held by the Social Survey Division of the Office of National Statistics and the National Centre for Social Research (NatCen: a non-commercial organization for designing, conducting and interpreting major social surveys)\(^30\). The survey is managed by the Department for Work and Pensions (DWP) Information and Analysis Directorate Income Analysis 1 (IAD IA 1), who also carry out data analysis for the DWP and other users of FRS data. My own main contact throughout the period in

\(^30\) See: [http://www.natcen.ac.uk/natcen/index.htm](http://www.natcen.ac.uk/natcen/index.htm).
which I was conducting this project was with the FRS team at the DWP, who would be using the documentation to assist in analysis of survey data.

**Information age government and the dissemination of statistics**

The decision that the FRS questionnaire documentation developed in this project should be accessible on the FRS intranet (and, ultimately, on the Web) was one of the earliest design choices made on his project, and it was one made largely before I took on the project and began work on the site. In a large part, this decision arose from the current orthodoxy about the benefits of the Web for improving and modernizing government practice, and from political initiatives aimed at putting government services online. In this section, I begin by giving an overview of current government e-Strategy, and then go on to review the online dissemination of social survey research output, in order to give contextual information for the decision to produce online documentation for the FRS.

Statements of government e-Strategy combine the language of information society theories (which I discussed in detail in chapter 3, pp. 42-56) with the goal of the modernization (and transformation) of government; ICTs are to play a significant and formative role in this project of modernization. As the early Cabinet Office strategy paper *e-government: a strategic framework for public services in the Information Age* (April 2000) states:

> 'The Information Age revolution has already brought huge changes to both manufacturing and service industries all over the world. It has driven down costs, brought suppliers closer to customers, and made them more responsive to their needs. The Government has launched initiatives to make the United Kingdom a world leader in e-commerce and to make access available to all. This third initiative, e-government, will ensure that government itself will play a full part in this radical transformation of our society'[^31].

A key document expressing these goals is the White Paper 'Modernising Government' (March 1999). This set out the government's plans for a long-term programme of change across the public sector, at local and central government level. The focus is on:

'delivering outcomes that matter, through better policy-making and through working together across organisational boundaries to make things simpler, easier and more effective from the user's point of view'\(^{32}\).

The White Paper set out commitments in five key areas: forward-looking policy-making; responsive public services; high-quality public services; valuing public service; and information age government, to which I wish to pay particular attention.

Information age government is the subject of chapter 5 of the White Paper; the specific goal in this area is that the government 'will use new technology to meet the needs of citizens and business, and not trail behind technological developments'\(^{33}\). The Paper argues that a 'revolution' has taken place in how leading companies do business: this has been done through using networked computing to refocus business activities on the customer; using IT to work with suppliers; and making 'innovative use of information to become learning organisations'\(^{34}\). The White Paper then goes on to argue that government has not kept up with these developments, and has not yet 'developed ways of ensuring that we maximise the benefits of IT for government as a whole'\(^{35}\). The White Paper then lays out the framework for bringing about a 'fundamental change' in the way government uses IT, to 'modernise the business of government itself, achieving joined up working between different parts of government changes; various other Cabinet Office documents refer to the 'knowledge economy'; the Cabinet Office website describes the 'UK Online Strategy' as 'the Government's comprehensive programme to lead the knowledge economy revolution' [http://e-government.cabinetoffice.gov.uk/EStrategy/EStrategy/fv/en]. I shall return to a fuller discussion of information society narratives in relation to the FRS documentation in chapter 7.

\(^{32}\) Information drawn from [http://www.cabinet-office.gov.uk/guidance/one/directory.asp?intID=75](http://www.cabinet-office.gov.uk/guidance/one/directory.asp?intID=75). Note also here the emphasis on working across inter-organizational boundaries: there is a substantial body of literature on collaborative and inter-organizational networking as the main form of post-bureaucratic organization (see, for example Heckscher, 1994).


\(^{34}\) See [http://www.archive.official-documents.co.uk/document/cm43/4310/4310-05.htm](http://www.archive.official-documents.co.uk/document/cm43/4310/4310-05.htm).

and providing new, efficient and convenient ways for citizens and businesses to communicate with government and to receive services\textsuperscript{36}. The e-Government Unit at the Cabinet Office states the specific goal of the Information Age government programme as follows:

"The Prime Minister has set a clear target that all Government services to the citizen and to business should be available online by 2005\textsuperscript{37}. This will involve the transformation of access to government services and services, provided in ways which match people's needs, not government structures\textsuperscript{38}."

Strategic responsibility for e-Strategy was originally located in the Office of the e-Envoy, set up in October 2000\textsuperscript{39}. Most of the work done by the Office of the e-Envoy has shifted to the e-Government Unit (eGU), located within the Cabinet Office, which now has the main responsibility for strategic IT development within government\textsuperscript{40}. The eGU has specific responsibilities in areas of developing strategy (developing policy and planning for ICT within government and providing an element of programme management for implementation), and providing architecture (providing policy, design, standards, governance, advice and guidance for ICT in central government; commissioning government-wide infrastructure and services; and addressing issues of systems integration with other levels of government, e.g. at local or devolved government level). A broad distinction in terms of use can be drawn between the two current flagship government websites. Directgov provides public service information across government [http://www.direct.gov.uk/Homepage/fs/en]. The Government Gateway [http://www.gateway.gov.uk/] is the website for registering to use online government services, i.e. is for conducting electronic transactions with

\textsuperscript{36} See http://www.archive.official-documents.co.uk/document/cm43/4310/4310-05.htm.
\textsuperscript{37} The original target was 2008.
\textsuperscript{39} See http://e-government.cabinetoffice.gov.uk/MediaCentre/BiographyArticle/fs/en?CONTENT_ID=4000130&chk=H2vS6x.
\textsuperscript{40} A new head of the eGU, Ian Watmore, was appointed in September 2004; the title of e-Envoy had gone, and the Unit has acquired a new emphasis on efficiency: 'The e-Government Unit works with departments to deliver efficiency savings while improving the delivery of public services by joining up electronic government services around the needs of customers' (see: http://e-government.cabinetoffice.gov.uk/Responsibilities/fs/en).
government. (There were three DWP services available via the Government Gateway, at the time of writing.)

I now want to examine the online presence of the FRS, on both the DWP site, and also other sites (primarily academic resources) via which information about the survey and its associated documentation is disseminated. (Note: these resource sites are aimed at academics and other specialists, which differentiates them to some degree from the government sites aimed at the general public; however, the use of the Web to disseminate information is the key point in common.)

The front page of the DWP website [http://www.dwp.gov.uk] links to a section on 'Statistics and Research' [http://www.dwp.gov.uk/asd/], which is the main page for the Information and Analysis Division (IAD). The sub-section on 'Research' [http://www.dwp.gov.uk/asd/asd5/index.asp] provides links to further information on e.g. the research and analysis conducted by the Social Research Division (see Table 4.2 above). The sub-section on 'Statistics' provides links (many of which are to pages on the ONS site) to: information on statistics collected on: benefits (e.g. related to working age; families and children; pensioners); quarterly statistical summaries of benefit statistics; and the Family Resources Survey. The FRS site [http://www.dwp.gov.uk/asd/frs/] contains PDF and HTML versions of the FRS Annual Reports, dating back to 1998-1999; links to FRS-related reports; and links to series based on FRS data, e.g. the Households Below Average Income series; Individual Incomes; Pensioners' Incomes Series. There are external links to the dataset at the Data Archive at Essex University [http://www.data-archive.ac.uk/], and to documentation held at the Question Bank at the University of Surrey [http://qb.soc.surrev.ac.uk/docs/home.htm].

The Data Archive, founded in 1967, is a centre for data acquisition, preservation, dissemination and promotion and hosts the largest collection of digital data in the social sciences and humanities in the UK. It houses several thousand datasets of interest to researchers across sectors and different disciplines. The FRS database has been deposited at the Data Archive each year since it began, and a large

---

41 Note that the material which I developed in the course of this project is not available on the publicly accessible site.

42 A consultation exercise (running until July 2004) on the 'modernisation' of DWP statistical publications put particular emphasis on the electronic dissemination of statistics related to benefits and welfare to work programmes. It proposes that the Internet should become the principal means of disseminating DWP statistics, with paper copies of publication no longer produced. See http://www.dwp.gov.uk/asd/asdl/stats_consultation/stats_consultation.asp.
amount of metadata and documentation related to the FRS is made available online through the Data Archive [http://www.data-archive.ac.uk/findingData/frsTitles.asp]. For example, for the year 1998-1999 (the survey year which I documented in my own project), the Data Archive provides a list of all variables grouped according to table (with a simple search according to table and then name), and holds a substantial amount of documentation available for users to examine online or to download. In Table 4.4, I provide a summary of documentation held by the Data Archive for the FRS for the year 1998-1999 (including notes on the format in which documentation is available).

The Question Bank holds questionnaires for large-scale social surveys. As well as a short introduction to the FRS [http://qb.soc.surrev.ac.uk/surveys/frs/frsintro.htm], and also (in PDF) two versions of the questionnaire, with or without BLAISE codes. The Economic and Social Data Service (ESDS) is a national data archiving and dissemination service that came into operation in January 2003. It is a collaborative initiative between four centres located at the Universities of Essex and Manchester. It provides access and support for a range of economic and social data, both quantitative and qualitative, across themes and disciplines. It aims to promote and encourage data usage in teaching and research. It has four specialist services: ESDS Access and Preservation; ESDS International (access to international data sources); ESDS Longitudinal (information on UK longitudinal data collections such as the British Household Panel Survey); and ESDS Government (aiming to facilitate use of large-scale government surveys). The FRS is supported by ESDS Government. Information on the site related to the FRS is primarily introductory; however, there are also extensive links to information about the FRS on external sites such as the ONS site, the Data Archive, and the Question Bank.

This section has given historical and organizational context for the Family Resources Survey. I discussed briefly the use of computers in UK government survey research, and the online dissemination of results. I now provide a more detailed discussion of the emergence and use of computer-assisted interviewing in social survey research, and the body of literature surrounding it.

43 See http://www.esds.ac.uk/.
44 See http://www.esds.ac.uk/government/frs/.
3 Introduction of CAI into social survey research

Overview

In this section I shall set the FRS in context as the first CAPI-based survey conducted by UK government. To this end, I give an overview of the chronology of the introduction of CAI into the survey research process. CAI had a slower uptake in Europe than in the United States, where computer-assisted telephone interviewing has been used from the 1970s. I summarize the reasons given in the literature for the introduction of computers into the survey research process, and then outline some of the concerns which have emerged as a result of their introduction, with a particular
emphasis on the area with which this project is particularly concerned - the documentation of electronic questionnaires.

Traditionally in survey research, a key document for the survey researcher and user has been the *paper questionnaire*, which contained clearly set out information on question wordings, response categories, routing, checks made by interviewers, and so on. However, large-scale social survey research for public policy is increasingly being carried out by interviewers using laptop computers and CAPI (computer-assisted personal interviewing) programs (de Leeuw and Nicholls, 1996). Many national surveys are now CAPI-based, having initially been paper-and-pencil questionnaires, but the FRS has been a CAPI-based survey since its inception in 1993, and has been developed using the CAPI program BLAISE. As a result, the paper questionnaire has now largely been replaced by the electronic questionnaire; for the FRS, written in BLAISE code. This is viewed by the interviewer on the laptop by a series of screens. These questions are not all necessarily visible, but are written into the program, to ensure that only the right people are asked the right questions. The paper questionnaire itself has been replaced by a relatively raw printout of the questions and the routing instructions. The latest version of the FRS 'questionnaire', produced by one of the survey agencies, is over 1000 pages long, and is not regarded as particularly user-friendly. (I give an example of the appearance of the printout of the electronic questionnaire for the FRS in Table 5.7 in the next chapter.)

The historical development of the introduction and use of computers in survey research is well documented (see Baker, 1992; Manners, 1990). In the United States, which has, in general, placed more emphasis on telephone interviewing (de Leeuw and Nicholls, 1996), computerized versions of telephone interviews (CATI - computer-assisted telephone interviewing) were in use from the 1970s. De Leeuw and Nicholls note that computer-assisted 'mail' surveys, including using the Internet for data collection, are more prominent in the USA. In Europe, the greater emphasis on face-to-face interviewing (de Leeuw and Nicholls, 1996), meant more advancement took place in the area of computer-assisted personal interviewing (CAPI). Early adopters of CAPI were Statistics Sweden in the early 1980s, followed by considerable research into the effects of the new technology on the survey process in both the private and public sectors. The increasing availability of laptops made CAPI a feasible option from the 1980s onwards. The Netherlands Central Bureau of Statistics (CBS) was the first to establish a large, full-scale ongoing CAPI survey in 1987, with the
Netherlands Labour Force Survey. In the UK, the Labour Force Survey was the first survey to use CAPI from 1990, having originally been a paper-and-pencil questionnaire; the Family Resources Survey (FRS - started in 1992) was the first full-scale, ongoing survey in the UK to be conceived as a purely computer-based survey. Most authors offer a taxonomy of the various forms of CAI (e.g. Snijkers, 1992; de Leeuw and Nicholls, 1996). Table 4.5 summarizes the variety of acronyms and types that appear in the literature.

Table 4.5 Taxonomies of CAI

<table>
<thead>
<tr>
<th>Type of CAI</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer-assisted telephone</td>
<td>The interviewer is seated at a computer terminal and asks the questions which appear on screen; the respondent's answer is typed directly into the computer. The most usual CATI setup is using a network, with supervisors present for quality control and to assist with problems; however, technological change makes it possible for a decentralized CATI survey to be carried out, for instance from interviewers' own homes. In terms of problems with CATI, Martin et al. (1993) point to the higher number of households in the UK which do not have telephones (roughly 12%) as compared to the US, and suggest that this could lead to some sampling problems. CATI has been, in general, more popular in the USA than in the UK and Europe (de Leeuw and Nicholls, 1996).</td>
</tr>
<tr>
<td>Computer-assisted personal interviewing</td>
<td>Interviewers visit respondents in their own homes with a laptop computer, and conduct a face-to-face interview. After the interview, the data is sent back to a central location, either on disk or by modem. New interviewer instructions and sample addresses can also be sent this way (Baker, 1992; Martin and Manners, 1995).</td>
</tr>
<tr>
<td>Computer-assisted self interviewing</td>
<td>Respondents themselves read and answer the questions on screen; the program guides the respondent through the questionnaire. An interviewer need not be present, although scenarios are emerging in which a CASI element is part of a broader CAPI interview, for example, when answering sensitive questions (Couper and Rowe, 1996).</td>
</tr>
</tbody>
</table>

Computer-assisted data collection (CADAC) methods have become ever more popular in survey data collection, increasingly replacing paper-and-pencil methods in academic, government and commercial environments (de Leeuw and Nicholls, 1996). Initial enthusiasm for these new methods stemmed from their potential for revolutionizing the data collection process. Manners (1990) and Saris (1991) identified some of the expected benefits of introducing computer-assisted interviewing (CAI) into large-scale surveys. Firstly, it is argued that CAI leads to better quality of data through the automatic routing of questions to prevent interviewers following the wrong route; by means of the computer carrying out more complex calculations; and through managing errors (e.g. invalid answers) by
automatically detecting them and checking with the respondent during the interview. Secondly, it is argued that CAI leads to improved speed and lower cost over paper-and-pencil interviewing, after the initial investment, as a result of eradicating the data inputting stage necessary with paper-and-pencil interviews. Concerns about the impact of introducing CAI have also been expressed; these surround the potential for negative effects on data quality, and costs and timescales, and also the extent to which the new technologies are acceptable to interviewers and respondents. I discuss this literature in more detail now.

**Measuring the impact of introducing CAI**

**Effects on data quality**

Nicholls *et al.* (1997) provide a review of the main research conducted on the effects of the transition to new data collection technologies on survey data quality. In the case of unit non-response (i.e. failure to obtain the requisite information from a designated sample unit), there had been some concern that CATI or CAPI respondents would object to having their information stored on a computer. However, several studies comparing refusal rates in both CAPI and CATI with those of equivalent paper-and-pencil control groups have typically found no significant differences (Baker *et al.*, 1995; Manners, 1990). Indeed, Baker *et al.* found a greater respondent willingness to disclose sensitive information (see also below on respondent and interviewer acceptance).

*Item non-response* is concerned with the extent to which respondents give poor answers, or the failure on the part of interviewers to ask questions. Nicholls *et al.* report that 'one of the most consistent conclusions of the CAI literature is that CAI can eliminate virtually all respondent and interviewer omissions of application items, but provides little or no reduction in rates of explicit refusals'. Baker *et al.* concur, in that they found lower item non-response for CAPI compared to PAPI (paper-and-pencil interviewing). This is attributed to the automation of branching by CAI programs, which eradicates the possibility of interviewer branching errors, particularly in very complex questionnaires with multiple branching dependent on respondents' answers to previous questions. Sebestik *et al.* (1988, cited in Weeks,
1992) suggest that 90% of errors made by PAPI interviewers were failures to record a required response and thus impossible to make with CAPI. Nonetheless, Baker et al. point out that the elimination of interviewing errors in executing branching instructions assumes that the CAPI program has been adequately designed and tested.

There was also concern that, whilst automating the process would remove the error in following the questionnaire, the software or hardware itself might introduce new error, for example, through keying errors and poor typing skills (Baker, 1992). Baker's article concludes that these concerns remain hypothetical and have not been found in empirical study. Dielman and Couper's (1995) empirical study of key presses when compared to audio tapes of the same interviews reports a 0.095% error rate, somewhat alleviating concern about the potential of typing errors introducing significant error into a CAPI survey.

*Effects on costs and timescales*

The introduction of CAI into the survey process theoretically offers the potential for speeding up the process by removing editing and inputting stages towards the end of the process and hence, also, reducing the cost of the process. However, additional costs emerge at the start of the process as a result of investing in new technology (laptop computers, etc.) and training interviewers to manage the new system. Sebestik et al. (1988, cited in Weeks, 1992), in an early study of the costs of introducing CAPI, found that training costs for CAPI were on average 18% higher than a comparable pen-and-paper study, and field data costs were 17% higher.

The significant point here is that costs primarily surrounded the transition from a paper-based interview to a CAPI system. Once the infrastructure is in place, the benefits of a CAPI system emerge. As Baker et al. (1995) sum up the effect of CAPI on costs: 'CAPI may initially be somewhat more expensive than PAPI, but the cost difference is likely to narrow as organizations and interviewers gain experience in using CAPI, especially if the costs of portable computers continue to decline' (p. 413). Nevertheless, Poynter (2000) is pessimistic about the future of CAPI, predicting that, by 2005, the high initial investment costs will lead to it being sidelined in survey research. He does, however, note that one exception will most likely be an increase in the use of handheld computers, because of their portability and
the fact that they can also be connected to the Internet. At the moment, however, limitations in terms of memory mean that handheld computers cannot yet cope with large-scale social surveys.

**Respondent and interviewer acceptance**

Baker (1992) outlines the potential barriers which could affect respondent and interviewer acceptance of new technology: concerns about confidentiality and seeing the computer as intrusive; and the pace of the interview being controlled by the speed of the computer leading to a loss of rapport and eye contact between interviewer and respondent. His own studies reported that, whilst a large majority of respondents were enthusiastic, a steady minority of 5% preferred paper-and-pencil versions of the interview. Most respondents thought the interviewer appeared more professional with the computer.

Baker *et al.* (1995) reported a greater respondent willingness to disclose sensitive information, i.e. that computerized interviewing led to more accurate reporting of sensitive data. In general, however, studies have tended to support the claim that computerization *per se* has less effect on the reporting of sensitive behaviour than whether or not the interview is self-administered (Jobe and Pratt, 1997; Wright *et al.*, 1998; Tourangeau and Smith, 1996). Wright *et al.* also investigated the effect of the age of the respondent on acceptability of computerized self-administered interviews, finding that younger respondents tended to have more positive attitudes towards and familiarity with computers than older respondents. This finding was supported by Couper and Rowe (1996), who found in a study that willingness to self-complete an interview rather than insisting on interviewer-assisted completion was related to age, level of education, and computer experience. Buetow *et al.*'s (1996) study of the use of CAPI among Australian GPs supported the finding that older patients were more likely to regard the computer with suspicion, and to prefer not to operate the computer independently. For many, this was related to problems in reading the screen.

In terms of interviewer acceptance, Wojcik and Baker (1995, cited in Nicholls *et al.*, 1997) describe broad interviewer acceptance of CAPI achieved by a combination of lighter and more powerful hardware, enhanced software, improved
data management and communications, and carefully developed interviewer training. Once trained, most interviewers preferred to use CAPI. Martin et al. (1993) report that interviewers reacted well to using the computer and handled assignments without major problems. Some non-typists found the keyboard difficult, but it was argued that these problems would disappear with experience (Couper et al., 1997).

Couper and Burt's (1994) study of interviewer acceptance of CAPI contextualizes the impact of introducing a new form of working to an experienced workforce. They note that 'computer anxiety is generally associated with age, and with women' and that, in many cases, CAPI involves the imposition of a new method onto a workforce which tends to fit the profile of people expected to be anxious about using computers. Their study of attitudes before and after using CAPI found that, in general, users were positive about the new technology, and that the key factor in determining attitudes was experience rather than interviewer attributes.

The primary finding of these studies to investigate the introduction of CAI seems to be that the most significant effects of the introduction of computers into the survey research process are connected to the social aspects of that introduction - the extent to which the transition between paper- and computer- based systems is managed. A number of emerging issues are also examined in the literature.

**Emerging issues**

**Human factors**

There has been increased emphasis in the literature on the changes in the roles of interviewers and other professionals in the research process brought about by the introduction of these systems. De Leeuw and Collins (1997) note that the focus is now shifting from data collection techniques and sum up that it is 'safe to assume that with well-trained interviewers and the same well-constructed questionnaire, both CAPI and CATI will perform well, and differences in data quality will be extremely small'. They suggest, however, that the 'human factors' of CAI have been neglected, under which heading they include such issues as: whether reading from a screen and typing require different perceptual and motor skills than going through a paper questionnaire; or the
significance that needs to be attached to the reports of interviewers that it is harder to grasp the overall structure of the questionnaire.

Similarly, Couper and Nicholls (1998) conclude that the first effects of the change to computer-based survey research have been primarily operational and are already well-documented, i.e. the speed and efficiency with which surveys are conducted, and the completeness and consistency of data collection. They suggest that more important consequences surround issues such as the roles humans play in the collection process, and the nature of survey documentation. Bateson and Hunter, for example, (1991a and 1991b) discuss the changing roles of professionals involved in survey research as a result of the move to new technologies, emphasizing the greater role played by the interviewee in achieving quality data, and the extent to which researchers now take on some of the role of expert programmers in being able to specify questionnaires and consistency checks.

**Web surveying**

As noted above, Poynter (2000) is pessimistic about the future of CAPI, predicting that, by 2005, the high initial investment costs will lead to it being sidelined in survey research. Recently, the literature on the use of computers in survey research has begun to shift from studies of the impact of CAPI to studies of the impact of the Web on social survey research (see Reips and Bosnjak, 2001; Batinic, Reips and Bosnjak, 2002). A brief survey of these recent studies is summarized in Table 4.6. A fuller review of this literature is beyond the immediate scope of this thesis; however, I include this summary to show: firstly, the continuing shift towards the use of the Web in social survey research; and, secondly, to show how this literature is concerned with issues similar to those examined in relation to the introduction of CAI, i.e. studying the effects of the new technological intervention, studying respondent and interviewer acceptance, and examining unanticipated effects of introducing web-based technology into the survey process.
Table 4.6 Literature on the effects of the Web on the survey process

<table>
<thead>
<tr>
<th>Topic</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAWI - computer-assisted web interviewing</td>
<td>- Wings and Snijders (1998)</td>
</tr>
<tr>
<td>(new term for taxonomy)</td>
<td>- Roos and Wings (2000)</td>
</tr>
<tr>
<td>Effects on data quality</td>
<td>- Tuten, Bosnjak and Glascoff (1999)</td>
</tr>
<tr>
<td></td>
<td>- Vehovar, Batagelj, Manfreda and Zalatel (1999)</td>
</tr>
<tr>
<td></td>
<td>- Vehovar, Manfreda, and Batagelj (1999)</td>
</tr>
<tr>
<td>Effects on costs and timescales</td>
<td>- Couper, Traugott and Lamias (2001)</td>
</tr>
<tr>
<td></td>
<td>- Klassen and Jacobs (2001)</td>
</tr>
<tr>
<td></td>
<td>- Manfreda, Vehovar and Batagelj (2001)</td>
</tr>
<tr>
<td></td>
<td>- Perkins and Yuan (2001)</td>
</tr>
<tr>
<td>Respondent and interviewer acceptance</td>
<td>- Tourangeau et al. (2000)</td>
</tr>
</tbody>
</table>

Documentation issues

Kelly (1999) writes that 'it has become more and more difficult for developers, interviewers, supervisors, and managers to keep control of the content and structure of CAI instruments'. It is easy, and hence tempting, to add new functions to the instruments which rapidly make the program and its documentation more complicated. CAI questionnaires are *programmed* rather than *written*: in the case of the FRS, the program is produced in BLAISE, a variant of the PASCAL programming language (Manners, 1990), which is not broadly understood or accessible to non-specialists. The process thus becomes less transparent, even to analysts with statistical expertise and familiarity with the survey. Clark and Maynard (1998) reflect on the needs of secondary analysts who are one more step removed from the creation of the questionnaire. Whereas the computerization of the survey process has made raw data files more easily accessible via the Internet, analysts also need access to the questionnaires on which the survey was based in order to track question routing and the context in which questions are asked. The lack of a formal or easily understood questionnaire is a bar to this process (Bethlehem and Manners, 1998).

As Martin and Manners (1995) point out, in a PAPI (paper-and-pencil interview) survey, the questionnaire itself is a vital document for researchers and others to use as a record of what the survey covers. CAPI software, however, varies in the type of questionnaire documentation that can be produced. As Kelly (1999) points out: 'the documentation of CAI instruments [becomes] a separate task, one that was
not necessary when paper questionnaires were used. Kelly outlines the three most common current approaches of documenting questionnaires: producing separate questionnaire specifications independent of the CAI program; manual editing of the program; and semi-automated documentation of the electronic questionnaire. These approaches are problematic to the extent that they either remain as opaque as the questionnaire they are trying to replace, or else have the potential to introduce error into the process. In addition, many continuous surveys face problems when updating documentation from year to year as new questions are introduced or old ones dropped.

There are a number of projects documenting electronic CAPI questionnaires. Current solutions have tended to emphasize documenting either the questionnaire, or the survey metadata (i.e. data about data; this information allows analysts to use datasets more effectively). A questionnaire generator called qgen, which processes questionnaire information automatically into HTML or XML, has been developed at the MRC Biostatistics Unit, University of Cambridge; however, this currently can handle only very small-scale questionnaires (Walker, 2000). In terms of documenting metadata, Card (2000; and see http://www.socio.com) offers a means of providing information on and searching for variable metadata from multiple survey datasets; however, searchable forms of the survey questionnaires are not offered. Card's conceptualization of a user interface for variable metadata informed the development of the layout for the variable metadata on the FRS website (see chapter 6; p. 152).

The Data Documentation Initiative (DDI) is an international consortium of academic researchers and government statistical offices (including the ESRC Data Archive, which holds records of FRS data). The DDI aims to establish an international criterion and methodology for the content, presentation, transport, and preservation of metadata about datasets in order to produce codebooks which are uniform, highly structured, and are easily searchable on the Web. To this end, the DDI has developed what is known as a Document Type Definition (DTD) for the markup of social science codebooks. The project has aimed to provide a markup for codebooks which lends itself to simultaneous use of multiple datasets, with the purpose of providing a specification that might improve the entire process of data collection, data dissemination, and data analysis. The DDI's work has already been put into use by

---

45 http://www.mrc-biu.cam.ac.uk/qgen/.
46 http://www.icpsr.umich.edu/DDI/.
several major international projects, e.g. the Networked Social Science Tools and Resources (NESSTAR)\(^{47}\). NESSTAR aims to provide a 'seamless interface' between user and data and its documentation, by integrating data discovery, usage, and dissemination tools. The purposes of NESSTAR are to allow users to locate multiple data sources in every participating data archive in a single search operation, to browse highly detailed structured metadata about these resources, and to download subsets of cases and/or variables in a variety of formats.

The DTD employs XML (Extensible Markup Language), a 'next-generation' markup language which is a dialect of a more general markup language, SGML (Standardized General Markup Language). The XML DTD is seen as the 'glue' that brings together the entire data collection, distribution, and analysis process - thereby standardizing the process of accessing all data and documentation related to a large number of surveys. However, XML is a technology still in flux, for which standards are currently in the process of being set\(^{48}\). As a consequence of this, not all web browsers are capable of processing XML easily.

The TADEQ project (Bethlehem and Manners, 1998) was a collaborative R&D project funded by the European Union ESPRIT Programme. TADEQ stands for a 'Tool for the Analysis and Documentation of Electronic Questionnaires'\(^{49}\). The project involved four National Statistical Institutes across Europe, and a research institute to provide software development for the project. One partner in the project was Statistics Netherlands (developers of the CAPI program BLAISE, in which the Family Resources Survey is programmed), and the Office of National Statistics was also a participant (who, with NatCen, currently hold the contract for conducting the FRS). The TADEQ project was therefore of direct relevance to this project, since it involved the developer of the technology in which the FRS questionnaire is programmed, and also one of the main bodies that conducts the survey\(^{50}\).


\(^{48}\) See the information on XML at the World Wide Web Consortium's website: http://www.w3.org/.


\(^{50}\) BLAISE users form their own community (holding, for example, regular conferences; see: http://www.blaisusers.org), drawn from government statistical offices across the world (including ONS and Statistics Netherlands, who developed BLAISE); other research bodies such as NatCen; and participants from universities (in particular, the Survey Research Centre at the University of Michigan, which has a strong tradition in this field). The TADEQ project is an exemplar of the overlap between organizations and players in these small communities concerned with investigating the effects of CAI on the survey research process, and issues surrounding the documentation of electronic questionnaires.
The TADEQ project positioned itself as providing a solution to problems emerging from the replacement of paper questionnaires with CAPI systems. While these programs enable the development of large and complex electronic questionnaires, it has consequently become more difficult to control their content and structure. The aim of the project was to develop a tool that could document and analyze the structure of electronic questionnaires, and that could make a human-readable presentation (on paper or electronically in hypertext format) of the electronic questionnaire. The project focused on the development of its own XML DTD: the Questionnaire Definition Language (QDL); the purpose of this was to ensure an open tool that was appropriate for the multiple surveys with which the project team was variously involved. A prototype of the tool was released in 2001. In January 2001, I conducted testing of the TADEQ software on behalf of ONS, with particular reference to the Family Resources Survey. In chapter 6 (p. 173), I discuss how this evaluation of the TADEQ tool affected the development of online documentation for the FRS; in chapter 8 (pp. 239-241), I consider the relationship between the TADEQ project and my own project from an actor-network perspective. In the next chapter, I shall describe in detail the structure and content of the FRS, its users, and its associated documentation.

4 Summary

In this chapter I have given a detailed overview of the wider background in which my project to develop online documentation for the FRS can be located, describing a number of organizational and technological issues. I gave a historical overview of the organization of government statistics in the UK describing, in particular, those departments which are concerned with FRS. I also reviewed government e-Strategy particularly as it related to the increased use of the Internet by UK government in the dissemination of survey research results. I reviewed literature concerned with the increased computerization of the survey research process and the resultant need for new forms of documentation that this literature identified.

Having located the FRS within this organizational and technological background, I move on in the next chapter to a fuller description of the FRS and the documentation that existed at the start of this project.
5 The Family Resources Survey: a CAPI-based survey

1 Overview of the Family Resources Survey

The Family Resources Survey (FRS) is a large and complex multi-level survey which collects information on the incomes and circumstances of around 22-25,000 private households in Great Britain. This includes household characteristics; income and receipt of Social Security benefits; tenure and housing costs; assets and savings; carers and those needing care; and employment. The FRS contract is currently held by ONS Social Survey Division and the National Centre for Social Research (NatCen), and has been since the first full survey year (1993/94). The project is managed by the UK Department for Work and Pensions (DWP) Information and Analysis Directorate IA 1 (IAD IA 1).

The FRS was launched in October 1992 to meet the information requirements of Department of Social Security (DSS) analysts. Traditionally, the department had relied on other government surveys, notably the Family Expenditure Survey (FES) and General Household Survey. However, these surveys have relatively small sample sizes, and therefore did not provide sufficiently reliable information on many groups in society which were of particular interest to the DSS. National benchmark surveys in the UK, such as the General Household Survey, the Family Expenditure Survey, the British Crime Survey, and the British Social Attitude Survey, are all now CAPI, having been paper-and-pencil questionnaires when they started. In contrast, the FRS has been a CAPI-based survey since its inception. The FRS uses the CAPI program BLAISE, the program developed by Statistics Netherlands and used by ONS and NatCen for almost all public sector surveys.

Households interviewed in the survey are asked a wide range of questions about their circumstances. Although some of the information collected is available elsewhere, the FRS provides new or much more detailed information in a number of areas and was set up to bring some topics together on one survey for the first time. The sample size is intended to allow more confidence in the analyses of smaller sub-groups. Although the FRS was designed with DSS's needs in mind, it also contains information that is of interest to other government departments and outside
researchers. The FRS databases are deposited at the Data Archive\textsuperscript{51} at Essex University and are made available directly to other government departments.

Table 5.1 (overleaf) summarizes key information about the aims, content, and principal bodies which conduct the FRS, with specific reference to the 1998-1999 survey. (As I discuss in more detail in chapter 6, I chose to limit myself to documenting the 1998-1999 survey, since this was the most complete survey at the time the project began.) I shall then go on to give some general information about methodology; the types of data collected; the questionnaire and the unit of analysis; and the users of the FRS. In the following sections I shall give detailed descriptions of the questionnaire, the dataset, and the questionnaire documentation.

\textsuperscript{51} http://www.data-archive.ac.uk/
Table 5.1 Summary of the FRS (1998-1999)

| Principal investigators | Department for Work and Pensions (DWP)  
| Office of National Statistics (ONS), Social Survey Division  
| National Centre for Social Research (NatCen) |
| Aims | The FRS aims to:  
- support the monitoring of the social security programme  
- support the costing and modelling of changes to National Insurance contributions and social security benefits  
- provide better information for the forecasting of benefit expenditure |
| Main topics | - Household characteristics (composition, tenure type)  
- tenure and housing costs including Council Tax, mortgages, insurance, water and sewerage rates  
- consumer durables  
- vehicles  
- use of health services  
- welfare/school milk and meals  
- educational grants and loans  
- children in education  
- informal care (given and received)  
- occupation and employment  
- health restrictions on work  
- children's health  
- wage details  
- self-employed earnings  
- travel to work  
- personal and occupational pension schemes  
- income and benefit receipt  
- income from pensions and trusts, royalties and allowances, maintenance and other sources  
- income tax payments and refunds  
- National Insurance contributions  
- earnings from odd jobs  
- children's earnings  
- interest and dividends  
- investments  
- National Savings products  
- assets |
| Coverage | - Dates of fieldwork: April 1997 - March 1998  
- Country: Great Britain  
- Spatial units: Government Office Regions  
- Observation units: Families/households |
| Universe sampled | - Location of units of observation: National  
- Population: Private households in Great Britain south of the Caledonian Canal |
| Methodology | - Time dimensions: Repeated cross-sectional study annual - analysed on a financial year basis  
- Sampling procedures: Multi-stage stratified random sample  
- Number of units: 24000 (target) 23484 (obtained)  
- Method of data collection: Face-to-face interview |
| Date of first release | 10 January 2000 |
| Date of last release | 09 March 2004 (5th edition) |
Methodology

The information in Table 5.2 relates to the 1998-1999 FRS, and is derived from the Annual Report for that year.

Table 5.2 FRS methodology (1998-1999)

| Framework for sample selection | The FRS uses a stratified clustered probability sample drawn from the Royal Mail’s small users Postcode Address File (PAF). The PAF is a list of all addresses where fewer than 50 items of mail are received a day, and is updated twice a year. The survey selects 1,680 postcode sectors with a probability of selection that is proportional to size. Each sector is known as a Primary Sampling Unit (PSU). The PSUs are stratified by 24 regions and also by three other variables derived from the Census. Stratifying ensures that proportions of the sample falling into each group reflect those of the population. Within each PSU a sample of addresses is selected. In 1998-1999, 23 addresses were selected per PSU. Each year, one half of the PSUs are retained from the previous year’s sample, but with new addresses chosen; while for the other half of the sample, a fresh selection of PSUs is made (which in turn will be retained for the following year). This is to improve comparability between years. |
| Data collection methods | The consortium of Social Survey Division (SSD) of the Office for National Statistics and the National Centre for Social Research have been conducting fieldwork for the FRS since 1992. Interviews are carried out jointly on behalf of the DSS by interviewers from the ONS and NatCen. Each month, the PSUs are systematically divided between the two organizations and then assigned to the field staff. Before interviewers make contact with the selected addresses, a letter is sent to the address, explaining that it has been chosen for the survey, and that an interviewer will call. Participation in the FRS is voluntary. The interviewers are asked to call at the address. A lower limit of four calls is set and these calls have to be made at different times of the day and on different days of the week. In 1998-1999, FRS interviewers averaged 7.7 calls per address before returning it as a non-contact. The average interview lasts around one hour and 20 minutes, but the time will vary according to the size of household and its circumstances. Interviewers new to the FRS are briefed on the questionnaire and an annual re-briefing is given to all interviewers on changes to the questionnaire. Those who have been working on the survey for some time also complete a written field report each year, describing their experiences with particular parts of the questionnaire, and commenting on how changes are received in the field. |
| Response | The FRS aims to interview all adults in a household. A household is defined as fully cooperating when it meets this requirement. In addition, to count as fully cooperating, there must be fewer than 13 'don’t know' or 'refusal' answers to monetary amount questions in the benefit unit schedule (i.e. excluding the assets section of the questionnaire). Proxy interviews are accepted only under restricted circumstances. In 1998-1999, for those households classed as fully cooperating, proxy responses were obtained for 14 per cent of adults. |

Data collected

Modelling social security benefit entitlement is central to many of the DWP uses of FRS information, and the data collected reflects this, focusing on income, including receipt of social security benefits, housing costs, and circumstances of household members, such as whether someone gives or receives care or has childcare costs. This focus also underlies the routing of some questionnaires. For example, detailed questions on the value of liquid assets held are only asked of those respondents who are willing to provide an estimate of the value of their total savings and report a figure between £1,500 and £20,000. On average, a third of households surveyed fall into this category. This range is wide enough to capture those who may be entitled to benefit on the basis of their capital but reduces the burden on the majority of respondents. Further questions address other areas relevant to DWP policy such as barriers to moving off benefits and into work, pension provision, and maintenance payment and receipt.

Summary of questionnaire and unit of analysis

The FRS is broken down into three main units: households, benefit units, and individuals. The definition of a household used in the FRS is a single person or group of people living at the same address who either share one meal a day or share the living accommodation, i.e. a living room. So, for example, a group of students with a shared living room would be counted as a single household even if they did not eat together, but a group of bedsits at the same address would not.

A household consists of one or more benefit units, which in turn consists of a number of individuals (adults and children). 'Benefit unit' is a standard DWP term which relates to the tighter family definition of a single adult or couple living as married and any dependent children. A dependent child is aged under 16 or under 19 if still in full-time education. So, for example, a man and wife living with their young children and an elderly parent would be one household but two benefit units. In the process of the interview, the computer calculates the benefit units that people belong
to, and they are subsequently interviewed according to those groupings in the benefit unit questionnaire(s). Table 5.3 summarizes what may constitute a benefit unit.

The DWP groups people into benefit units in order to define dependency within families/households and thus eligibility for certain benefits and the amount they are entitled to receive. The benefit unit is the basic unit for analysis of FRS data.

Table 5.3 Summary of the FRS unit of analysis: the benefit unit

<table>
<thead>
<tr>
<th>Benefit unit types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A married or cohabiting couple*, with dependent children in the household.</td>
</tr>
<tr>
<td>2. A married or cohabiting couple*, with no dependent children in the household.</td>
</tr>
<tr>
<td>3. A man or woman with no wife/husband/partner* in the household, but with dependent children.</td>
</tr>
<tr>
<td>4. One person only: i.e. a man or woman with no wife/husband/partner* in the household, and with no dependent children.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Children under 16: included with parents or legal guardian.</td>
</tr>
<tr>
<td>b) Children aged 16-18 in full-time further education and living parent/legal guardian: included with parents or legal guardian. The parent(s) should still be receiving Child Benefit for them.</td>
</tr>
<tr>
<td>c) Children aged 16-18 not in full-time further education, and children 19 or over: treated as adults, in a benefit unit of their own (or with partner). No Child Benefit will be received for them.</td>
</tr>
<tr>
<td>d) Foster children (16+): if covered by a Local Authority maintenance allowance, treated as separate benefit unit; if not, treated as b) or c) above. Foster children under 18: treated as a) above.</td>
</tr>
</tbody>
</table>

* Same-sex cohabiting couples are considered by DWP rules to be in separate benefit units from each other. They are initially coded as 'cohabiting', but the program will compute that they are in separate benefit units.

Users of FRS data

Although the primary users of FRS data are within the DWP, the survey is also used outside the department. FRS data are incorporated into the Policy Simulation Model (PSM), used extensively by DWP economists for policy evaluation and costing of policy options. Responses are uprated to current prices, benefits and earnings levels, and calibrated to DWP departmental report forecasts of benefit caseload and expenditure. Using FRS data has made it possible to model some aspects of the benefit system which could not be done previously, e.g. income-related benefits, severe disability premiums, and allowances for childcare costs.

In addition to incorporation into formal modelling, FRS data play a role in the analysis of patterns of benefit receipt for policy monitoring and benefit forecasting.
Examples are the extent of multiple benefit receipt and the distribution of individual benefits. Data are also used in figures for take-up of income-related benefits. Figures are based on a combination of administrative and survey data. The aim of the analysis of FRS data is to establish how many interviewees who say they are not receiving benefits are in fact entitled to them. The access to metadata, especially on imputation\textsuperscript{52}, also informs analysts' judgements. FRS data are analyzed to produce analyses of incomes using Households Below Average Income (HBAI) methodology. The HBAI dataset also forms the basis of the Pensioners' Income Series, the DWP's analysis of trends in components and levels of pensioners' incomes.

The FRS has also been used as a sampling frame for follow-up studies to look at particular groups. For example, the Disability Survey, which re-interviewed over 7000 disabled respondents who appeared in the FRS between July 1996 and March 1997. The survey provided a detailed picture of the type and severity of disability, extra needs, and participation in leisure activities of people with disabilities. Merged with FRS information, a major use of the results by the DWP is to measure and analyze receipt of disability benefits and gather information to enable more accurate forecasting of expenditure.

Having provided an overview of the FRS, its users, and a general account of the data collected, I now move on to give a more detailed description of the questionnaire, the dataset, and the questionnaire documentation, all of which information is necessary for a full understanding of the project on which the current project is based.

\footnote{52 The process whereby missing values are computed.}
2 Description of the FRS questionnaire

Parallel blocks

The FRS interview consists of a household schedule; a benefit unit schedule repeated for as many benefit units as there are in the household; and an assets block for respondents with savings between specified levels. Each of these is known as a parallel block. (There are also two other parallel blocks: the recall block and admin block. These provide data on respondents and on the route taken through the questionnaire in a particular instance.)

Blocks of questions

The question instructions are broken down into sections: at the top level by parallel blocks, then within the Household and Benefit Unit schedules by blocks of questions on different subjects. The Assets Block is a separate question block in its own right.

The blocks of questions correspond to the way the interview program is divided up. Within the paper-based documentation, each of these blocks of questions is given a separate chapter, with a unique header at the top of the page. Each block has a name, a shorthand version of the content of the block. The contents page shows all the block names. (Some question blocks are divided into sub-blocks, and some variables are 'between blocks'.)

Question names

In a computer-assisted interview, there are no question numbers as such. Instead, each question is given a name. Parallel blocks, blocks of questions and question names are therefore the reference system used on the FRS. Table 5.4 provides a summary of the question blocks within the FRS questionnaire.
Table 5.4 Summary of question blocks in the FRS

<table>
<thead>
<tr>
<th>Block</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household schedule</strong></td>
<td></td>
</tr>
<tr>
<td>QNames</td>
<td>Household members</td>
</tr>
<tr>
<td>HHG</td>
<td>Household grid</td>
</tr>
<tr>
<td>QEthnic</td>
<td>Householder and head of household</td>
</tr>
<tr>
<td>QAccomDat</td>
<td>Benefit unit allocation</td>
</tr>
<tr>
<td>QRenting</td>
<td>Tenure and address information</td>
</tr>
<tr>
<td>QAccomH</td>
<td>Details of rented accommodation</td>
</tr>
<tr>
<td>QOwner</td>
<td>Owned accommodation and mortgage details</td>
</tr>
<tr>
<td>QInsur</td>
<td>Household insurance policies</td>
</tr>
<tr>
<td>QCounTax</td>
<td>Council Tax</td>
</tr>
<tr>
<td>QAccomCharge</td>
<td>Property charges</td>
</tr>
<tr>
<td>QWaterSew</td>
<td>Water and sewerage</td>
</tr>
<tr>
<td>QLodger</td>
<td>Intra-household contributions in conventional households</td>
</tr>
<tr>
<td>QShare</td>
<td>Rent and contributions within shared households</td>
</tr>
<tr>
<td>QProperty</td>
<td>Income from subletting</td>
</tr>
<tr>
<td>QPolities</td>
<td>Insurance policies</td>
</tr>
<tr>
<td>QModCons</td>
<td>Household durables</td>
</tr>
<tr>
<td>QTVehic</td>
<td>Vehicles: ownership and use</td>
</tr>
<tr>
<td>QWellare</td>
<td>Use of NHS services, free prescriptions, welfare milk and school milk and meals</td>
</tr>
<tr>
<td>QCare</td>
<td>Childcare</td>
</tr>
<tr>
<td>QCare</td>
<td>Help given and received</td>
</tr>
<tr>
<td><strong>End of the household schedule</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Benefit unit schedule</strong></td>
<td></td>
</tr>
<tr>
<td>QHealth</td>
<td>Health and ability to work</td>
</tr>
<tr>
<td>QEduc/QChEduc</td>
<td>Education, grants and loans</td>
</tr>
<tr>
<td>QNHHCh</td>
<td>Children outside the household</td>
</tr>
<tr>
<td>QCurst</td>
<td>Employment status</td>
</tr>
<tr>
<td>QJobDes</td>
<td>Job description</td>
</tr>
<tr>
<td>QEmpJob</td>
<td>Employee pay details</td>
</tr>
<tr>
<td>QSelfJob</td>
<td>Self-employed earnings</td>
</tr>
<tr>
<td>QTravel</td>
<td>Travel to work</td>
</tr>
<tr>
<td>QPens</td>
<td>Occupational and personal pensions</td>
</tr>
<tr>
<td>QBenefit</td>
<td>State and other benefits and pensions</td>
</tr>
<tr>
<td>QIncA</td>
<td>Income from pensions, trusts, royalties</td>
</tr>
<tr>
<td>QIncB</td>
<td>Maintenance, allowances, other income</td>
</tr>
<tr>
<td>QChinc</td>
<td>Children’s income</td>
</tr>
<tr>
<td>QAdint</td>
<td>Adults’ savings and investments</td>
</tr>
<tr>
<td>QChint</td>
<td>Children’s savings and investments</td>
</tr>
<tr>
<td></td>
<td>Total assets/change in income</td>
</tr>
<tr>
<td></td>
<td>End of the Benefit Unit Schedule</td>
</tr>
<tr>
<td>Question</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>QCurrAC</td>
<td>Amount in current account(s)</td>
</tr>
<tr>
<td>QASaveAC</td>
<td>Details of accounts</td>
</tr>
<tr>
<td>QAEquity</td>
<td>Investments/shares/bonds: details and value</td>
</tr>
<tr>
<td>QCertif</td>
<td>National Savings Certificates: details and value</td>
</tr>
<tr>
<td>QPGB</td>
<td>Pensioners' Guaranteed Income Bonds: details and value</td>
</tr>
<tr>
<td>QSaye</td>
<td>SAYE schemes: details and value</td>
</tr>
<tr>
<td>QPremium</td>
<td>Premium Bonds: details and value</td>
</tr>
<tr>
<td>QNSIB</td>
<td>National Savings Income Bonds: details and value</td>
</tr>
<tr>
<td>QABonds</td>
<td>Bonds: details and value</td>
</tr>
<tr>
<td>QCBonds</td>
<td>Bonds: details and value</td>
</tr>
<tr>
<td>QFirstOp</td>
<td>Bonds: details and value</td>
</tr>
<tr>
<td>QYPlan</td>
<td>Bonds: details and value</td>
</tr>
</tbody>
</table>

**Checks and warnings in the FRS**

In the FRS questionnaire, there are certain checks programmed which look for unlikely or inconsistent answers, as well as the usual limits to the ranges given for numerical responses.

- **Hard checks:** at these checks, the computer will stop the questionnaire, inform the interviewer of the problem and require an answer to be changed in order to proceed. They are normally for situations that are logically impossible, such as the year a property was bought by the household being before any household member was born.

- **Soft checks:** these occur when unusual but possible answers are entered. Here a warning screen querying the situation appears. Interviewers can suppress the warning and continue. A notepad facility exists in the questionnaire to mark that this has occurred.
3 Description of the structure of the FRS database

The FRS data exist in a series of hierarchical tables and also in the form of a flatfile.

Hierarchical tables

The FRS database consists of approximately 24 hierarchical, normalized\(^{53}\) tables, each table relating to a particular level (e.g. household, benefit unit) or type of information (e.g. pensions). Three main versions exist: the first covering edited and imputed fully co-operating households (the main database used by analysts); the second covering unedited data, as received from the survey contractors; and the third covering partially co-operating households (held for reference). Other datasets with different structures exist for non-responding households and other types of information relating to the data, e.g. a transactions dataset of edits applied.

General purpose tables covering household, benefit unit, adult and child information contain a record for each household, benefit unit, adult and child in the sample. Other, more specialized tables contain records on the basis of the relevant routing of the questionnaire/circumstances of the respondent. For example the OWNER table will contain only records for those households who are owner occupiers. Similarly, the BENEFITS\(^{54}\) table will only contain records for benefits that are received by the respondent. Individual records are uniquely identified by a combination of one or more key variables.

The highest level in the hierarchy is the household level. Records in this table (HOUSEHOL) are identified by the key variable SERNUM (serial number). The serial number is formed by concatenating sampling and location codes to form a unique identifier for each household within a survey year. It is further qualified by extra digits to denote the survey year, thereby extending uniqueness across all years. Each household may consist of a number of benefit units, records for which are identified by the additional key variable BENUNIT. Each benefit unit (and household) will

---

\(^{53}\) Normalization is the process whereby the database is restructured to eliminate duplication and to keep redundancy (number of skipped - not asked - values in any given table) to a minimum.

\(^{54}\) Table names are given in capital letters (BENEFITS), variable names in capitals and italics (BENUNIT).
consist of a number of adults and/or children, whose records are each identified by a person number as the third key. (Note that \textit{SERNUM} and \textit{PERSON} are enough uniquely to identify individuals in the sample.)

Below these main levels, other key variables exist depending on the table of interest. For example, an individual may receive a number of state and non-state benefits. Information about each benefit is held in a separate record in the \textit{BENEFITS} table. The key variables \textit{SERNUM}, \textit{BENUNIT}, \textit{PERSON}, plus the additional key, \textit{BENEFIT}, uniquely identify each record.

Similarly, a household that is buying its house with a mortgage may have one or more endowment policy/PEP/unit trust investments covering its loan. Information on these policies is held in the \textit{ENDOWMNT} table. Each record is referenced by \textit{SERNUM}, \textit{MORTSEQ} (mortgage sequence number) and \textit{ENDOWSEQ} (policy sequence number relating to that mortgage).

Key variables of the same name have the same coding frame across the database and can be used to relate information from different tables. For example, to look at the type of direct payments made by respondents alongside the total amount that is paid, the tables \textit{BENEFITS} and \textit{DSSPAY} would be related by \textit{SERNUM}, \textit{BENUNIT}, \textit{PERSON}, \textit{BENEFIT}. Similarly, to look at the age of recipients of benefits, the \textit{BENEFITS} and \textit{ADULT} tables would be related using \textit{SERNUM}, \textit{BENUNIT}, \textit{PERSON}. Or to look at contributions made by someone outside the household to mortgages, the tables \textit{MORTGAGE} and \textit{MORTCONT} would be related by \textit{SERNUM}, \textit{MORTSEQ}. Table 5.5 provides an overview of the information held in each of the tables of the FRS dataset, together with key variables used to identify individual records.
## Table 5.5 Tables in the FRS dataset

<table>
<thead>
<tr>
<th>Table name</th>
<th>Description</th>
<th>Key variables</th>
</tr>
</thead>
</table>
| ACCOUNTS   | Income from interest/dividend-bearing assets and savings together with (for a subset of records) the value of National Savings products for the accounts/investments held by adults and children. Each record relates to a type of investment (current account, savings account etc). Adults/children may have more than one type of investment, each record giving the total interest/dividends received (if they have more than one account of that type). | - SERNUM  
- BENUNIT  
- PERSON  
- ACCOUNT (account type, held by each person) |
| ADMIN      | Household level fieldwork administrative data. Each record relates to a household in the sample. | - SERNUM |
| ADULT      | Responses to various questions asked of adults. Each record relates to an adult in the sample (complete coverage of all adults). | - SERNUM  
- BENUNIT  
- PERSON |
| ASSETS     | Value and other information about assets and savings held by adults and children. Unlike the accounts table, each record relates to an individual investment (savings account, TESSA, PEP, shares with a single company etc), except for certain National Savings Products where the total value of specific types (e.g. income bonds, capital/deposit bonds) are recorded together. Adults/children may hold more than one investment of the same type. Data are collected for the subset of adults and children routed into the assets block of the FRS questionnaire. ASSETTYPE has the same coding frame as ACCOUNT on the ACCOUNTS table. SEQ is the sequence number through one or a group of assets, as determined by the block of the questionnaire, e.g. one block covers NSB ordinary and investment accounts, TESSAs, and other types of saving. | - SERNUM  
- BENUNIT  
- PERSON  
- ASSETTYPE (asset type)  
- SEQ (sequence number for that group of assets) |
| BENEFITS   | Details of amount of and other information related to state and non-state benefits received by adults in the sample. Also includes information on the total amount of direct payments as part of IS/JSA, as well as information on Social Fund loans held and future receipt of some benefits. Each record relates to an individual benefit. An adult may receive more than one benefit but not more than one of the same type. | - SERNUM  
- BENUNIT  
- PERSON  
- BENEFIT (benefit type, held by each person) |
| BENUNIT    | Benefit unit level data. Each record relates to a benefit unit in the sample (complete coverage of all benefit units). | - SERNUM  
- BENUNIT |
| CARE       | Information on those needing care. Each record relates to an individual in the household or various categories of individual outside the household (e.g. parent or child outside the household/client of a voluntary organization). NEEDPER is therefore equivalent to person for household members, but has additional codes for non-household members. For these cases, BENUNIT is set to 1. | - SERNUM  
- BENUNIT  
- NEEDPER (person receiving the care) |
| CHILD      | Responses to various questions relating to children. Each record relates to a child in the sample (complete coverage of all children). Information is collected by proxy from responsible adults. | - SERNUM  
- BENUNIT  
- PERSON |
| DSSPAY     | Information on the payments deducted from Income Support/Jobseeker's Allowance by the DSS to pay directly for different items (rent arrears, water charges, fines, maintenance payments etc). Each record relates to a type of deduction. Benefit type is either IS or JSA direct payments, using the same keys as in the BENEFITS table. All variables in this table are key variables. Up to 10 different direct payment types can be identified at DSSPAY. | - SERNUM  
- BENUNIT  
- PERSON  
- BENEFIT (benefit type)  
- DSSPAY (type of DSS direct payment, within each benefit) |
| ENDOWMNT   | Information on endowments/pension plans/PEPs/Unit Trusts etc being used to cover the mortgage. Each record relates to an individual policy. A household buying their property with a mortgage may have more than one policy covering more than one loan (to a maximum of 4 policies per loan). | - SERNUM  
- MORTSEQ (mortgage sequence number)  
- ENDOWSEQ (endowment policy sequence number, within each mortgage) |
| EXTCARE   | Information on children aged 16-24 living outside the household/benefit unit who are currently receiving full- or part-time education. Each record relates to an external child to a benefit unit. Each benefit unit may have more than one 16-24 year old living outside the household (to a maximum of 4). | - SERNUM  
- BENUNIT  
- EXTSEQ (sequence number of children living outside the benefit unit) |
| HOUSEHOL   | Information collected at the household level. Each record relates to a household (complete coverage of all households). | - SERNUM |
| INSURANC   | Information on insurance policies held by household members. Each record relates to an individual policy. A household may have more than one insurance policy (to a maximum of 6). | - SERNUM  
- INSSEQ (sequence number of insurance policy within household) |
**Flatfile**

Whereas data for the same benefit unit is split into numerous tables in the hierarchical dataset, the flatfile contains rows for each benefit unit in the year's dataset. The following example (see Figure 5.1) shows how the data for a particular household is held in both formats.

On the hierarchical dataset, the data for this household is linked by a serial number on the HOUSEHOL table. For example, in this case of a household made up
of a family of four, one benefit unit with information at this level is stored on the BENUNIT table. Information about the husband and wife is then stored on the ADULT table, with a separate record for each person. Similarly, information on their two children is held in the CHILD table (one record for each child).

The flatfile takes all this information about this benefit unit (all four family members) and stores it on one single line, starting with household variables, then benefit unit variables, head of household variables, spouse variables, child 1 variables, etc.

Figure 5.1 FRS datasets flow chart

Variables have different names on either dataset, despite holding identical information. Variables from the hierarchical datasets are mapped to the flatfile using suffixes to distinguish the person to whom the variable relates where necessary. Table 5.5 maps variables from the hierarchical dataset to the flatfile, continuing the example from above:
Table 5.5  Mapping variables from the hierarchical datasets to the flatfile

<table>
<thead>
<tr>
<th>Hierarchical variable (table)</th>
<th>Flatfile variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of husband</td>
<td>AGE (ADULT)</td>
</tr>
<tr>
<td>Age of wife</td>
<td>AGE (ADULT)</td>
</tr>
<tr>
<td>Age of first child</td>
<td>AGE (CHILD)</td>
</tr>
<tr>
<td>Age of second child</td>
<td>AGE (CHILD)</td>
</tr>
<tr>
<td>Amount of child benefit received</td>
<td>BENAMT (BENEFITS)</td>
</tr>
<tr>
<td>Amount of family credit received</td>
<td>BENAMT (BENEFITS)</td>
</tr>
<tr>
<td>Type of vehicle owned</td>
<td>VEHIC (VEHICLE)</td>
</tr>
<tr>
<td>Bto...</td>
<td></td>
</tr>
</tbody>
</table>

4 Description of the FRS documentation

FRS documentation in place at the start of the project fell into three categories, existing in both paper and electronic form. There is general information on the survey, metadata related to variables in the hierarchical dataset, and the survey questionnaire.

General information

Background information on the FRS is currently documented in a Guide produced each year by IAD IA 1. This provides preliminary information on the background of the survey (history of the survey, response rates); a description of the structure of the dataset; and programming information for SAS.

Metadata

The hierarchical dataset of the FRS is made up of 24 tables, with each of almost 1500 variables associated with one of the tables. Information (metadata) on each variable includes:

---

55 The documentation that I describe in this section, pp. 126-129, is the documentation that existed prior to my own project starting. The documentation which I developed in my own project is the subject of chapter 6.
• The name of the SAS format associated with the variable.
• The label associated with the variable.
• Numerical values associated with the SAS format.
• The value/label of the SAS format associated with the variable.
• The maximum and minimum values the variable can take.
• Whether the variable is derived.
• References to the questionnaire, the source question associated with the variable, the block of questions where the variable can be found, and the sequence in which questions are asked.
• Whether the variable is an integer, floating point or date type variable.
• Description of the type of variable, e.g. frequency, categorical, etc.

This information is currently held on an Excel file of more than 6000 lines, which is searchable, but not easily navigable. A major task for the current project was to develop a way of presenting this information in a more user-friendly way, whilst retaining the search functionality of the Excel file.

Questionnaire

In the case of the FRS, the electronic questionnaire, written in BLAISE code, is viewed by the interviewer on the laptop by a series of screens of questions which are not visible, but written into the program. This is to ensure that only the right people are asked the right questions, and can involve the nesting of questions (for example, between the household and the individual level). The paper questionnaire has been replaced by printout in a relatively raw form of the questions and the routing instructions. The latest version of the FRS 'questionnaire', known as the BLAISE automatic documentation (BAD), is over 1000 pages long in Word, and users need a certain amount of knowledge of BLAISE to be able to use this documentation.

As outlined above, the questionnaire has a highly complex structure, and not all questions are asked of all respondents. Much of this process is automated: e.g., respondents are routed only to those questions which are relevant to them; the same questions are asked again for different household members; questions with alternate
wordings (e.g. is/was, his/her) are automatically tailored to the situation; computations can be made in the course of the interview. None of these processes is visible to the interviewer, as they would be with a paper questionnaire. The BAD attempts to show some of these processes, but is largely unintelligible to a non-expert user.

Table 5.7 shows sections of the BAD documenting questions from the FRS. I have included annotations in the right-hand column to explain the content of the text. These examples have been chosen to show how the BAD documents three components of the questionnaire documentation:

1. **Question text:** The FRS involves proxy interviews, i.e. answers may be given to questions on behalf of people. The software automatically inserts pronouns and verb tenses which are appropriate to whether it is a personal or proxy interview.

2. **Answer types:** Variables on the FRS can be any of a number of types (e.g. numeric real, numeric integer, open, closed, date, and time variables), and can have a number of answer formats related to their type (i.e. integer ranges; numeric ranges; text string lengths; full details of the numbers and values of options available in closed variables; information on whether special answers were available and what these answers were). Users of a questionnaire require all this information.

3. **Conditions: checks and computations:** Checks on information provided are either hard (which prevent the interview from continuing until the information has been verified, or is consistent with prior answers), or soft, which provide an alert to the interviewer, but can be overridden to allow the interview to proceed. Users need this information to follow the route of a questionnaire. In addition, the electronic questionnaire allows computations to be conducted within the course of the interview easily. Users of the questionnaire need to know the conditions under which values are computed. Conditions for computations and routing also need to be shown so that the user can follow the path that an interview takes.

Users of the BAD are able to use the styles functionality in Word, for example, to hide certain parts of the questionnaire which are not immediately relevant (either types of information shown, or sections of the questionnaire). The questionnaire can
also be searched for text like any Word document. Nonetheless, as this example of a single question makes clear, the paper questionnaire supplied by the BAD is complex to read and navigate. Appendix 2 contains a longer sample of text from the BAD.

Table 5.7 **BLAISE Automatic Documentation**

<table>
<thead>
<tr>
<th>Sample text from the BAD</th>
<th>Explanation of text</th>
</tr>
</thead>
</table>
| **Ask if:** QAccomdat TENure IN [Part .. Squatting]  
  **And:** AccJob = Yes | **Conditions** under which the question is asked: here it is dependent on the answers to two questions (QAccomdat.Tenure and AccJob). |
| AccJobPer QRenting  
  Who is that?  
  CODE ALL THAT APPLY. | **Question text:** the question that is put to the respondent.  
  **Notes to the interviewer.**  
  **Answers:** number of options available. Note the proxy text (^DMName) which relates back to information given earlier in the interview. The interviewer would not see this on screen, since the information would be supplied automatically by the software. |
| SET [14] OF  
  (1) ^DMName[1]  
  (2) ^DMName[2]  
  (3) ^DMName[3]  
  (4) ^DMName[4]  
  (5) ^DMName[5]  
  (6) ^DMName[6]  
  (7) ^DMName[7]  
  (8) ^DMName[8]  
  (9) ^DMName[9]  
  (10) ^DMName[10]  
  (11) ^DMName[11]  
  (12) ^DMName[12]  
  (13) ^DMName[13]  
  (14) ^DMName[14] | **Checks** made on the answers supplied, and the questions and values which are used to make this check.  
  **Text shown to interviewer if the answer does not match conditions (note proxy text again).** |
| **Check if:** QAccomdat.Tenure IN [Part .. Squatting]  
  **And:** AccJob = Yes  
  **And:** In loop FOR Index := 1 TO 14  
  **And:** Index IN AccJobPer  
  PRec[1].Depend[Index] = Adult  
  Code ^Index is not valid for this question. |
In this chapter, I have given a detailed description of the Family Resources Survey, with particular reference to the survey year with which my project was concerned (1998-1999). I finished with a discussion of the FRS documentation.

In the following chapter, I describe my project to produce online documentation for the FRS. The structure of both the FRS dataset and questionnaire as I have described them informed the structure of the online documentation; for example, I used the tables of the hierarchical dataset to structure the layout of the table pages on the website (see chapter 6, p. 150), and I organized the online version of the questionnaire according to the question blocks (see chapter 6, p. 151). The information which I have given on users and existing forms of practice will become particularly relevant when I discuss the user survey I conducted (see chapter 6, pp. 192-198); the issue of users' needs, as and how I perceived them, is one which I shall take up at various points in the next chapter, and develop in my discussions of the site's 'audience', throughout chapter 7.
6 Online documentation for the Family Resources Survey:  
a case study of a hypertext application

1 Introduction

Overview of the chapter

In this chapter, I shall describe in detail the work I carried out to produce online documentation for the Family Resources Survey. Before beginning this discussion, I want to note that the description that follows in the bulk of this chapter is, by necessity, very technical in its references to the minutiae of both the FRS and the HTML and JavaScript which I wrote to document the survey questionnaire. I have saved fuller analysis of the social processes at work throughout the project until chapter 7; however, in order to prevent this chapter being entirely disconnected from the social context in which the documentation was produced, I want firstly to give some information about my working practices throughout the project, including some brief discussion of the people on the FRS team with whom I worked to produce the documentation, how the FRS team changed during the course of the project (with some preliminary reflections on how this affected the project work), where the work was conducted, and so on. I conclude section 1 of this chapter with an outline of the project. This describes briefly what, in retrospect, can be seen as its two main stages: the production of pilot documentation, and the production of full documentation. I shall outline the five main tasks that I undertook across the pilot and full documentation stages: these five tasks will form the basis of my organization of the rest of this chapter.

The bulk of this chapter is a very detailed description of the work I undertook during the project, which I present as follows. In section 2, I give a short description of the completed FRS documentation, as presented on the DWP intranet and on the CD-ROM which accompanies this thesis. Sections 3-7 go into full detail about the five main tasks which I carried out during the project. Throughout these sections, I shall indicate which other FRS team members were involved in the process, and how I carried out the work, in order to give some contextual information.
Throughout the time that I was working on the documentation (October 1999-August 2002), the FRS team consisted of between 3-5 statistical officers, headed by a project manager. In line with standard civil service practice, which aims to give junior staff broad experience within a department, the statistical officers varied across the time I was working with the FRS team and were in place, on average, for 18 months. Because of this staff turnover, the nature of the work I was conducting, and (eventually) the relatively limited time that I spent on site at the DWP (see below), I had only a small degree of contact with the wider FRS team, and the section in general. Contact here mainly took the form of organized meetings. In September 2000, I gave a presentation to the FRS Users Group, which consisted of users of the datasets within the FRS team and across ASD (about 10-12 people). I reported on the preliminary work done documenting the metadata for the website, and I received some feedback on layout and content. In July 2002, after completing work on the full questionnaire documentation, I conducted a brief user survey, interviewing FRS team members and others within the division who were users of the FRS.

For the duration of my project, however, my main point of contact was with the two individuals who held the post of project manager of the FRS, i.e. who were responsible for the management and development of the FRS. Mid-way through the project (2001), the original project manager moved on to another civil service position, and was replaced by a new project manager, who had previously been a statistical officer on the FRS team. There was a short hiatus in development at this point when the new project manager took over; after this, he and I took the opportunity to discuss how the project was going to progress, and we drew into our discussions a member of the department’s IT staff to guide future development on the project. This break point in the project is one which I will use to organize my forthcoming discussion of the documentation into a discussion of the production of the pilot documentation, followed by the full documentation. The break was also significant in marking a change in my own perception of the project, and in the kind of work I was carrying out, shifting the project away from my twin focus on design and formulating academic questions towards the users of the site. I return to this
subject in chapter 7, particularly when I discuss my changing perceptions of the 'audience' for whom I was carrying out the work.

In the first few weeks of the project, I went on the induction day that all new ASD staff attended, in order to learn more about the structure of ASD in general, and I also went on a SAS course run by the department in order to get a better sense of the types of statistical work that might be carried out by the FRS team (and so might have a bearing on how the documentation could be used). I went into the department offices in central London once a week and worked on site there. Being on site was convenient in terms of access to FRS documents and the questionnaire, in order to have discussions with the project manager at the earliest stages of the project, and, as the online documentation was developed, to be able to ensure its compatibility with the FRS intranet. However, working as, in effect, a part-time member of staff contributed to later difficulties which I had in making the transition from writing the website to writing the thesis (see chapter 8, pp. 243-244). The work that I carried out at this stage of the project I shall refer to as developing the pilot documentation: I shall discuss it in more detail below but, broadly, it consisted of determining the types of documentation that were already available, and coming up with initial design principles for the site; designing and implementing pilot documentation (in particular, of the metadata); and the design and implementation of search facilities. The work I carried out consisted primarily of, firstly, designing the preliminary structure that the site would take and, secondly, HTML and JavaScript coding of online versions of the documentation.

After the short hiatus in the project, and the arrival of the new project manager, I started work on what I shall refer to as the full documentation. Again, I discuss this in more detail below, but this consisted of designing and implementing an online version of the full questionnaire, conducting a small user survey, and then implementing the feedback obtained during this survey. Again, the work involved HTML and JavaScript coding; however, this work was substantially carried out away from the DWP. In part, this was a practical matter related to a shortage of space at the DWP's offices; in part, the work I was doing was no longer as reliant on access to information held at the FRS, and I was able to work on coding the site as satisfactorily from my own workplace. At this stage, I went into the DWP offices on a regular basis (every 4-6 weeks), to meet the FRS project manager and a member of the IT staff who had become more closely involved with the project at this stage. The project manager
and I invited this person to join our discussions because the future development of the project would be in part the responsibility of the IT department, and we wanted input on what they would find useful from the project. The IT staff member indicated that his department rarely had the resources to get user feedback on the design and content of their applications. The last task which I conducted was therefore a brief user survey of the documentation website: users from the FRS team and other users in the division were given access to the documentation, and were either interviewed by me, or provided me with feedback via email which I used to construct a 'wishlist' of changes to be made to the site (some of which I implemented, others of which were passed on to the member of the IT staff for later implementation).

Many factors influenced the form which the documentation eventually took, and I shall raise these in passing throughout the description that follows, although my full analysis is in chapters 7 and 8. Significant throughout development, particularly in the second stage, were issues related to user needs: how I perceived them; how they were constructed in sources I used such as internal documents or HTML manuals; how they were communicated to me by the project managers on behalf of staff who would be using the documentation; how they were communicated by users themselves. User needs are an important theme throughout this chapter, and I develop this theme in chapter 7, where I discuss my changing perceptions of the 'audience' of the site, and in chapter 8, where I characterize the 'story' of the second stage of the project as being concerned with creating a 'user-driven prototype'.

I was constrained in some of the development choices which I could make by pragmatic issues such as the time available to me and my own technical expertise (I had some basic HTML at the start of this project, as well as some experience in statistical analysis tools). There were also other issues arising from how this project related to projects and work being done both internally and externally, such as the ongoing support which IT staff at the DWP gave to analysts, projects such as TADEQ, and work done by organizations such as those setting standards for data documentation. I discuss issues arising from these in depth in chapter 8 when I consider how I went about 'translating' the project, but I shall refer to them in the description that follows in this chapter.
Overview of the project

The project to develop documentation for the Family Resources Survey can be seen as falling into two stages. Firstly, I developed pilot documentation, to try out preliminary versions of the design, layout, and features of the site.

The second stage of the project was an extension of this, to provide full documentation for the survey questionnaire, and to give IT staff user feedback on the design. Table 6.1 provides a brief overview of the main tasks which I carried out during the project.

Table 6.1 Project summary for documenting the Family Resources Summary

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Pilot documentation</th>
<th>Full documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>Determine existing documentation and outline design principles for site</td>
<td>Design and implement full documentation</td>
</tr>
<tr>
<td>Two</td>
<td>Design and test pilot documentation</td>
<td>Acquire user feedback and implement changes based upon it</td>
</tr>
<tr>
<td>Three</td>
<td>Design search facilities</td>
<td></td>
</tr>
</tbody>
</table>

In the next section, I give an overview of the completed website, describing its structure and organization, and a summary of its constituent parts. In sections 3–7, I give a detailed description of the development of the pilot and the full documentation. These sections are organized around the tasks outlined in the table above.

The documentation which I developed during this project is presented on a CD-ROM included with this thesis, and the reader is encouraged to examine this alongside the text. Throughout the text, file names from this CD-ROM are given in square brackets and blue text, directing the reader towards example pages from the site relevant to the points under discussion in the text. All material on the site and CD-ROM is Crown copyright.
2 Overview of the online documentation for the FRS

The documentation for the Family Resources Survey which I developed in the course of my project is presented online on the DWP intranet. The site gathers together material describing the FRS survey and the dataset, and presents searchable documentation of the metadata and documentation of the questionnaire related to the 1998-1999 FRS. There are help files to assist department members in navigating and using the features of the site.

The website's front page [docfront.htm] gathers together links to information under two headings:

- General documentation (i.e. information which is related to all survey years).
- Version-specific documentation (i.e. information related to a particular survey year).

Figure 6.1 shows the structure of the completed website. (Note: solid lines on this diagram indicate single pages and links; broken lines indicate that the box or line represents multiple pages or links; lines and dots indicate that pages and links do not yet exist, but are accounted for in the structure of the site.) All of the webpages shown here (and presented on the CD-ROM which accompanies this thesis) were created by me, working with DSS/DWP staff, during the course of this project, and did not exist prior to the start of the project. The content of the pages was drawn from existing material and documentation, which was not in HTML form, as I discuss in greater detail below (pp. 142-145).

---

56 File names in square brackets and blue text indicate files on the CD-ROM version of the documentation website supplied with this thesis. This enables the reader to examine individual pages of the site.
Figure 6.1 Structure of completed FRS website

Home page

General Documentation
- Background information
- Structure of database
- Programming examples
- Imputation information
- Style guide to questionnaire
- Questionnaire instructions

Version-specific Documentation
- 1998-1999

Questionnaire Documentation
- Questionnaire home page
- Household schedule
- Benefit unit schedule
- Assets block

Database Documentation
- Full description
- Search facilities

Metadata Documentation
- Benefit key
- Benefit map
- Period codes
- Usage codes

Forthcoming years

All pages in site

Table pages

Variable pages

137
General documentation

This links to pages gathering together information which holds for all years, including:

- Background: survey information; overview and response rates [frsback.htm].
- Structure: where the datasets are stored within the department; full description of both the hierarchical and flatfile datasets [frsstruct.htm].
- Programming examples: SAS examples for hierarchical and flatfile datasets [frsprog.htm].
- Imputation: details on how missing values are computed [frsimpute.htm].
- Style guide: how to navigate the paper-based BLAISE automatic documentation using the styles function in Word [frsstvles.htm].
- Questionnaire: description and instructions for using the BLAISE electronic questionnaire [frsquse.htm].

Version-specific documentation

This forms the bulk of the online documentation, and provides year-on-year documentation of the database, the metadata, and the questionnaire. Information pertains to a particular year.

Database documentation

The database documentation covers information on each table in the dataset; provides links to detailed metadata related to each variable in the dataset; and links to search facilities:

- Table descriptions: a single page gathers together links to pages for each table in the hierarchical dataset. Each table page consists of links to pages for each
individual variable in that table. Variable links are grouped together according to the main variable types: key (identifying), data, derived, and system variables [see, e.g. adult.htm]. (A fuller description of the table structure is given in chapter 5, Table 5.5, pp. 123-124 [structure of the database: frsstruct.htm; table descriptions: frstable.htm].)

- Variable metadata: each variable page (there are nearly 1500 in total) contains metadata related specifically to the variable. (See chapter 5, pp. 126-127 for a full description of the metadata [or, e.g. adch.htm].) There are links from each page to the general metadata pages where necessary (see below) and from the questionnaire back to variable metadata pages.

- There are three variable search facilities [frshdtopicsearch.htm]:
  1. Classification search: a search facility which enables the user to choose from 19 first-level search classifications and various second-level classifications. Links are given to lists of variables, and from there to individual variable pages.
  2. Topic search: a search facility enabling users to choose their own search terms by inputting keywords. Search results provide links to individual variable pages.
  3. Name search: alphabetical listing of all variables, by table, with links to individual pages.
Metadata documentation

These pages collect supporting metadata documentation which holds for multiple variables, and hence is linked to from most variable pages. This includes information on:

- Benefit key: code numbers mapped alongside the benefits they represent [frshdbenkey.htm].
- Benefit map: mapping of benefit code numbers alongside relevant questions [frshdbenmap.htm].
- Period codes: many of the questions on the FRS ask for amounts received/paid and to what period they relate (e.g. benefit receipt, Council Tax payments). In these cases, the amounts reported were converted to weekly equivalents using the given period code. This page maps the period codes alongside the period they represent [frshdrcode.htm].
- Usage: details of variable type (e.g. categorical, etc.) [frshduusage.htm].

Questionnaire documentation

From the questionnaire documentation front page [frquest.htm], the documentation of the questionnaire is split across three separate pages, each covering one of the parallel blocks: household schedule, benefit schedule, and assets block.

Within each parallel block page, the documentation is organized according to blocks of questions which correspond to the way the interview program is divided up. Each block has a name, a shorthand version of the content of the block. These block names are listed on the left-hand side of the page. Clicking on these names moves the user down the page to the relevant part of the questionnaire. Within each question block, the documentation is organized around question names.

In order to make these large pages manageable, users can hide or show the documentation text in two ways: using global instructions to hide away types of text; or by managing text within individual question blocks or around individual question names.
• *Using global instructions*

At the top of each parallel block page, there are eight global instructions given in blue text. A single click on the blue text performs a specific function on the whole documentation text:

1. Show or hide conditions - these control whether or not condition text is shown.
2. Show or hide question text - these control whether or not question text is shown.
3. Show or hide answer type - these control whether or not answer types are shown.
4. Show or hide all - reduces documentation to a list of question blocks, or refreshes the page to show all text.

• *Managing individual blocks*

The user can also reduce or expand the text of the questionnaire within individual blocks. Clicking on the large, bold black text hides away the contents of a whole question block. Clicking on this text again expands the block. Clicking on the bold blue text hides away the conditions, question text, and answer types around an individual question. Clicking on the text again expands this information. The global instructions override functions performed on individual blocks. If an individual block is hidden away, a global instruction to hide or show e.g. all conditions will still work when that individual block is expanded again.

*Summary*

In this section I have given an overview of the online documentation for the FRS which I developed in this project. In the following sections, I discuss the development in greater detail.
3 Task One: Determine existing documentation and outline design principles for site

In the earliest stages of the project, I determined what documentation already existed, and in what forms. There were several pieces of paper-based documentation, and also a number of files including the BLAISE automatic documentation, and a large variety of smaller pieces of documentation built up in an *ad hoc* fashion as the FRS team needed it, or as other users requested it. At this point, I was working on site one day a week, primarily to have access to this information, stored on the ASD computer network. In addition, in the course of our weekly meetings, the FRS project manager and I discussed some basic design principles for the site, which I drew up using her knowledge of which aspects of the site would be most useful for FRS users in the department. I also drew guidance on site design from HTML and hypertext manuals.

*Types of documentation*

*Paper-based documentation*

The key introductory document for staff to the FRS is the *Guide to the Family Resources Survey*. This is a 23-page booklet which provides preliminary information on the background and structure of the survey. The *Guide* covers a variety of topics:

- Background to the survey: basic information on history and response rates.
- Structure of the database: description of the hierarchical and flatfile datasets (the survey dataset is produced in two forms: a hierarchical dataset in which data are organized into 24 tables and, despite a more complex structure, is easier to use once learnt, and a flatfile dataset which is more accessible to the new user).
- Programming examples for both the hierarchical and flatfile datasets.
- Details of imputation (i.e. the process whereby missing values are computed).
- Style guide to the current questionnaire documentation: instructions for navigating around the current paper-based version of the questionnaire.


**Metadata documentation**

The hierarchical dataset is made up of 24 tables, with each of nearly 1500 variables related to one of these tables. Metadata (i.e. data about data) for each of these variables are held on a single Excel file which consists of nearly 6000 lines of data, which is searchable, but not easily navigable. In discussions with the project manager, I learned that the FRS team used this file frequently. We agreed that an immediately useful task would be to put this information online, retaining the search functionality of the Excel file. Table 6.2 lists the 24 tables and the number of variables associated with each. Each line in the original spreadsheet contains detailed information on individual variables.

**Table 6.2 Tables in the FRS hierarchical dataset (FRS survey 1998-1999)**

<table>
<thead>
<tr>
<th>Table name</th>
<th>No. of variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCOUNTS</td>
<td>9</td>
</tr>
<tr>
<td>ADMIN</td>
<td>32</td>
</tr>
<tr>
<td>ADULT</td>
<td>433</td>
</tr>
<tr>
<td>ASSETS</td>
<td>17</td>
</tr>
<tr>
<td>BENEFITS</td>
<td>30</td>
</tr>
<tr>
<td>BENUNIT</td>
<td>84</td>
</tr>
<tr>
<td>CARE</td>
<td>45</td>
</tr>
<tr>
<td>CHILD</td>
<td>166</td>
</tr>
<tr>
<td>DSSPAY</td>
<td>6</td>
</tr>
<tr>
<td>ENDOWMNT</td>
<td>8</td>
</tr>
<tr>
<td>EXTCHILD</td>
<td>11</td>
</tr>
<tr>
<td>HOUSEHOL</td>
<td>221</td>
</tr>
<tr>
<td>INSURANC</td>
<td>29</td>
</tr>
<tr>
<td>JOB</td>
<td>158</td>
</tr>
<tr>
<td>MAINT</td>
<td>25</td>
</tr>
<tr>
<td>MORTCONT</td>
<td>8</td>
</tr>
<tr>
<td>MORTGAGE</td>
<td>53</td>
</tr>
<tr>
<td>ODDJOB</td>
<td>9</td>
</tr>
<tr>
<td>OWNER</td>
<td>23</td>
</tr>
<tr>
<td>PENAMT</td>
<td>7</td>
</tr>
<tr>
<td>PENSION</td>
<td>26</td>
</tr>
<tr>
<td>RENTCONT</td>
<td>7</td>
</tr>
<tr>
<td>RENTER</td>
<td>53</td>
</tr>
<tr>
<td>VEHICLE</td>
<td>5</td>
</tr>
</tbody>
</table>

**Total**: 1440
The specific metadata for each variable contained in the Excel spreadsheet are as follows:

- The name of the SAS format associated with the variable.
- The label associated with the variable.
- Numerical values associated with the SAS format.
- The value/label of the SAS format associated with the variable.
- The maximum and minimum values the variable can take.
- Whether the variable is derived.
- References to the questionnaire, the question linked to the variable, the block where the variable can be found, the sequence in which questions are asked.
- Whether the variable is an integer, floating point or date type variable.
- Description of the type of variable, e.g. frequency, categorical, etc.

**Questionnaire documentation**

I have discussed the existing BAD in chapter 5 (see pp. 127-129); to summarize: the BLAISE automatic documentation (BAD) is a Word document of more than 1000 pages. At its highest level, it is organized according to the three sections of the FRS. First, the household schedule: addressed to one person in the household (usually the head, although other members of the household are encouraged to be present) and asking about household-level information, such as the relationship of individuals to each other, tenure and housing costs. Second, the benefit unit schedule: addressed to each adult in turn and asking questions about employment, benefits, pensions, investments and other income. Finally, the assets block asks about the value of investments for relevant respondents. Within the first two, lower-level blocks are broken into new sub-questionnaires.

I set aside documenting the questionnaire itself at this stage to focus on the metadata. However, in meetings I had with staff at ONS who produced the FRS questionnaire, it became clear that FRS users at the DSS were not aware that the BAD could be searched using facilities in Word. The project manager and I decided that providing instructions on how to do this would be an interim and helpful measure.
Design principles

Drawing on the project manager's knowledge of FRS users, and using HTML and hypertext manuals, I decided on three design principles for the documentation site:

- **Accessible via the Web.** As I had discovered, existing FRS documentation was in many disparate forms (documents, spreadsheets, graphics etc.); one of my main goals was to present it in a single place, on the FRS website on the DSS intranet (i.e. a private network shielded from the wider Internet; this also meant that the documentation could be made available externally via the Web). Since we wanted to use the intranet to disseminate this information, I chose HTML as the primary tool for development.

- **Anticipating users.** Shirk (1988) and Girill and Luk (1992) suggest that the most significant feature in the design of hypertext-based documentation is that it should not confound user expectations. I adopted this principle for the FRS documentation in three ways. Firstly, when designing the pages, I used the layout of other ASD 3E pages available on the DWP intranet. Secondly, I attempted to organize the information on the site in such a way that users would not be forced to 'hunt out' information, i.e. it should emulate the organization of the existing documentation. Finally, in transferring information from paper-based and other sources, I did not want to lose functionality that was already in place, e.g. the ability to search the metadata Excel file. In summary, then, my perception of user expectations was formed from conventions of hypertext derived from manuals and other general literature; the format of existing documentation; and from discussing existing forms of user practice with the project manager.

- **Updating documentation.** The FRS is an annual survey, and changes to the survey questionnaire are made each year. In addition, junior team members would change on a frequent basis. Consequently, my design had to allow for easy updates, and this process had to be readily explicable to new staff members.
4 Task Two: Design and test pilot documentation

In discussion with the FRS project manager, I had outlined two primary purposes for the initial documentation. Firstly, I wanted to develop a design for the layout and organization of the site and for the individual pages. Secondly, I wanted to develop a hypertext format for the Excel metadata documentation.

I chose to limit myself to documenting the 1998-1999 survey (known within the department as FRS 35), since this was the most complete survey at the time that I began work on the documentation website. With design and layout one of my primary interests at this stage, the simpler forms of the documentation were the main focus, i.e. the paper-based documentation and instructions for using the BAD. At this stage, I was working both on and off site, and was making most of the design choices independently, with occasional references to the FRS project manager to check, for example, details of terminology.

At this stage, from an academic perspective, I was exploring how the various taxonomies of hypertext worked in practice, and I drew extensively on the work of Hunter (1999); see my more detailed discussion in chapter 7, pp. 209-210.

**Paper-based documentation**

I presented the various sections of the *Guide* as separate web pages. Figure 6.2 shows the front page of the completed pilot documentation website (the link 'Structure' is in red to show a selected hyperlink). I adopted the layout of the page from existing ASD 3E web pages. The links on the left of the page are to other sections of that website. I added links from this front page as further parts of the site were developed (e.g. to the questionnaire documentation). Note the link ('Style guide', under 'Introductory documentation') to the page providing advice on how to use the BAD.

The five sections of the *Guide* (survey background; database structure information; programming examples; imputation information; guide to the paper-based documentation) are linked to under 'Introductory documentation' and I presented this information on separate pages. Figure 6.3 shows the page describing the structure of the FRS database. Within this page, I provided links to the separate
sub-sections at the top of the page (e.g. the link 'Hierarchical Tables' is in red, and clicking on this moves the reader to the section lower down the page).

Figure 6.2 Pilot documentation website front page
The FRS database consists of 24 hierarchical, normalized tables, each table relating to a particular level (e.g. household, benefit unit) or type of information (e.g. pensions). Three main versions exist: the first covering edited and imputed fully cooperating households (the main database used by analysts), the second covering unedited data, as received from the survey contractors; and the third covering partially cooperating households (held for reference). Other data sets with different structures exist for non-responding households and other types of information relating to the data, e.g. the flatfile data set out of date.

The FRS datasets are stored in `data1/frs/frs9*` on UNIX box 14, where `*` is the survey year (e.g. 67 for 1996-97), and `*` is the letter for the latest version of that release. These are copied to `data4/frs/frs9**` on UNIX box 18.

When data is updated a new release is created for users. The old release is not deleted, so can still be accessed, but the new updated release should be the data which is used by all users. The releases which should be used for each FRS year are:

- 1998-99 frs989a
- 1997-8 frs978d
- 1996-7 frs967d
- 1995-6 frs956f
- 1994-5 frs945p (NB. better not to use frs945q)

The FRS data exists in both a series of hierarchical tables and also in the form of a flatfile.

I repeated this layout for all of the Guide, with the exception of the pages demonstrating programming examples. I presented these on the website on a single page rather than the five pages which this section takes up in the Guide. These programs are given in the Guide to enable new users to make the transition from using the flatfile dataset to the more complex hierarchical dataset (FRS team members find the flatfile easier to use, but less effective in statistical analysis). The SAS programs in the Guide enable comparison of the two methods of programming in order to learn.
by example. In the paper-based documentation, these examples stretch across five pages, necessitating movement backwards and forwards through the text. When these examples are constructed in HTML (see Figure 6.4 [prog_eg_1.htm]), it becomes possible to present the two examples in a web browser so that a user can examine the different examples of programming simultaneously, using frames and scrolling within the browser to align parts of the two programs.

Figure 6.4 Programming examples

<table>
<thead>
<tr>
<th>EXAMPLE FLATFILE PROGRAM 1</th>
<th>EXAMPLE HIERARCHICAL PROGRAM 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>/* This program uses the FRS flatfile - ASD14 */</td>
<td>/* This program uses the hierarchical FRS dataset - ASD14 */</td>
</tr>
<tr>
<td>/* To find the employment status of the head of a benefit unit where the head is in receipt of DLA mobility - compare between 3 survey years */</td>
<td>/* To find the employment status of the head of a benefit unit where the head is in receipt of DLA mobility - compare between 3 survey years */</td>
</tr>
<tr>
<td>/* 1996/97 FRS */</td>
<td>/* 1996/97 FRS */</td>
</tr>
<tr>
<td>rsubmit; libname frs67 '/datal/frs/frs967d'; endrsubmit;</td>
<td>rsubmit; libname frs67 '/datal/frs/frs967d'; endrsubmit;</td>
</tr>
<tr>
<td>rsubmit; data test; set frs67.frs967 (keep=sernum benunit qdlacahd qdlamohd empbhd gross); flag=0; if qdlamohd=1 then flag=1;</td>
<td>rsubmit;</td>
</tr>
</tbody>
</table>

The column on the left shows an example of flatfile programming. The column on the right shows the same program for the hierarchical dataset. Scroll between the two to compare.

To return to Programming Information, use the Back button.
**Metadata documentation**

**Documenting the tables**

As I discussed above, the hierarchical dataset is made up of 24 tables, with each of nearly 1500 variables related to one of these tables (see Table 6.2 above for a list of these tables), and metadata on each variable is held in an Excel spreadsheet.

Figure 6.5 shows the ACCOUNTS table web page for the pilot documentation, with the link to the variable $ADCH$ selected, and shown in red. I also provided links to pages dedicated to the other variables associated with this table ($ACCINT$, $ACCOUNT$, etc.) and repeated this pattern for all the tables in the hierarchical dataset (with links to the other 23 table pages on the left-hand side of these pages). Note how closely, therefore, the structure of these pages mirrors the structure of the hierarchical dataset.

I used stylesheets to ensure that each page remained uniform. Stylesheets are a means whereby consistent information, such as the appearance of particular forms of text, e.g. a heading, can be located in a single document, and each web page can be directed to use the information from this document, rather than every page being coded up individually. I included links from each variable page to definitions of the categories used, and also information on benefits and how the variables are used.
ACCOUNTS table web page for the pilot documentation

ACCOUNTS

Income from interest/dividend bearing assets and savings together with (for a subset of records) the value of National Savings products for the accounts/investments held by adults and children. Each record relates to a type of investment (current account, savings account etc). Adults/children may have more than one type of investment, each record giving the total interest/dividends received (if they have more than one account of that type). For National Savings products, if the adult/child is not routed into the assets questions, a banded figure for the value of the investment is collected (those entering the assets block will have an accounts record but with this variable skipped).

Variables

ACCINT ACCOUNT ACCTAX
ADCH BENUNIT NSAMT
PERSON SERNUM _MONTH_
When coming up with a design for documenting the variable metadata, I drew on an account of a similar project described in Card (2000). In Card’s example, a single web page was dedicated to each variable, with links provided from the table with which the variable was associated. Figure 6.6 outlines more fully the structure I devised for this part of the pilot website: it shows the path for reaching one variable (ADCH) off the ACCOUNTS table. Each column in the original spreadsheet contains specific types of metadata (see above). (Note: solid lines on this diagram indicate single pages and links; lines and dots indicate that the box or line represents multiple pages or links.)

In the Excel spreadsheet, individual variables run across a handful of rows; e.g., the spreadsheet entry for the variable ADCH under the ACCOUNTS table is made up of three lines of text (in a larger document of nearly 6000 lines). The user must locate the appropriate variable by searching for text, and then must refer to a paper document which lists which type of information is contained in each lettered column.
For example, to discover what type of variable \textit{ADCH} is, the user of the spreadsheet would have to search for the variable, locate the appropriate column (using the FRS \textit{Guide} to find out which column contains the appropriate information), and then check the definition of '(I)'. Figure 6.7 shows the on-screen appearance of the Excel spreadsheet with the \textit{ADCH} cell selected (note how the scrollbars show the amount of scrolling necessary to look at all the data related to this variable, as well as the metadata for all variables).

This can be compared with Table 6.3, which summarizes the metadata for the variable \textit{ADCH} (split across two lines for space constraints). In contrast, Figure 6.8 (p. 155) presents all this information on a single page in HTML. As the selected hypertext link (in red) shows, users can click to access definitions of each type of metadata. These pages in turn contain information on: the benefits which the codes on the variable pages represent; details on period code (related to the process of converting monetary amounts into weekly values); what is represented by the usage codes. Coding individual pages in HTML for each variable (almost 1500 in total) was a substantial piece of work, and finessing the layout for these individual pages was the subject of ongoing discussions between me and both FRS project managers for the duration of the project. However, the basic structural design of one variable to one page of HTML did not change.

\textit{Search facilities}

One function of the Excel spreadsheet was that the metadata could be easily searched by key words. In line with my design principles, and also since the project manager had indicated that this was a popular function with users, I developed some preliminary means of searching the hypertext version of the metadata.

Card (2000) had indicated two types of search facility users of social science data are likely to want: by variable name and according to topic. Consequently, I included a very simple search facility by variable name for the pilot documentation: a single page with links to pages listing all variables beginning with the same letter; however, it relied on users knowing the name of the variable they were trying to find. There were nearly 1500 variables in the 1998-1999 FRS, and users of the documentation come from outside the FRS team. This meant, according to the project
manager, that they did not have the detailed familiarity with variable names that came from using them on a day-to-day basis. I therefore also decided to devise a topic search facility (see below, pp. 162-169).

**Figure 6.7 Spreadsheet information for variable ADCH**

<table>
<thead>
<tr>
<th>COLUMNS</th>
<th>LABELS</th>
<th>VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>TABLE</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>ACCOUNT</td>
<td>Interest received</td>
</tr>
<tr>
<td>A3</td>
<td>ACCOUNT</td>
<td>Account Type</td>
</tr>
<tr>
<td>A4</td>
<td>ACCOUNT</td>
<td>NSB Ordinary account</td>
</tr>
<tr>
<td>A5</td>
<td>ACCOUNT</td>
<td>NSB Investment account</td>
</tr>
<tr>
<td>A6</td>
<td>ACCOUNT</td>
<td>TESSA</td>
</tr>
<tr>
<td>A7</td>
<td>ACCOUNT</td>
<td>Savings, investments etc</td>
</tr>
<tr>
<td>A8</td>
<td>ACCOUNT</td>
<td>Government Gilts</td>
</tr>
<tr>
<td>A9</td>
<td>ACCOUNT</td>
<td>Stock</td>
</tr>
<tr>
<td>A10</td>
<td>ACCOUNT</td>
<td>Unit Investment Trusts</td>
</tr>
<tr>
<td>A11</td>
<td>ACCOUNT</td>
<td>Shares, Bonds etc</td>
</tr>
<tr>
<td>A12</td>
<td>ACCOUNT</td>
<td>PEP</td>
</tr>
<tr>
<td>A13</td>
<td>ACCOUNT</td>
<td>National Savings capital bonds</td>
</tr>
<tr>
<td>A14</td>
<td>ACCOUNT</td>
<td>National Savings Income bonds</td>
</tr>
<tr>
<td>A15</td>
<td>ACCOUNT</td>
<td>National Savings deposit bonds</td>
</tr>
<tr>
<td>A16</td>
<td>ACCOUNT</td>
<td>First Option bonds</td>
</tr>
<tr>
<td>A17</td>
<td>ACCOUNT</td>
<td>Yearly Plan</td>
</tr>
<tr>
<td>A18</td>
<td>ACCOUNT</td>
<td>Children's Bonus bonds</td>
</tr>
<tr>
<td>A19</td>
<td>ACCOUNT</td>
<td>Before tax</td>
</tr>
<tr>
<td>A20</td>
<td>ACCOUNT</td>
<td>After tax</td>
</tr>
<tr>
<td>A21</td>
<td>ACCOUNT</td>
<td>Whether Adult or Child Account</td>
</tr>
<tr>
<td>A22</td>
<td>ACCOUNT</td>
<td>Benefit Unit</td>
</tr>
<tr>
<td>A23</td>
<td>ACCOUNT</td>
<td>Value of National Savings investment</td>
</tr>
</tbody>
</table>

Excel spreadsheet showing the data for variable ADCH.
Table 6.3 Full spreadsheet information for variable ADCH

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Details</td>
<td>TABLE</td>
<td>VARIABLE</td>
<td>VAR_FMT</td>
<td>LABEL</td>
<td>FRMVALUE</td>
<td>FMTVALUE</td>
<td>MINVAL</td>
</tr>
<tr>
<td>Metadata</td>
<td>ACCOUNTS</td>
<td>ADCH</td>
<td>ACS_302X</td>
<td>Child or adult a/c?</td>
<td>1</td>
<td>Adult</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>J</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Details</td>
<td>MAXVAL</td>
<td>DERIVED</td>
<td>BENEFIT</td>
<td>QUESTION</td>
<td>TYPE</td>
<td>BLOCK</td>
<td>USAGE</td>
</tr>
<tr>
<td>Metadata</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>ADCH</td>
<td>(I)</td>
<td>BINTREST</td>
<td>C</td>
</tr>
</tbody>
</table>

Figure 6.8 HTML information for variable ADCH

ADCH

Whether adult or child account

Minimum value 1
Maximum value 2
Period code variable N/A
Benefit key Non-benefit variable
Question ADCH
Type Integer
Block BINTREST
Sequence of questions N/A
Usage C

SAS format ACS_302X

Values Labels
1 Adult
2 Child
Figure 6.9 Structure of pilot FRS website

[Diagram showing the structure of a pilot FRS website with categories such as Introductory Documentation, Detailed Documentation, Background information, Structure of database, Programming examples, Imputation information, Questionnaire information, Hierarchical dataset, Variable searches, Accounts, Adult, Assets, Benefits, Benefit, Care, Child, Child, Display, Endowment, ExtChild, Household, Insuranc, Job, M4nt, MortCont, Mortgage, OddJob, Owner, Pension, RentCont, Renter, Vehicle, etc.]
Outcomes of pilot project

Figure 6.9 shows the structure of the pilot website. (Note: solid lines indicate single pages and links; lines and dots indicate that the box or line represents multiple pages or links.) After I had completed a substantial amount of the pilot documentation, including putting online all of the Guide and part of the metadata documentation, I gave a presentation to the FRS Users Group on the work that I had been doing (see section 1, above). The purpose of this was so that I could gain feedback on content, design, and layout; it also served to provide legitimacy for my work within the department.

The issues which emerged from this presentation and my ongoing discussions with the FRS project manager fell into two categories: feature-related issues, i.e. suggestions for additions to the functionality of the site; and technical/structural issues, i.e. concerned with the design or layout of the site.

New features

- **Extended search facilities for the metadata documentation.** Drawing on Card (2000), I had already considered adding further search facilities in order to make the metadata documentation more useable. In particular, I wanted to add the ability to search the information by topic and classification. The FRS project manager agreed that the search facilities should be extended in this way.

  This proved to be an extensive sub-project in its own right. It involved the development and implementation of a classification scheme for the search function. In addition, I had to learn some JavaScript in order to be able to carry out the work. I shall discuss how I developed these search facilities in greater depth in the following section 5.
Technical/structural issues

- **Problems encountered in naming files.** As I created individual pages for each variable, I realized that a handful of variables shared a name with a table (e.g. there is an ACCOUNTS table and an ACCOUNTS variable and a CARE table and a CARE variable). Since I was storing all the HTML files in the same location, and was naming them according to their variable or table name, this would lead to duplication in file names (which did not matter in the Excel file). The project manager indicated that these common variable names were being removed through various iterations of the survey, but duplications still remained in FRS 35.

  My interim solution was to give precedence to table files, and to indicate variable files with the suffix '_var': thus, the file containing the information for the ACCOUNTS table was named 'accounts.htm'; the file for the ACCOUNTS variable was named 'accounts_var.htm'. This, I hoped, would solve the technical problem without confounding user expectations too much.

  Another, related problem emerged when I realized that there was duplication in the names of some variables under different tables. Some of these variables were simply repeated (e.g. _MONTH_) and a single variable HTML page would suffice: the ability of hypertext to link to a single file from multiple locations assisted in avoiding duplication here. However, some of these variables were related to different tables. Again, this would lead to duplication in file names. My solution - again, trying not to work against user expectations - was to give the variable name a suffix indicating with which table it was associated. For example, two variables entitled ANYMON are associated with both the ADULT and the ASSETS tables. The file names were thus 'anymon_adult.htm' and 'anymon_assets.htm' respectively.
• *Increased access to metadata documentation.* I had originally only provided links to the information on usage codes, benefit codes and period codes through the variable pages. I added new links on the site's front page to make these commonly used pages more immediately accessible for users.

• *Modifications to layout on table and variable pages.* Throughout this stage of development, I had obtained feedback from the project manager at our regular meetings on the layout of both the table and variable pages. As I had been coding the table pages, it became clear that these pages could contain lists of up to more than 400 variables (see Table 6.2), and both the project manager and I agreed that a plain alphabetical list of variables was insufficiently user-friendly.

The metadata Excel file (which was generated automatically from the BLAISE program) categorized variables in four ways: key variables (variables used to locate individual responses - the same across all tables); data variables (which hold data information); derived variables (variables which are created in the process of using the questionnaire, calculated from one or more other variables); system variables (which store system-related information). I reorganized the table pages to reflect this differentiation between the variables, and added links within the pages to enable navigation up and down the page. In chapter 7, I shall discuss the classification system (and the one devised by the project manager for the search facilities) in terms of Latour's (1990) work on 'cascades of inscription' (i.e. the process whereby classifications take on new forms without altering their basic structure).

Figure 6.10 [accounts.htm] shows the new layout for the ACCOUNTS page (note that there are no derived variables in the ACCOUNTS table; 'data' is highlighted to show a link). Comparing this with Figure 6.5 above shows the changes made. (Note also the changed logo from the Department of Social Security to the Department for Work and Pensions.) I made similar changes to the layout of the individual variable pages, to group together metadata which were more closely associated with each other. Figure 6.11 shows these changes to the ADCH page [adch_accounts.htm]; Figure 6.8 shows the original layout.
ACCOUNTS

Income from interest/dividend bearing assets and savings together with (for a subset of records) the value of National Savings products for the accounts/investments held by adults and children. Each record relates to a type of investment (current account, savings account etc). Adults/children may have more than one type of investment, each record giving the total interest/dividends received (if they have more than one account of that type). For National Savings products, if the adult/child is not routed into the assets questions, a banded figure for the value of the investment is collected (those entering the assets block will have an accounts record but with this variable skipped).

Variables

Key Data System

Key variables

1. SERNUM Serial number
2. BENUNIT Benefit Unit
3. PERSON Person number within household
4. ACCOUNT Account type

Data variables

ACCOUNT Interest received
ACCTAX Is that interest before or after tax
ADCH Whether adult or child account
NSAMT Value of National Savings investment

System variables

_MONTH_ Month code (source)
Figure 6.11 *New layout for ADCH page*

<table>
<thead>
<tr>
<th>Label</th>
<th>Whether adult or child account</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key</td>
<td>N/A</td>
</tr>
<tr>
<td>Source</td>
<td>Base</td>
</tr>
<tr>
<td>Question</td>
<td>ADCH</td>
</tr>
<tr>
<td>Block</td>
<td>BINTREST</td>
</tr>
<tr>
<td>Minimum value</td>
<td>1</td>
</tr>
<tr>
<td>Maximum value</td>
<td>2</td>
</tr>
<tr>
<td>Type</td>
<td>Integer</td>
</tr>
<tr>
<td>Usage</td>
<td>C</td>
</tr>
<tr>
<td>Format</td>
<td>ACS_302X</td>
</tr>
<tr>
<td>Period code variable</td>
<td>N/A</td>
</tr>
<tr>
<td>Benefit key</td>
<td>Non-benefit variable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Values</th>
<th>Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adult</td>
</tr>
<tr>
<td>2</td>
<td>Child</td>
</tr>
</tbody>
</table>
5 Task Three: Design search facilities

As I wrote above, Card (2000) indicates two types of search facility which users of social science data are likely to want: by variable name and according to topic. The simple search facility I had included in the pilot documentation enabled users to find variables according to name (by grouping links to individual pages in alphabetical order). However, this depended on users knowing exactly which variable they were looking for. I next turned to developing a means to search for variables according to topic. This would permit a much more open-ended search which was not dependent on users' knowing the names of variables for which they were searching.

The first stage in developing the search facilities was to work out a classification system for likely topics. The FRS project manager constructed this, based on her knowledge of how users searched the FRS metadata. She isolated twenty top-level categories, some of which had subcategories. This classification system was intended to reflect specific topic interests users were likely to have (variables related to council tax, to employment, to assets and savings, to benefits etc.). She classified each variable on the 1998-1999 FRS dataset according to this system. Table 6.4 shows the topic classification system for the variable topic search, with first-level and second-level categories shown.

Based on this classification system, I devised two ways of searching the metadata. Firstly, I coded means whereby users could enter terms into a search engine, in order to find relevant variables. Secondly, I coded a search tool whereby users could choose, using a drop-down box, a first-level then a second-level classification, and then be linked to related variables. I shall discuss this process in terms of Latour's (1990) work on 'inscription' in chapter 7, where I also discuss the construction of the search facilities in terms of collaborative and anonymous authorship in a bureaucratic setting.

Figure 6.12 shows the page from the site on which the search facilities are located (with the search term 'interest' entered under topic search). Figure 6.13 shows the results of this search, which open up in a new window. Figure 6.14 shows the drop-down box for the classification scheme under search by classification (Figure 6.12 shows the second-level classifications for the first-level classification 'Assets and savings'; Figure 6.14 shows the second-level classification for the first-level
classification 'Demographic characteristics'). [Search facilities are accessible from: frshdtopicsearch.htm.]

Table 6.4 *Topic classification for variable topic search*

<table>
<thead>
<tr>
<th>Top-level classification</th>
<th>Second-level (where applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Assets and savings</td>
<td>Accounts and Investments held</td>
</tr>
<tr>
<td></td>
<td>Capital value</td>
</tr>
<tr>
<td></td>
<td>Interest and dividends</td>
</tr>
<tr>
<td>2. Care</td>
<td>Childcare</td>
</tr>
<tr>
<td></td>
<td>Informal care</td>
</tr>
<tr>
<td>3. Consumer durables</td>
<td></td>
</tr>
<tr>
<td>4. Council Tax</td>
<td></td>
</tr>
<tr>
<td>5. Demographic characteristics</td>
<td></td>
</tr>
<tr>
<td>6. Employment</td>
<td>Earnings</td>
</tr>
<tr>
<td></td>
<td>Employment Status</td>
</tr>
<tr>
<td></td>
<td>Job description</td>
</tr>
<tr>
<td></td>
<td>Self-employment income</td>
</tr>
<tr>
<td>7. Health</td>
<td>Children's health</td>
</tr>
<tr>
<td></td>
<td>Health restrictions on work</td>
</tr>
<tr>
<td>8. Housing Benefit</td>
<td></td>
</tr>
<tr>
<td>9. Housing costs</td>
<td>Charges on property</td>
</tr>
<tr>
<td></td>
<td>Mortgages</td>
</tr>
<tr>
<td></td>
<td>Rent</td>
</tr>
<tr>
<td></td>
<td>Structure and contents insurance</td>
</tr>
<tr>
<td></td>
<td>Water and sewerage</td>
</tr>
<tr>
<td>10. Informal care</td>
<td></td>
</tr>
<tr>
<td>11. Insurance policies</td>
<td></td>
</tr>
<tr>
<td>12. Maintenance</td>
<td>Maintenance paid</td>
</tr>
<tr>
<td></td>
<td>Maintenance received</td>
</tr>
<tr>
<td>13. NHS services</td>
<td></td>
</tr>
<tr>
<td>14. Non-state pensions</td>
<td>Income from pensions, trusts and annuities</td>
</tr>
<tr>
<td></td>
<td>Pension scheme membership</td>
</tr>
<tr>
<td>15. Other Income</td>
<td>Children's earnings</td>
</tr>
<tr>
<td></td>
<td>Educational grants/loans</td>
</tr>
<tr>
<td></td>
<td>Income from property</td>
</tr>
<tr>
<td></td>
<td>Non-state benefits</td>
</tr>
<tr>
<td></td>
<td>Odd jobs</td>
</tr>
<tr>
<td></td>
<td>Royalties and allowances</td>
</tr>
<tr>
<td></td>
<td>Welfare/school milk/meals</td>
</tr>
<tr>
<td>16. Qualifications</td>
<td></td>
</tr>
<tr>
<td>17. Social Security Benefits/Tax Credits</td>
<td></td>
</tr>
<tr>
<td>18. Tenure</td>
<td></td>
</tr>
<tr>
<td>19. Travel to work</td>
<td></td>
</tr>
<tr>
<td>20. Unclassified</td>
<td></td>
</tr>
</tbody>
</table>
For the search tool for searching by topic, I adapted a piece of freeware JavaScrip which carried out the kind of search I wanted to implement on the site. This involved my coding the script so that it contained all variable names associated with the relevant classification topics, and then connecting this information to the appropriate HTML variable page. This was another fairly laborious coding process, and I worked on it both on and off site with little external input other than to indicate my progress to the FRS project manager.

Table 6.5 shows the workings of the JavaScript for five different variables. Lines beginning 'title' show the search terms entered for each variable (compare with the topic classification system for 'Assets and Savings'); lines marked 'desc' give the variable description (the label from the metadata); lines marked 'links' point to the variable HTML file. A user conducting a search using the word 'interest' would be given the result shown in Figure 6.13 below, providing links to the variables ACCINT and ACCTAX.

**Table 6.5 JavaScript showing how search terms are entered for variables in the topic search**

```
title[2]="assets savings interest dividends accint"
desc[2]="Interest received"
links[2]="accint.htm"
matched[2]=0

title[3]="assets savings accounts account"
desc[3]="Account type"
links[3]="account.htm"
matched[3]=0

title[4]="assets savings interest dividends tax acctax"
desc[4]="Is that interest before or after tax?"
links[4]="acctax.htm"
matched[4]=0

title[5]="assets savings adult child adch"
desc[5]="Whether an adult or child account"
links[5]="adch_accounts.htm"
matched[5]=0

title[6]="assets savings capital value national nsamt"
desc[6]="Value of National Savings Investments"
links[6]="nsamt.htm"
matched[6]=0
```
Figure 6.12 Search facilities page

Variable search by classification
You can carry out a search according to classification.
Selecting a topic from the first level drop-down box will offer you second-level topics.
Click on the links to access details about which variables are associated with these topics.

Select first level: 1. Assets and savings

Second level:
1. Accounts and investments held
2. Capital value
3. Interest and dividends
4. Other

Variable search by topic
This page links to all variables according to topic.
You can input multiple keywords by separating each keyword with a +. The search is non-case sensitive.

Keywords: | interests |
Click To Start Search

Variable search by name
This list links to all variables alphabetically.

A B C D E
F G H I J
K L M N O
P R S T U
V W Y
Figure 6.13 Results of search on 'Interest'
Variable search by classification

You can carry out a search according to classification.

Selecting a topic from the first level drop-down box will offer you second-level topics.

Click on the links to access details about which variables are associated with these topics.

Select first level:

5. Demographic characteristics

Second level:

1. Adult
2. Child
3. Other

Variable search by name

This list links to all variables alphabetically.

A B C D E F G H I J K L M N O P R S T U V W Y
I also developed the classification search facility using JavaScript, based on examples from web design manuals. Table 6.6 shows the code I developed for producing this function. The first half of the code shows how the first-level classifications were put into the program as search terms. The second levels were grouped together - the code here shows only the second-level classifications for 'Assets and savings' (i.e. Accounts and investments held; Capital value; Interest and dividends; Other - see Table 6.4 above).

Table 6.6 HTML and JavaScript code showing how search terms are entered for variables in the classification search

```html
<p class="prose">Select first level:
<select onchange="firstLevelItemChanged(this.selectedIndex)">
  <option>1. Assets and savings
  <option>2. Care
  <option>3. Consumer durables
  <option>4. Council tax
  <option>5. Demographic characteristics
  <option>6. Employment
  <option>7. Health
  <option>8. Housing benefit
  <option>9. Housing costs
  <option>10. Insurance policies
  <option>11. Maintenance
  <option>12. NHS services
  <option>13. Non-state pensions
  <option>14. Other income
  <option>15. Other social security benefits/tax credits
  <option>16. Qualifications
  <option>17. Tenure
  <option>18. Travel to work
</select></p>
<p class="prose">Second level:</p>
<div id="div1" class="prose">
  <ol>
    <li><a href="topicsearchassets1.htm">Accounts and investments held</a></li>
    <li>Capital value</li>
    <li>Interest and dividends</li>
    <li>Other</li>
  </ol>
</div>
```

168
Figure 6.15 shows the output a user would receive after choosing the first-level classification 'Assets and Savings' and the second-level classification 'Interest and dividends'. I wrote new pages in HTML such as this (this one is topicsearchassets3.htm, as shown in the code above) to be the output of the search. These pages listed and linked to the relevant variables under each classification; here the links are to the variables ACCINT and ACCTAX.

Figure 6.15 Output of classification search
6 Task Four: Design and implement full documentation

Revisiting the purpose of the project

The coding, testing, and finessing of the search facilities was being completed at the time that the FRS project manager left, and the new manager was appointed. As I stated in the introduction to this chapter, there was a brief hiatus in development at this time, and then the new project manager and I met to clarify our aims for the project and to discuss how we wanted the project to proceed. I saw the most pressing task as being to devise a design and layout for the online questionnaire documentation.

From this point onwards, we included in our discussions a member of the ASD IT department, in order to get his advice on how my designs could be best implemented from the point of view of the IT staff who would share some responsibility for maintaining the documentation after this project ended. The outcome of our discussions was that we defined the project in terms of providing a prototype, using one survey year as an example, upon which year-by-year extensions could be built. The member of the IT department was most interested in receiving user feedback (which time constraints did not usually allow technical staff to get) on a prototype which gave users a representation of how documentation might look, rather than in having the project provide specific technical solutions, which constituted the IT department's own area of expertise. The remainder of the development work was therefore primarily dedicated to building a fully documented version of a single survey year, and acquiring feedback from the FRS project manager and the IT staff member throughout this process. Both individuals also had comments related to the layout of the site, and refinements to this were ongoing throughout the remainder of the project. Soon after the new project manager took over this job, I began to work primarily off site; this was in part because of a shortage of space at the DWP, and also partly because I was able to work on the questionnaire documentation as easily off site. I continued to go to the DWP on a regular basis (every 4-6 weeks), sending files containing changes to the site for review by the project manager and the IT staff member in advance of our meetings via email, or by posting sections of the site online.
My work at this point, then, became concerned with issues of implementation, and what I could reasonably achieve in the time remaining to the project. I was not under very pressing time constraints, but the project did have a set time period (3 years), which was, by this point, half-way through. The project manager, the IT staff member and I took this into account as we were determining tasks from this point on. We were, in particular, concerned with what users wanted for the site. This was a very significant point for the project, where the definition of the project shifted markedly from an emphasis on design principles and investigating academic models of hypertext towards an emphasis on perceived user needs (see chapter 8, pp. 241-243).

Prior to the new FRS project manager taking over, I had conducted software testing for the ONS of the TADEQ tool, which had been developed as part of a project aiming to provide a standard format for online documentation of social survey questionnaires. As my own project work shifted its focus towards developing the online questionnaire documentation for the FRS, I reviewed the various standards which were emerging at this time, including TADEQ. I shall briefly describe these, and explain why I elected not to use them. This can be seen as another point in the development where the purpose of the project was redefined (I discuss this in the conclusion to this chapter).

**Review and decide on standards**

Standards set by communities for developing computer applications are a significant way in which this development can be constrained. As discussed in chapter 4 (pp. 93-94), two markup standards for electronic questionnaire documentation have been developed, both of which were potentially appropriate for use in documenting the FRS. Both standards are based on XML (*Extensible Markup Language*), a 'next-generation' markup language which is a dialect of a more general markup language, SGML (*Standardized General Markup Language*).

The first standard is the Data Documentation Initiative Document Type Definition (DDI DTD), which I had encountered in my preliminary researches of how other government statistical bodies were responding to the need they perceived for documentation for electronic questionnaires. The Data Documentation Initiative
(DDI)\textsuperscript{57} is an international consortium of academic researchers and government statistical offices (including the ESRC Data Archive, which holds records of FRS data). The DDI aims to establish an international criterion and methodology for the content, presentation, transport, and preservation of metadata about datasets in order to produce codebooks which are uniform, highly structured, and are easily searchable on the Web. To this end, the DDI has developed what is known as a Document Type Definition (DTD) for the markup of social science codebooks.

The second XML DTD, the Questionnaire Definition Language (QDL) was developed as part of the TADEQ project. This project (Bethlehem and Manners, 1998) was a collaborative R&D project funded by the European Union ESPRIT Programme, involving National Statistical Institutes, research institutes, and commercial marketing research organizations (TADEQ: a Tool for the Analysis and Documentation of Electronic Questionnaires). The project was aimed at developing a tool to make a human-readable presentation (on paper or electronically in hypertext format) of the electronic questionnaire. The project focused on the development of its own XML DTD: the Questionnaire Definition Language (QDL), and a prototype of the tool was released in 2001. A main player in the TADEQ project is the Social Survey Division of the Office of National Statistics (ONS), who have conducted the field research for the FRS since its inception.

I decided not to use either of the XML standards available. My decision was based on two main criteria. Firstly, XML is a technology still in flux, the standards for which themselves are currently in the process of being set\textsuperscript{58}. As a consequence of this, not all web browsers are capable of processing XML easily.

Secondly, I had found, during testing I had conducted of the TADEQ tool, that the standardization process led to some loss of detail familiar to users of the FRS. For example, in the metadata documentation, the terminology used for the answer types by TADEQ differed from that used in the Excel metadata documentation produced specifically for the FRS. TADEQ groups variables according to six types (numeric, real, numeric integer, open, closed, date, and time variables). The Excel metadata file generated from BLAISE for the FRS organizes this information differently: eleven types of variable are listed in this documentation (categorical, date, frequency, key,

\textsuperscript{57} See http://www.icpsr.umich.edu/DDI/codebook.html.

\textsuperscript{58} See the information on XML at the World Wide Web Consortium's website: http://www.w3.org/.
monetary, period code, quantitative, string, date components, weekly, and system variables). In addition, the variables are categorized according to usage. While TADEQ provides the same information, the necessity to provide a solution which applied generally to a broad range of surveys meant that survey-specific information was lost.

I therefore came to the decision that, since a standardized version of the FRS based on the QDL DTD was available through the TADEQ tool, there was scope for the development of an online version of the FRS documentation dedicated solely to the FRS. I would also suggest that this was in part my attempt to give this project a distinctive identity beyond the potential solutions which were being offered by the TADEQ tool, i.e. this was another point of redefinition for project; I shall return to this in chapter 8, pp. 239-241.

However, the TADEQ tool did influence my development work, specifically in certain aspects of the presentation of the questionnaire documentation. The user of TADEQ is presented with an online version of the questionnaire which they can reduce (e.g. by hiding all conditions) to make the documentation more manageable, and then expand in order to examine sections of the questionnaire in greater detail. I adapted this presentational aspect into the design of the questionnaire for the current project, as I shall describe in more detail below.

**Design and implement questionnaire documentation**

In order to make the questionnaire documentation speedier to use, I broke it up and organized it according to its three main constituent parallel blocks: the household schedule, the benefit unit schedule, and the assets block (see chapter 5, Table 5.4, pp. 119-120 for detailed information on the content of these blocks).

Figure 6.16 shows my design for the front page of the section of the website devoted to the questionnaire documentation [fisuquest.htm], divided into the three parallel blocks. (Note the links on the left-hand side to the documentation home page, the help file, and the full documentation for the 1998-1999 survey. These structural changes to the organization of the site were carried out after discussions with the FRS project manager and the IT staff member, and I will discuss them in more detail below.)
I shall use the example of the assets block to describe in more detail how I designed and implemented the questionnaire documentation. The assets block for the 1998-1999 survey consists of 12 question blocks, as summarized in Table 6.7.
Table 6.7 Question blocks in the assets block

<table>
<thead>
<tr>
<th></th>
<th>Assets block</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCurrAC</td>
<td>Amount in current account(s)</td>
</tr>
<tr>
<td>QASaveAC</td>
<td>Details of accounts</td>
</tr>
<tr>
<td>QAEquity</td>
<td>Investments/shares/bonds: details and value</td>
</tr>
<tr>
<td>QACertif</td>
<td>National Savings Certificates: details and value</td>
</tr>
<tr>
<td>QPG1B</td>
<td>Pensioners' Guaranteed Income Bonds: details and value</td>
</tr>
<tr>
<td>QSaye</td>
<td>SAYE schemes: details and value</td>
</tr>
<tr>
<td>QPremium</td>
<td>Premium Bonds: details and value</td>
</tr>
<tr>
<td>QNSIB</td>
<td>National Savings Income Bonds: details and value</td>
</tr>
<tr>
<td>QABonds</td>
<td>Bonds: details and value</td>
</tr>
<tr>
<td>QCBonds</td>
<td>Bonds: details and value</td>
</tr>
<tr>
<td>QFirstOp</td>
<td>Bonds: details and value</td>
</tr>
<tr>
<td>QYPlan</td>
<td>Bonds: details and value</td>
</tr>
<tr>
<td></td>
<td>End of the assets block</td>
</tr>
</tbody>
</table>

Within each block, the documentation is made up of three main elements. These are (detailing the information which each element contains):

1. **Question text**: The FRS involves proxy interviews, i.e. answers may be given to questions on behalf of people. The software automatically inserts pronouns and verb tenses which are appropriate to whether it is a personal or proxy interview.

2. **Answer types**: Variables on the FRS can be any of a number of types (e.g. numeric real, numeric integer, open, closed, date, and time variables), and have a number of answer formats related to their type (i.e. integer ranges; numeric ranges; text string lengths; full details of the numbers and values of options available in closed variables; information on whether special answers were available and what these answers were).

3. **Conditions: checks and computations**: Checks on information provided are either hard (which prevent the interview from continuing until the information has been verified, or is consistent with prior answers), or soft, which provide an alert to the interviewer, but can be overridden to allow the interview to proceed. In addition, the electronic questionnaire allows computations to be conducted within the course of the interview easily. Showing conditions for computations and details of checks allows a user to follow the route of the questionnaire or the path a particular interview took.
Figure 6.17 Documentation for the assets block - all elements hidden

Figure 6.17 shows my original version of the online documentation of the assets block. All elements in the documentation are hidden. I organized the page as a table, with the left-hand cell containing links, and the right hand cell (the bulk of the page) containing the global instructions and the questionnaire documentation itself.

I grouped the question block names together on the left-hand side, and these provide links which enable the user to move up and down the page to the relevant question block (FRS35.Assets.QPremium[] is highlighted in red to show that it is a hyperlink). Other links on the left-hand side of the page take the user out of the questionnaire documentation - either to the documentation home page, or to the help
file, which I coded to open in a separate window. I also linked to the section of the help file directly concerned with using the questionnaire documentation.

I grouped the global instructions at the top of the page: these enable the user to show or hide each or all of the constituent elements of the questionnaire by clicking on the relevant instruction.

The main body of the page shows the list of question blocks in bold, black type. Clicking on this text opens up the elements of an individual question block for use.

Figure 6.18 shows the documentation of the assets block with question text showing for the question block FRS35.Assets.OCurrAC[].Account. The question names are given in blue text under the bold, black heading of the question block. Within this question block, then, there are 9 questions in total.

Clicking on the blue text expands the individual question to show all the elements related to that question, including routing, the full text of the question asked in the interview, and details of the types of answers that can be given. Figure 6.19 shows the documentation with all items expanded for the question FRS35.Assets.OCurrAC[].Account[].AnyMon. Above the blue text of the question name is the routing to show the conditions under which the question is asked. Below the blue text of the question name is the full text of the question asked (including notes to the interviewer) and the answer types available for the interviewer to input. Clicking on the blue text FRS35.Assets.OCurrAC[].Account[].AnyMon would contract the question and hide all this information; clicking on any other blue text question name would expand this information for that particular question. Clicking on the bold black text FRS35.Assets.OCurrAC[].Account[] would contract and hide all the question names; clicking on any other bold black text would expand that particular question block.

I made sure that the styles of the text for the conditions, question text, and answer types follow as closely as possible the styles used in the BAD so that FRS users would be presented with a familiar style.
Figure 6.18 Documentation for the assets block - question text showing
Figure 6.19 Documentation for the assets block - all elements showing for one question

```
ASK IF: In loop FOR PersCont := 1 TO HHSizex
   AND: AssElig[PersCont] = 2
   AND: PA[PersCont][1].Asset = Yes
   AND: In loop FOR pcount := 1 TO 4
       AND: (pcount = 1) OR ((pcount > 1) AND (Account[pcount - 1].Mbre = Yes))
       AND: pcount = 1
   FRS35.Assets.QCurrAC[].Account[].AnyMon

   Now I'd like to ask you about your current account(s): At the end of last (month/pay period), did you have any money left in your current account, after your household expenditure?

   INTERVIEWER: THIS INCLUDES ANY JOINT ACCOUNTS
   (1) Yes - money in (one or more) accounts
   (2) No - no money in any current account
   (3) No longer have any current account(s)

   FRS35.Assets.QCurrAC[].Account[].AccName
```
Having shown the layout and functionality of the documentation, I shall now give an example to show how I wrote the documentation. Table 6.8 shows the HTML and JavaScript code which produces the documentation for the question

FRS35.Assets.OCurA[]].Account[].AnyMon (Figure 6.19). The <div>...</div> tags break the file into divisions upon which specific functions can be carried out.

The <div> and <p> tags contain 'attributes'. The class attribute directs the page towards a stylesheet set up to control the appearance of the text of the various elements in the documentation. In this case, the style names show the constituent element from the questionnaire: 'fullblockname', 'questionname', 'condition', 'condition2', 'questiontext', 'answertype'. The attribute id gives the division a unique name (an identity). I chose to label divisions in this way because it would allow each division name within the documentation to be stored on a database: from the perspective of the IT staff, this would then enable them to move towards the automatic production of the documentation. From a theoretical perspective, I would also note how any automation of the documentation arising from the way that I labelled the divisions would therefore be re-inscribing the attribute names - which I had, in turn, derived from the BAD.
Table 6.8 HTML and JavaScript code for producing FRS35.Assets.QCurrAC[].Account[].AnyMon

1. `<div class="fullblockname" id="FRS35.Assets.QCurrAC[].Account[]">`  
   `<span onMouseOver="window.status='Click here to expand and contract block'; return true" onMouseOut="window.status='"; return true" onclick="showHide('FRS35.Assets.QCurrAC[].Account[].Body')">`  
   `FRS35.Assets.QCurrAC[].Account[]</span>`

2. `<div id="FRS35.Assets.QCurrAC[].Account[].Body">`

3. `<div class="questionname" id="FRS35.Assets.QCurrAC[].Account[].AnyMon">`

4. `<div class="condition" id="FRS35.Assets.QCurrAC[].Account[].AnyMon.Condition">`  
   `<p>ASK TP: In loop FOR PersCont := 1 TO HHSSize</p>`  
   `<p class="condition2" AND: AssElig[PersCont] = 2</p>`  
   `<p class="condition2" AND: Ph[PersCont].A[1].Asset = Yes</p>`  
   `<p class="condition2" AND: In loop FOR pcount := 1 TO 4</p>`  
   `<p class="condition2" AND: (pcount = 1) OR ((pcount &gt; 1) AND (Account[pcount - 1].More = Yes))</p>`  
   `<p class="condition2" AND: ppcount = 1</p>`

5. `<p class="questionname">`  
   `<span onMouseOver="window.status='Click here to expand and contract section'; return true" onMouseOut="window.status='"; return true" onclick="conditionalShowHide('FRS35.Assets.QCurrAC[].Account[].AnyMon')">`  
   `FRS35.Assets.QCurrAC[].Account[].AnyMon</span>`

6. `<div class="questioncontext" id="FRS35.Assets.QCurrAC[].Account[].AnyMon.QuestionText">`  
   `<p>Now I'd like to ask you about your current account(s):</p>`  
   `<p>At the end of last (month/pay period), did you have any money left in your current account, after your household expenditure?</p>`  
   `<p>INTERVIEWER: THIS INCLUDES ANY JOINT ACCOUNTS</p>`

7. `<div class="answertype" id="FRS35.Assets.QCurrAC[].Account[].AnyMon.AnswerType">`  
   `<p>(1) Yes - money in (one or more) account(s)</p>`  
   `<p>(2) No - no money in any current account</p>`  
   `<p>(3) No longer have any current account(s)</p>`

8. `<hr>`

9. `</div>`
Notes to Table 6.8

1. Division which indicates by means of the id attribute that all that follows is related to the question block FRS35.Assets.OCurrACfl.Accountn. The stylesheet class ‘fullblockname’ produces the bold black text as shown in Figure 6.17. The JavaScript commands within the <span>... </span> tags enable the function whereby clicking on that text expands and contracts the block.

2. Division which is identified by means of the id attribute as containing the elements of FRS35.Assets.OCurrACfl.Accountn.Body. The .Body suffix acts as an indicator, when the JavaScript program runs, that all that follows within this division should have specific functions carried out on it; i.e. this marks out the area that will be expanded and contracted when the block name FRS35.Assets.OCurrACfl.Accountn is clicked.

3. Division which indicates by means of the id attribute that all that follows is related to the question FRS35.Assets.OCurrACfl.Accountn.AnyMon. (As shown in Table 6.7, there are 12 question blocks within the assets block. For reasons of space, this example only shows the block OCurrACfl.Accountn; other blocks are built up in the same way, and nested within the FRS35.Assets.OCurrACfl.Accountn.Body division.)

4. Division which indicates by means of the id attribute that all which follows is related to the condition FRS35.Assets.OCurrACfl.Accountn.AnyMon.Condition. The .Condition suffix acts as an indicator, when the JavaScript program runs, that all that follows is subject to the global instruction 'Show and hide conditions'. The stylesheet class ‘condition’ produces the italic text and the stylesheet class ‘condition2’ produces the indented italic text (see Figure 6.19).

5. Division which indicates by means of the id attribute that all that follows is related to the question FRS35.Assets.OCurrACfl.Accountn.AnyMon. The stylesheet class ‘questionname’ produces the blue text (see Figure 6.18). The JavaScript commands within the <span>... </span> tags enable the function whereby clicking on that text expands and contracts the question.

6. Division which indicates by means of the id attribute that all which follows is related to the condition FRS35.Assets.OCurrACfl.Accountn.AnyMon.QuestionText. The .QuestionText suffix acts as an indicator, when the JavaScript program runs, that all that follows is subject to the global instruction 'Show and hide question text'. The stylesheet class ‘questiontext’ produces the plain text (see Figure 6.19).

7. Division which indicates by means of the id attribute that all which follows is related to the condition FRS35.Assets.OCurrACfl.Accountn.AnyMon.AnswerType. The .AnswerType suffix acts as an indicator, when the JavaScript program runs, that all that follows is subject to the global instruction 'Show and hide answer type'. The stylesheet class ‘answertype’ produces the reduced plain text (see Figure 6.19).


182
7 Task Five: Acquire and act on user feedback

I obtained user feedback on the full documentation in two ways. Firstly, the FRS project manager and the IT staff member reviewed a beta version of the full documentation, and I implemented their feedback on the site. These changes were predominantly concerned with restructuring the organization of the site. I then made this version of the site generally available for users within ASD to evaluate, and conducted short interviews and discussions with users, implementing some of the changes they suggested.

Restructuring the site

So far, I had been expanding the site in a relatively piecemeal fashion, adding new features as they were developed, and following a largely hierarchical structure (see Figure 6.9; compare with Figure 6.1). At this stage, I reorganized the structure of the site: the content had greatly increased, and this would also allow for future year-by-year additions to the site. I outlined three areas for restructuring: I put new layers in the organizational structure; I wrote help files and linked to them from each page; and I put links in place between the questionnaire and metadata documentation.

New layers in structure

- The new front page (Figure 6.20 [docufront.html]) demonstrates the main principle which informed my restructuring of the site: grouping together documentation which holds for each year, and grouping together documentation specific to each year (compare this with Figure 6.2, which shows the pilot documentation home page). My intention was that links could be added to this page as new survey years were documented.
These pages gather together some of the documentation available which relates to the Family Resources Survey. General documentation gathers together information which holds for each survey, year-on-year. Version specific documentation is gathered together by survey year.

Help for using this documentation is here. The pages are under constant development. Please contact Una McCormack with any comments or suggestions.

**General documentation**

- Information related to each survey year.
- Background to the survey
- Structure of the database
- Programming examples
- Imputation flags
- Style guide to the current documentation
- Questionnaire description and instructions

**Version specific documentation**

Year-on-year documentation of the questionnaire, the database and the metadata. Click on the appropriate survey year.

- 1998-1999

© Crown
Within the general documentation, this led in effect to a new organizational 'layer' in the hierarchy of the site. Figure 6.21 shows the new page for the structure of the database [frstruct.htm]; Figure 6.3 shows this page in the pilot documentation. On the pilot page, the links on the left-hand side led to pages external to the documentation site. On this restructured page, I replaced the links on the left-hand side with links to pages that fall within the hierarchical layer of the general documentation. Links to the external pages remain on the documentation home page (amended to reflect the departmental reorganization).

In addition, I made the version-specific pages and the questionnaire home page more interconnected. Figure 6.22 shows the 1998-1999 documentation home page [1998_99-front.htm], with links to both the questionnaire documentation specific to that year, and to the general questionnaire documentation home page (Figure 6.16 [frquest.htm]). This page links back to specific years (I organized it to allow additions on a year-by-year basis). Again, my intention was that links on the left-hand side of both these pages could be added on a year-by-year basis; this constituted another 'layer' in the structure of the site (see Figure 6.1).
Structure of the FRS Database

Introduction
Hierarchical Tables
Flatfile

Introduction

The FRS datasets are stored in /data1/frsYYyy on UNIX box 14 (Londo03), where YYyy is the survey year (e.g. 0001 for 2000-01), and ~ is the letter for the latest version of that release. These are copied to /repdata.sas_surveys/frs/frsYYyy on Londo02.

When data is updated a new release is created for users. The old release is not deleted, so can still be accessed, but the new updated release should be the data which is used by all users. For further information on latest releases, click here.

The FRS data exists in both a series of hierarchical tables and also in the form of a flatfile.

Hierarchical Tables

Background
Key variables
Normalization principle
Table description

Background

The FRS database consists of approx. 24 hierarchical, normalized tables, each table relating to a particular level (e.g. household, benefit unit) or type of information (e.g. demographic, survey data).
Figure 6.22 1998-1999 documentation home page

FRS 1998-1999
Full documentation of the 1998-1999 FRS.

Questionnaire documentation
Questionnaire: detailed documentation of
1. Household schedule
2. Benefit unit schedule
3. Assets block

Database documentation
Hierarchical dataset: full description
Search hierarchical dataset: by classification, topic, and variable name

Metadata documentation
Benefit key
Benefit map
Period codes
Usage codes

© Crown
Help file

Because the site had now become relatively complex, I wrote a help file to assist users in their navigation of the site. I put a link to this help file from all pages, regardless of layer within the new site structure (see Figures 6.21 and 6.22: the positions of the links to the documentation home page and the help file are kept constant on all pages; see also Figure 6.17 [fishassets.htm]).

Rather than have the help file open in the same window, I had it open up in a new window (see Figure 6.23 [docuhelp.htm]), using the target="_blank" attribute within the anchor <a> tag. HTML manuals on the whole discourage use of this tag; extensive use of this attribute is considered bad HTML style (see Niederst, 1999: 137), since it potentially leaves users with a mess of open windows. However, user manuals also state that it can be done sparingly, and I decided that a help file was more useful if it could be examined alongside the website, rather than the user having to click backwards and forwards.

In addition, I anchored links to the help file within the help page itself, so that clicking on the help link would take the user directly to the appropriate part of the page. For example, Figure 6.16 shows the questionnaire documentation home page; clicking on the help link on that page takes the user directly to help on using the questionnaire documentation (see Figure 6.24). This is known as 'context-sensitive help'.
FRS documentation help

What is on this site?
Using the search facilities
Using the questionnaire documentation

What is on this site?

General documentation

Pages that gather together information which holds for each survey, year on year. This includes:

- Background:
  survey information, overview and response rates
- Structure:
  where the datasets are stored, full description of both the hierarchical and flatfile datasets
- Programming examples:
  for hierarchical and flatfile datasets
- Imputation:
  flags and programming examples
- Style guide:
  how to navigate the paper documentation using the styles function in Word
- Questionnaire:
  detailed description and instructions for using the questionnaire

Version specific documentation

Year-on-year documentation of the:

- Questionnaire:
  routing and question documentation of assets block, and household and benefit schedules
The documentation of the questionnaire is split into three separate pages, each covering the parallel blocks: household schedule, benefit schedule, and assets block.

Within each parallel block page, the documentation is organized according to blocks of questions which correspond to the way the interview program is divided up. Each block has a name, a shorthand version of the content of the block. These block names are listed on the left hand side of the page. Clicking on these names links you to the relevant part of the questionnaire.

Within each question block, the documentation is organized around question names.

The documentation text can be hidden or shown in two ways:

- Using global instructions to hide away all types of, e.g. condition text.
- Managing the text within individual question blocks or around individual question names.

Global instructions

At the top of each parallel block page, there are four global instructions given in blue text. A single click on the blue text performs a specific function on the whole documentation text:

- Show and hide conditions - controls whether or not condition text is shown.
- Show and hide question text - controls whether or not question text is shown.
- Show and hide answer type - controls whether or not answers are shown.
- Show and hide all - reduces documentation to a list of question blocks, or refreshes the page to show all text.

Managing individual blocks

The text of the questionnaire can also be reduced and expanded within individual blocks.
**Links between questionnaire and metadata**

I added further interconnectivity between pages on the site by putting in place links between the questionnaire documentation and the individual variable pages on the metadata documentation. Figure 6.25 shows the questionnaire documentation with the link to the variable *ANYMON* highlighted in red, showing an active link (compare this with the original version of the questionnaire, shown in Figure 6.19). The variable page opens in a new window (using the `target="_blank"` attribute).

**Figure 6.25 Questionnaire documentation showing link to variable ANYMON**

```plaintext
ASK IF: In loop FOR PersCont := 1 TO HHSize
    AND: AssEhg[PersCont] = 2
    AND: PA[PersCont][A][1].Asset = Yes
    AND: In loop FOR pcount := 1 TO 4
        AND: (pcount - 1) OR ((pcount > 1) AND (Account[pcount - 1].More = Yes))
        AND: pcount = 1

FRS35.Assets.QCurrAC[].Account[].AnyMon

Now I'd like to ask you about your current account(s):
At the end of last (month/pay period), did you have any money left in your current account, after your household expenditure?

INTERVIEWER: THIS INCLUDES ANY JOINT ACCOUNTS

Click here for associated variable

(1) Yes - money in (one or more) account(s)
(2) No - no money in any current account
(3) No longer have any current account(s)
```
ASD staff user feedback

I made a prototype of the site available to users on the ASD intranet across several weeks in June/July 2002. This prototype consisted of: a full version of the introductory documentation; a full version of the metadata documentation, including the search facilities; an initial version of the questionnaire documentation, covering the assets block; and a version of the help file.

I interviewed one user from the FRS team and conducted short interviews with three other members of ASD from outside ASD IA 1. I obtained written feedback via email from three other staff members and users of the FRS from outside ASD IA 1. The users from outside the FRS team were analysts developing the Policy Simulation Model (PSM), producing the Households Below Average Income (HBAI) report, and in pensions analysis. I have discussed already for what purposes FRS data is used by these analysts (see chapter 5; pp. 116-117); in these interviews I asked respondents which parts of the site and the documentation were most relevant to their particular needs. The FRS team member was most likely to use the whole site, particularly the questionnaire documentation59. Outside the FRS team, people responded that they were more likely to make use of the hierarchical documentation, particularly the variable metadata documentation. For example, analysts on the HBAI take variables from the FRS and adapt them for the purposes of their particular measures; documentation of the variable metadata is therefore useful to allow analysts on the HBAI to trace their own definitions against variables on the FRS60. The questionnaire documentation was seen as the type of detailed knowledge of the survey which the FRS team would have; this expertise could be drawn on by analysts external to the FRS team61. In the remainder of this section I summarize the feedback which relates to the final development work done on the documentation in order to bring the description of the development to a conclusion.

Concerning the content and the organization of the site, I asked respondents to comment on specific issues related to: the help file, i.e. whether it was sufficiently detailed, and whether opening the help in a separate window was helpful or

59 Interview with JS, 22/07/02.
60 Interview with LC, 22/07/02.
61 Interview with LC, 22/07/02.
confusing; whether opening separate windows for variable pages linked from table pages was helpful or confusing; whether navigation of the questionnaire documentation was intuitive or confusing; and whether links from the questionnaire documentation back to variable pages were useful, and whether or not these should open in separate windows. I summarize the suggestions for additions to content and amendments to the organization of the site, then describe the final changes made to the site on the basis of this feedback.

**Suggestions for additions to content**

One user suggested that the programming section would be more useful if it had very small, specific pieces of SAS code, which they could copy and paste\(^{62}\).

As discussed above, non-FRS team users found the hierarchical documentation the most significant and useful part of the documentation. The close relationship of the new documentation to the Excel metadata file was apparent, and one user expressed a marked preference for the new web-based layout\(^{63}\). Various suggestions were made by one user as to how the hierarchical documentation might be usefully expanded, e.g. providing information on which variables are fed into derived variables. Some of the omissions noted arose partly because of the fact that this was a single year's documentation, e.g. there was a suggestion for a 'history' of variables (giving information on, for example, how long a question has been asked for, has it ever taken other names, has the wording of the question differed - information needed for longitudinal analysis). This user suggested also that tracking 'problems' with variables would be helpful, e.g. giving information on low response, whether it is an income variable and subject to misreporting, degree of imputation, etc\(^{64}\).

Some expansion of explanations was suggested; several users noted that abbreviations were used at various points on the site which would be clear to FRS team members but not to members of the department outside of the FRS team\(^{65}\). One respondent suggested linking from technical words to a glossary, i.e. to be able to

---

\(^{62}\) Written communication from KH, 27/07/02.
\(^{63}\) Interview with LC, 22/07/02.
\(^{64}\) Summary of feedback obtained in written communication from KH, 27/07/02.
\(^{65}\) Written communications from SC and KH, 27/07/02.
click on a word and open up a new window explaining these abbreviations or technical terms.  

**Suggestions for amendments to site organization**

In general, however, the organization and structure of the site met with users' approval. One user indicated that she had been clear that placing the links around the site in a column on the left set off from the rest of the page signalled that these were for navigating the site. Another user commented that it was clear that the site had been ordered and layered broadly hierarchically, and gave this his explicit approval. Both of these points relate to *conventions* in the design of sites.

Answering my specific question about opening multiple windows, several users stated that they would be able to navigate and control the site if multiple windows opened, e.g. if each variable page linked from the table pages were to open as a new window. It was also noted by several users that this would be the most useful way to present the information, so that they could compare multiple variables simultaneously. Similarly, another user preferred a separate window opening for the help file; the help file met with his approval, particularly its context sensitivity, i.e. linking directly to the relevant section of the page. All of this feedback contrasts with the advice generally given in HTML design literature (see above, p. 188), that opening multiple windows can lead to confusion and should be avoided. Interviewed users stated that they were confident enough of their own web literacy for this not to be a concern. (See my discussion of choosing between the competing 'authorities' of users and manuals in chapter 7, pp. 229-234.)

This discrepancy between suggested practice in the manuals and user experience might result from a number of factors. Firstly, HTML manuals may well be giving their advice in the most generalized terms, and not necessarily for highly computer literate users. The users to whom I presented this documentation were computer literate: they are trained to use, daily, complex statistical analysis tools for

---

66 Written communication from EH, 27/07/02.
67 Interview with LC, 22/07/02.
68 Interview with JS, 22/07/02.
69 Interviews with PSM staff, 22/07/02; interview with JS, 22/07/02.
70 Interview with JS, 22/07/02.
the PC. Whilst these skills do not map directly onto Web literacy (a definition of which might include, for example, being able to find information quickly on the Web), there is an overlap with users being familiar with windows-based GUIs, which are the general interface for both Web browsers and statistical tools used in the department such as SPSS. The generalized information given in HTML manuals might not also hold well for presenting specialized information such as the hierarchically structured metadata, and might also conflict with existing expectations on how the information will be presented, inherited from previous documentation (such as the Excel data file). Users who preferred multiple windows open were stating a preference presumably based on prior experience of using the metadata in certain ways and formats, and perhaps using this to make a statement based on how they wanted to use it in future.

One user expressed confusion about the table pages listing the variables as to where the derived variables were to be found - the links at the top of the page to move the user about the page were not immediately clear as such.\textsuperscript{71} Since I knew from discussions with both FRS project managers that derived variables were of such significance to users, I needed to give this particular attention – I fixed it by the addition of a short explanatory note at the top of each table page, see Figure 6.26.

\textsuperscript{71} Interview with LC, 22/07/02.
As I had expected, given the questionnaire documentation's comparatively less developed state, this was the part of the site which attracted most suggestions for change. (This feedback was primarily received from the FRS team member.) Suggested changes to the questionnaire documentation related to:
• **Navigation.** Several respondents said that the instructions for using the show and hide functions were confusing, and their effects unclear. One user suggested that a layout similar to that used on other pages (i.e. with these navigation facilities on the left-hand side) might help, and that frames could be used to keep the instructions for showing and hiding text present on the page, when the documentation itself was being scrolled through.

• **Structure of the site.** The user from the FRS team found the links between the questionnaire documentation and the variable metadata documentation particularly useful, and he preferred that the individual variable pages opened in new windows.

---

**Acting on feedback from users**

I found that the suggestions for additions to content fell into two categories. Firstly, there was existing documentation that I could put online straightaway, e.g. examples of programming code. Secondly, there were documentation initiatives which were under development by the FRS team but for which documentation did not yet exist, e.g. a history of variables and problems related to variables. I passed on suggestions as to how this might be implemented on the site to the FRS team.

I implemented various amendments to the layout of the questionnaire documentation. Figure 6.27 shows the final appearance of the questionnaire documentation (compare to Figure 6.18). I used the `<div>` tag to divide the page into three sections: this allowed the navigational links on the left-hand side to remain in place rather than scrolling off the page when the documentation itself is being examined. I put the instructions for showing and hiding text into a separate section at the top of the page, to distinguish them from the navigational links, and I separated the two functions (show and hide) to make them clearer. In addition, I altered the background colour of the section in which the questionnaire documentation itself is presented, to distinguish it from the other sections, and to make it more readable.

---

72 Written communication from SC, 27/07/02; interviews with PSM staff, 22/07/02.
73 Interview with JS, 22/07/02.
74 Interview with JS, 22/07/02.
Figure 6.27 Final layout for questionnaire documentation
8 Summary

In this chapter, I have given a detailed description of the development work I carried out to build online documentation for the Family Resources Survey. I gave a detailed overview of the completed documentation, which covers general survey documentation, and version-specific (year-by-year) documentation of the questionnaire and the metadata. I gave a detailed account of the development work, which I described as involving five main tasks: determining existing documentation and outlining design principles for the site; designing and testing pilot documentation; designing search facilities; designing and implementing full documentation; and acquiring and acting on user feedback. Various kinds of work done on the site, however, e.g. finessing the layout of the variable metadata HTML pages, consisted of tasks that were ongoing for the duration of the project.

I organized this chapter around what, in retrospect, can be seen as two distinct phases in the development: the development of pilot documentation, and the development of full documentation. The two phases which I described fall on either side of a major point of change for the FRS as a whole, when the long-term project manager moved on and another project manager was appointed from within the FRS team. Prior to this, I had been concentrating on devising means of putting online various pieces of existing documentation, and had devoted most of my attention to an online, searchable version of the Excel metadata documentation. After the change in personnel, there were several changes to my own project. In our discussions, the new project manager, the IT staff member and I shifted the focus of the project to developing a user-driven prototype of the questionnaire documentation. I took the decision not to use existing and general standards, but to position the project in terms of it providing a solution more specific to the FRS. I also began to work increasingly off site.

In the following two chapters, I shall look more fully at some of the social processes of the development. In particular, I shall draw on an actor-network perspective, and I shall organize my discussion around three main themes: the 'authorship' of the site; 'influences' upon the text; and, my own role at various points as 'translator' of the site into both bureaucratic and academic terms.
7 Technology as text: narrative and authorship

1 Introduction

In this chapter, I shall examine in greater depth some of the social processes involved in the writing of the documentation for the FRS website. Before outlining in more detail what my analysis entails, I want to reiterate that this is intended to be a hermeneutic analysis: that is, it is concerned with interpreting the technology as a text (in both its production and reception), and also with the rhetorical uses of technology throughout the production of that text. In this chapter, then, I shall be looking at the influences upon the development of the website (i.e. I shall be giving an account of the ways in which the development of the online documentation can be described as being socially shaped). In particular, I want to examine occasions where specific decisions were made that influenced the form and content of the documentation (for example, my choice to use a specific HTML tag which allowed multiple windows to open), and the justifications which were behind these choices (in this example, why I chose to follow the preference for multiple windows expressed in feedback I acquired from users over the advice given in HTML manuals to avoid opening multiple windows).

Since I am concerned here with the social shaping of the documentation, I should refer again to work done by Edge (1995; see chapter 3, p. 58). In his discussion of the social factors which can shape the innovation process, he draws attention to the flexibility of the innovation process and the choices made during this process. More broadly, he describes how technological development is socially embedded: social factors may influence selection between available technological possibilities; they may permit only one area of 'possible' technological development to be explored to the extent that it becomes difficult to talk of 'alternatives'; they may operate by creating a particular environment (e.g. market) or intellectual climate where only certain technological configurations succeed; they may shape technological development by the specific embodiment of social models into the technology. Several factors emerged as particularly significant in the development - the writing - of the FRS documentation website; for example, the extent to which the project - and my own activity throughout - was embedded in wider (and very stable) networks, both organizational and technical.
The narrative of this chapter largely follows the progress of the project itself and, as the chapter goes on, I want to show how my own perception of the 'audience' for the work I was carrying out altered perceptibly. While building the pilot site, I drew heavily on theoretical work conducted by Hunter (1999), in order to test the applicability of the various taxonomies of hypertext which appeared in the literature. As I discuss in section 2, many of the design decisions I made at this stage were driven by my sense of having an academic audience, and of building the site in response to the various claims about hypertext made in the literature. However, as I described in chapter 6 (p. 171) and discuss in section 3 of this chapter, as the project went on my design decisions became increasingly driven by my perceptions of 'user expectations'; expectations which I derived both from the experience of the FRS team manager, and my own interactions with the team members who would be using the documentation site. I reflect upon issues related to my perceptions of the 'audience' for the work at various points throughout the course of this chapter, and particularly in section 2.

Whilst this is a hermeneutic analysis, nevertheless, throughout the chapter I want to draw on the tools of another theoretical perspective: actor-network theory (ANT). For example, one theme which I shall foreground throughout is that of the writing of 'narrative', drawing upon Law's (1999) presentation of ANT as a form of 'storytelling'. My focus on narrative derives, also, from the Gadamerian emphasis on the 'giving of accounts', which I examined in my discussion of hermeneutic methodology in chapter 2 (pp. 31-35). It is consistent with my focus in this chapter on the production of a bureaucratic text. I will examine several kinds of 'narrative' in the course of this chapter. In section 2, I shall examine the impact - or, the influence - of academic accounts of hypertext. I shall consider the 'narratives' which I myself was constructing throughout the production of the site in order to justify certain development decisions; for example, as I describe in chapter 8, my decision over whether or not to use standards for online documentation. My discussion of these narratives is closely related to the points I make throughout about the 'audience' which the work was addressing; i.e. at whom these justifications were aimed.

I draw upon the theoretical tools of ANT in particular to use two concepts. Firstly, an actor-network perspective allows me to consider 'actants' on the process of writing the documentation. The idea that all actants (whether natural, human, or technological) should be treated as having potentially equal significance, allows me to
give the influential academic narrative accounts (of hypertext and the Web) the status of actants (in hermeneutic terms, I shall be examining influences upon the text). I outline some of these narrative actants, and indicate ways in which they influenced, or shaped, the documentation. In addition and, again, drawing together the actor-network perspective on actants and the hermeneutic interest in the text, I discuss a variety of issues related to the authorship of the text. In particular, I explore what 'authorship' means within a bureaucratic setting such as the DWP, in which objectivity and distance are significant elements of the culture. I draw upon examples from the development of the FRS documentation to discuss a variety of issues related to the production - the writing - of a bureaucratic text. The second specific concept which I intend to draw upon from actor-network theory is that of 'translation'; I shall return to this in chapter 8.

In drawing upon actor-network theory to inform my analysis, I am again mapping issues related to the shaping and construction of technology onto the production and reception of texts. I am also bringing my own analytical orientation into dialogue with another: this is consistent with Law's (1999) reflections on actor-network theory:

"Is there such a thing as "actor-network theory" at all? Answer: yes. We can certainly make a story that tells of unity. But the answer is also no, for it is just as easy to tell tales of a kind of diaspora, of interaction with other "theories", of confusion, or if you prefer, of complexity, overlap and partial connections" (p. 4; his emphasis).

This point is consistent with my discussion in chapter 3 (pp. 68-69) of the debate between proponents of ANT and sociologists of science such as Collins and Yearley, of the 'family resemblances' between the approaches, and of the significance being in the debate itself and its opportunities for generating knowledge. Using theoretical concepts from ANT shows up what I think is a flaw in textually-orientated approaches, which is that the emphasis on the multiple possible readings of a text can lead to an author herself becoming distanced from a text. I think that this is evidenced, in part, by the emphasis on resistant readings and textual appropriation in much contemporary writing on text and media reception (which I discuss in chapter 3, pp. 61-62). It might also serve as a partial explanation for my own absence from my
original drafts of chapter 6 of this thesis - the description of the project - which I wrote originally in the tone of a standard scientific report. (I address other possible explanations for this throughout, particularly examining the reasons for the 'voice' that I adopted during the writing of the documentation website and which was compelling enough to survive into the writing of this thesis.)

In this section, I have introduced the themes I shall be discussing and the theoretical perspectives I shall be drawing upon in the rest of the chapter. The rest of the chapter is organized in this way: in section 2, I will look at textual influences upon the writing of the text or, in terms of ANT, at narratives as actants; in particular, academic narratives of hypertext. I shall discuss places where I made rhetorical use of technology, and consider the 'audience' to which these justifications were addressed. In section 3, I am concerned with the authorship of the text. I show how the writing of the documentation website was both anonymous and collaborative, and I also draw upon Latour's (1990) idea of 'cascades of inscription' to show that there were strong organizational and technical actants limiting the scope for authorship on the documentation website.

2 Textual influences, and narratives as actants

I wrote a variety of different texts throughout the course of the project. As well as the FRS documentation website itself - i.e. the HTML and JavaScript that I wrote - I also produced a number of other texts. For example, producing the CD-ROM included with this thesis required me to adjust small parts of the text so that links to other parts of the DWP website were not broken. Other texts that I wrote included an executive summary of the feedback which I received from my interviews with users; the presentation I gave to the FRS Users Group; also, the summaries of my work progress that I sent out via email to the FRS manager in the latter stages of the project, when I was working off site, and which formed the basis for our discussions in our meetings. My primary focus in this section, however, is the writing of the FRS documentation website itself.

While I was writing the website, I drew on the following texts: the FRS Guide; the BAD; the TADEQ tool; hypertext manuals; standards documents; Card's (2000) article (which supplied the format by which each variable was represented as a single
web page; see chapter 6, p. 152 above); academic accounts of hypertext such as those provided by Mitra and Cohen (1999) and Hunter (1999); and, tacit departmental knowledge which I acquired through formal and informal discussions whilst I was on site at the DWP. However, in this chapter, I want to consider narratives not just as influences, but also as actants. This gives, I believe, a broader, social perspective on the discrete influences on the FRS documentation that I described above.

Characterizing narratives as actants evokes the wider networks in which these narratives are produced.

For example, one of the earliest, seemingly most straightforward, decisions made about the FRS documentation was that it should be put onto the DSS intranet.\(^75\)

This choice bounded the technical scope of the project in certain ways; more specifically, it directed me immediately towards using HTML when I began work on the documentation. It was a decision which, I would argue, arose from the confluence of the organizational and technological trends that I described in chapter 4; i.e. from the increased computerization of the survey research process (which created a need for new forms of documentation), and the increased use of the Internet by UK government in the dissemination of survey research material. This shift to using the Internet for both survey results and documentation in turn arose from accounts of the 'information society' which I described in chapter 3 (pp. 42-56); more particularly, from those political initiatives such as e-Strategy (see my discussion in chapter 4, pp. 94-99).

Callon (1987), in his study of the VEL project (an attempt in the 1970s by French industry and government to produce an electric car and the associated and required infrastructure), refers to the engineers involved in the project as engineer-sociologists: i.e. that in the design and development of the new technology, they were also drawn into making sociological analyses. Callon writes:

'[The engineers] went from electrochemistry to political science without transition. The analysis that they proposed was both remarkably incisive and fully elaborated. Five years after the "great cultural revolution" of May 1968 and one year before the first oil crisis,

---

\(^75\) I am using the passive voice intentionally here to emphasize that this was a decision that primarily resulted from the network of influences I am about to describe, rather than from an individual and isolated decision.
they outlined the course of an evolutionary movement that would propel French society from the industrial to the post-industrial age. This change was to occur through pressure from new social movements and with the expected help of electrons' (p. 86).

Thus, the engineers on the VEL project, Callon argues, were not just architects of technological advances but also proponents of social change: the technological innovations they were proposing went hand-in-hand with (what they perceived as) social improvements.

I do not want to overstate this in relation to the FRS documentation website. However, where the engineer-sociologists of the VEL project were influenced by post-industrialist theories, the work conducted by myself and other FRS team members on the site (and, by extension, within the wider actor-network of government social survey research) is located within the discourse of information society narratives. This led to a specific and fundamental decision about the form of the documentation being made at a very early stage; i.e. that it should be available on the department intranet (with a view, eventually, to putting it on the Web). This is an example of what Edge (1995) is describing when he questions, in general, whether innovation processes are flexible, and whether social factors permit one area of possible technological development to be explored, rather than alternatives.76

One other point which Callon’s study raises relates to the heterogeneity of the engineers which he studied; i.e. the diversity of roles that they perform. A similar point could be made for the FRS team: that they are simultaneously professional statisticians (see chapter 4, p. 91) working within the context of a scientific discourse, as well as civil servants mediating between political initiatives (such as e-Strategy) and the objectivity demanded in a bureaucratic setting. The same point about heterogeneity applies to my own experiences on the project: as someone funded by the DSS/DWP and treated as a part-time member of staff; and, at the same time, a student attempting to formulate academic questions and produce a thesis. Whilst the

76 Taking into account the technological and organizational infrastructure that supports the FRS (i.e. taking into account the extent of the immutability of the networks in which it is embedded), it is hard to see what alternatives there were other than to put the FRS documentation online. Unlike other surveys managed by the department, which shifted from being paper-based to CAPI-based, the FRS had been conceived as a CAPI-based survey from the outset; a paper-based version of questionnaire did not exist, other than the BAD, which was generated automatically from the BLAISE code.
primary purpose of the project from the point of view of the collaborative sponsors (the DSS/DWP) was to develop online documentation for the FRS, I simultaneously had another, academic purpose. The heterogeneity of my own roles meant that the academic literature I was reviewing played a specific role in the development of the site. This was particularly the case in the first stage of the project (the pilot stage), when academic accounts of hypertext were very significant in the design choices I was making. In the remainder of this section, I shall describe in more detail the effect of the academic literature on the documentation website. This discussion will, I hope, give a sense of how, particularly during the pilot stage, I tried to frame the development work in such a way that it also allowed me to form and pose academic questions: how the project was simultaneously an academic project, as well as a project aimed at providing specific outcomes for the DSS.

I described the pilot project in detail in chapter 6 (see p. 135); to summarize the main points, I determined the content of the existing documentation and outlined design principles for the site; I then put online the paper-based documentation and the metadata documentation, and I developed search facilities for the metadata documentation. Throughout the pilot stage, I worked closely with the FRS project manager, consulted internal documents, and used HTML manuals. I drew upon the academic literature on hypertext, and I want now to show how this literature guided aspects of the development - how it influenced the development - and how my academic questions were, in turn, altered as the work on the site progressed. (I should note here that this is an example from a specific stage in the project, at the pilot stage of the development). At this point, I was using the academic literature on hypertext in an attempt to formulate research questions for the project: academic questions surrounding the 'nature' of hypertext, using models from one source to challenge the assumptions I saw in the wider literature. Ultimately, however, I was unable to sustain this argument; or, rather, the kinds of questions that became more relevant were less to do with resolving debates over technological and social determinism and more about discourses surrounding technology.

One consequence of my focus at this time on models of hypertext was that my design decisions were often motivated by my sense that there would be an academic audience to whom I was attempting to justify these decisions, and I will discuss this at various points in the example that follows. Eventually, not only the academic focus altered, but also the sense of the audience of the site changed; I became increasingly
conscious of the users of the site. I shall discuss this further below (section 3). In the remainder of this section, I want to show how my own 'story' of the project at the pilot stage emerged in response to the narratives of hypertext in the literature, and how I had a sense of an academic audience to whom I was justifying my design decisions.

'Classics' from the literature on hypertext such as Bolter (1991) and Landow (1992) characterize the technology in terms of it being the next rung on a textual evolutionary ladder (see my discussion in chapter 3, pp. 73-79). Mitra and Cohen (1999) outlined a classification of the distinguishing features of hypertext with particular reference to the Web (see chapter 3, pp. 78-79), highlighting such features as intertextuality, non-linearity, how the reader of hypertext is active, how hypertext is ephemeral (since links can be broken and pages deleted). I was dissatisfied with these accounts of hypertext for a number of reasons: firstly, from a theoretical dislike of teleological accounts of technology; secondly, because of literature from other fields such as media and cultural studies which had, since the 1980s, been emphasizing the active nature of book reading (see, for example, Radway, 1984); thirdly, because I had encountered online texts such as electronic journals which shared more features with their paper-based counterparts than with the descriptions of hypertext given by Mitra and Cohen (see chapter 3, pp. 80-81); and, fourthly, because I was conscious of the linear structure of the FRS database as evidenced by the name of the hierarchical dataset.

Nevertheless, my preliminary work on the FRS documentation site involved putting online the paper-based documentation, and I was able to find examples of differences between the linear presentation of the FRS Guide (see chapter 6, pp. 146-149) and my online versions of the same material. For example, Figure 7.1 [frstable.htm] shows my presentation of the structure of the hierarchical dataset in electronic format. In this case, the information in the paper-based Guide took four pages across which the user was obliged to move backwards and forwards. On the web page, I set a table at the top containing each of the 24 table names, and made each table name a hyperlink to the section further down the page which provided information on that particular table.

Another example of non-linear representation of text can be seen on the pages demonstrating programming examples. In the Guide, these examples stretch across five pages, and users must move backwards and forwards through the text. When I put these examples online (see Figure 7.2 [prog_eg_1.htm]), I was able to present them so
that a user could examine the different examples of programming simultaneously, using frames and scrolling within the browser to align parts of the programs.

**Figure 7.1 Non-linear presentation of the Guide - Table descriptions**
Having created examples on the FRS website which supported Mitra and Cohen's description, I was naturally keen to be able to provide opposing examples. At this stage in the development, then, my chief academic concern became finding a 'better fit' description of the site than that offered by the accounts of hypertext in Bolter, Landow, and Mitra and Cohen. I turned to a classification of hypertext applications provided by Hunter (1999; discussed in detail in chapter 3, pp. 81-82), and this became very influential in both the development work on the site, and the analytical work which I was simultaneously conducting. To recap, in her attempt to counter the notion that hypertext is implicitly flexible, relative, and non-hierarchical,
Hunter developed a classification of hypertext applications, using case studies of projects as examples. She suggests that these projects can be grouped under four approaches: topic-driven hypermedia texts, central text hypermedia, multi-document hypermedia, and hypermedia nests (p. 113):

- **Topic-driven hypermedia texts** are best seen as providing a multi-dimensional filing system for existing information which is already highly categorized and hierarchical.

- **Central text hypermedia** provides an informational shell around a central text, a text which may be a person/writer or a literary artefact.

- **Multi-document hypermedia** presents a multiple document archive, using hypertext predominantly as a presentation device.

- **Hypermedia nests** contain a small number of closely related, multiply linked texts.

At this point, then, I began consciously to conceptualize the website according to this classification, in order to explore the extent to which Hunter's classification worked in practice; in effect, I began to tell a story about my own work which was responding to the 'classic' literature on hypertext by using Hunter's taxonomy to guide my design choices. Moreover, I had a strong sense of the academic audience to which this work was addressed, and my 'story' of the project, or my justifications for the design choices I was making, was addressed to that hypothetical audience; I will indicate this at various points in the description of the development that follows.

Having put the *FRS Guide* online, I reconceptualized the section of the website presenting introductory and background information on the FRS in terms of Hunter's account of *multi-document hypermedia*. She describes this as 'a multiple document archive, using hypertext predominantly as a presentation device' (p. 117). Such hypermedia are primarily used for learning purposes. For my putative academic audience, then, I was able to cast my online presentation of the *Guide* as a collection of documents or chapters on themes related to the FRS; a multiple information source that provided an introduction to the FRS.

The work that I conducted next on the site was to put the metadata documentation online. Several sources were influential in the design decisions that I
made here. As described above (see chapter 6, p. 152), I drew on a similar project by Card (2000), in which a single web page was dedicated to each variable, with links provided from the table with which the variable was associated. In addition to using this prior example of a similar application, I drew support for my design decisions from the principle I had outlined of not confounding user expectations (see chapter 6, p. 145). I had taken this from literature on developing online documentation: Girill and Luk (1992) argue (p. 573) that freeing information from its 'traditional constraints' leads to its own problems; that users can become disorientated, and even 'lost in hyperspace':

'if the underlying information has a stable, well-understood, and well-known structure, allow the network to reflect that structure...
We believe that documentation systems designers can preserve the benefits of hierarchical access yet avoid its known weaknesses not by forsaking the use of text structure (as pure hypertext does) but rather by exploiting it in a different way' (pp. 574-6).

Again, note how I was making a design choice with attention to the support I could draw for it from academic and practical literature on hypertext; I was primarily concerned with addressing the academic questions that I was formulating at this point.

Most significantly at this point, I was guided by Hunter's description of topic-driven hypermedia texts. Hunter describes this approach as, in effect, providing 'a multi-dimensional filing system' for existing information which is already highly categorized and hierarchical. The information presented is rigorously categorized and largely pre-determined, and the user has limited or even no impact on the content. As Hunter summarizes:

'Topic-driven hypertexts are less problematic to construct than many other applications because by definition they count on a stable approach to material that does not unduly disrupt the expected order. In many ways they are more flexible and speedy enumerative bibliographic systems... Such hypertexts are set up... for educated specialist users, to better present the formal, corporately held directions

211
of information and make more effective the sense of a necessary
answer or conclusion' (p. 114).

These applications are therefore primarily aimed at presenting organized
information for the use of specialists who want their needs to be anticipated. This
information is intended for specialist users with specific expectations as to how it will
be presented. As Hunter summarizes about her own project:

'[T]he mere pre-existence of the information to be put into the
hypertext meant that substantial categorising and hierachising had
already taken place, which could not be disrupted without unhelpfully
disordering the expectations of users' (p. 113; author's emphasis).

All of these sources enabled me to justify my organization of the metadata
documentation in a hierarchical fashion; i.e. I had used academic and practical
accounts of hypertext to provide support for a design decision which, in turn, I could
use as a counter-example against claims in the literature that hypertext was
necessarily non-linear and associative. I therefore conceptualized the variable
metadata in just the way Hunter described: as information already highly formalized
and hierarchically organized, as the name of the dataset implies. Figure 7.3 shows the
structure I devised for this part of the pilot website, showing the path for reaching one
variable (ADCH) off the ACCOUNTS table. (Note: solid lines on this diagram
indicate single pages and links; lines and dots indicate that the box or line represents
multiple pages or links. See Figure 6.9 for a full map of the pilot website, showing in
detail its hierarchical structure.)
In the choices that I made in the design of the documentation for the variable metadata, then, I drew on a number of resources to support my design decisions. I used a prior example of a similar project (Card, 2000) and I drew upon literature on site design (Girill and Luk, 1992). However, my role was not just as designer of the site - I was also attempting to formulate academic questions to test the accuracy of the descriptions of hypertext that appeared in the literature, and I was conscious of a potential academic audience to whom these questions and decisions would need to be justified. In designing the metadata documentation, for example, I was looking for examples from the development to differentiate this site from the literature that imbues hypertext applications with specific characteristics. In addition to the other sources, I used Hunter's classification to provide grounds for the design decisions that led to me structuring the metadata documentation hierarchically. At this, the pilot stage, I was using the site development to explore questions related to the nature of hypertext that had arisen from the debate which I perceived in the literature. The structure of the site, at this stage, was greatly influenced by my attempts to answer these questions; and my sense of having to respond to the literature in the field (in Gadamerian terms, of operating within a tradition), led me to tell the 'story' of the project to myself as one of justifying design decisions to a potential academic audience. My design decisions can be characterized as rhetorical moves made to persuade a perceived audience.
After I had completed the pilot documentation, and during the hiatus in the development (see chapter 6, p. 170), I continued to conceptualize the remaining work I had to carry out on the site in terms of Hunter's classification of hypertext systems. I intended to design the online questionnaire documentation in terms of 'central text hypermedia' which, as Hunter describes them, provide 'an information shell surrounding the central text, a text which may be a person/writer or a literary artefact' (1999: 115). In this case the central text would be the FRS questionnaire. Ultimately, I abandoned this approach to the design of the online questionnaire documentation. In part, this was because I had become dissatisfied with the importance in Hunter's classification attached to application design; significantly, however, my academic questions had begun to shift away from proving or disproving that particular taxonomies provided a 'best fit'. This coincided with the second phase of the development, in which I shifted the emphasis in the project away from design principles towards getting user feedback (see chapter 6, pp. 170, 183-198, and discussed in section 3, below). The 'audience' for the site - the people to whom my 'story' of the site was being addressed - altered perceptibly in the second phase of the project, from an academic audience to the people who would be using the site.

In this section, I have looked at influences upon the text; more particularly, I characterized narratives as actants, and I looked at the effect of discourses of the information society and academic accounts of hypertext on the development of the FRS documentation website. The very early decision to put the documentation online was part of a broader trend in the online dissemination of survey material as part of government e-Strategy; I used academic accounts of hypertext to guide my design decisions during the pilot stage of the project. In this section, I also drew out the heterogeneity of my own role as both designer of the site and a student attempting to ask academic questions (I shall discuss this further in chapter 8). The other theme which I raised in this section was of the audience for the site: at the pilot stage, this was a potential academic audience, and I made rhetorical moves to justify my design decisions to this academic audience.

Having looked at some of the influences on the text, I now want to examine the process of writing the text, taking into account the wider bureaucratic setting in which the project was carried out.
3 Authorship in a bureaucratic context

Introduction

Drawing together the actor-network perspective on actants and the hermeneutic interest in the text, I now want to discuss a variety of issues related to the authorship of the text of the FRS documentation website. I am intentionally discussing the 'authors' rather than the 'writers' of the site, because the nuances of the word are concerned with the author as an acknowledged source - an authority - and also imply a distinctive, individual tone. These themes underlie the rest of this section, in which I draw out the contrasting process of authoring texts in a bureaucratic context.

Authorship of the FRS documentation website was a collaborative and anonymous process, and there were several strong influences (or networks) that drew the content and structure of the site in particular directions. (As an aside, the phrase 'authoring language' refers to a language, such as HTML, that enables the development of documents and computer applications without the need for detailed low-level programming.)

I begin by looking at how an objectivity of tone was achieved in the writing of the FRS documentation (I am intentionally using the passive voice here). In part, this was through making anonymous my own authorship of the site; in part, it was achieved through collaborating with other department members to author parts of the site. I shall give examples of anonymous text on the site, and examples of situations during the site's development where the authorship of the website was shared: i.e. where other department members had a significant impact on the site's form and content. At various points throughout the chapter, too, I attempt to draw out some of the more immutable networks that lay behind and informed the authoring of the site. This is particularly relevant in my discussion of the development of the classification system for the search facilities, where I draw on Latour's (1990) notion of 'cascades of inscription' to show how an already existing system of classification was re-inscribed upon the site. I also examine a point where two 'authorities' on how to present online information came into conflict (these authorities were HTML manuals, and the tacit knowledge which users of the site had expressed to me), and I had to choose which one to use to guide my decision over presentation. This section is therefore concerned with the writing of a bureaucratic text, and the limits on authorship in this context.
Anonymous authorship

My authorship of the site is almost entirely anonymous: my name appears once on the FRS documentation website - on the website front page [docufont.htm]. The two 'signatures' that appear on almost every page are institutional: the DWP logo at the top of the page, and a 'Crown copyright' graphic that appears at the bottom of most pages; see Figure 7.4 (the questionnaire documentation front page)\textsuperscript{77}. I derived much of the text for the site from pre-existing, internal documentation. Firstly, I took a substantial part of the text of the site from automatically generated texts\textsuperscript{78}. The questionnaire documentation which I put online was taken from the BAD, i.e. the BLAISE Automatic Documentation. The BAD is generated directly by staff at ONS from the CAPI program BLAISE, in which the FRS is written, and which was developed by Statistics Netherland. This means that a large amount of the content that was put onto the FRS documentation website was the result of substantial prior work, conducted long before this project was begun - it is the product of an existing network of both organizational and technical actants, such as the CAPI program itself. The Excel file which provided the content for the variable metadata pages is also generated automatically from the BLAISE questionnaire by its programmers at the ONS, and so is the product of these external networks. This file was a very significant part of the project: it provided the layout (the form) of the variable metadata pages which I devised and reworked, in conjunction with the FRS project manager and the IT staff member (see my discussion of 'cascades of inscription', below). It is consistent, I think, with the actor-network principle of treating human and non-human actants as equal to consider these automated texts 'authors' of the site.

\textsuperscript{77} Examples of institutional signatures can be found in other bureaucratic environments, and not necessarily those with a scientific ethos. For example, at the BBC during the 1960s and 1970s, the in-house production of drama scripts, combined with Writers' Guild (i.e. trade union) regulations protecting freelance writers, could lead to a situation where a made-up name was given on-screen credit for a script. This might happen, for example, when rewrites were so extensive that the writer asked for his or her name to be taken off the script, but where the in-house staff member who had rewritten it could not take a credit under Writers' Guild rules. This combination of union rules and bureaucratic procedures at the BBC at this time resulted, under certain circumstances, in an institutionalized anonymity of authorship.

\textsuperscript{78} I also designed the interface of the online version of the questionnaire in such a way that the IT department at the DWP would be able to generate future 'editions' of it automatically from the HTML I wrote.
Figure 7.4 Institutional signatures

The second source of internal documentation which I used to provide text for the site was paper-based documentation such as the Guide. This had no author credit, and the tone in which it was written was distanced, and made extensive use of the passive voice: see the text in Figure 7.5.
Imputation flags

Imputation is the process whereby missing values, for chosen variables, are edited to valid values. The main objective of imputation is to maximize the information available to users for analysis. Furthermore, the imputation carried out simplifies the analysis for users and helps to secure the uniformity of analysis created from the FRS datasets.

The editing process is carried out on cases where a value has been given but following checking is believed to be incorrect, for example, where a respondent has given a benefit amount which is not valid.

When imputation and editing flags are undertaken, transactions are applied to a copy of the dataset to ensure that the original dataset can still be accessed in its unaltered state. Transactions change the old value in the copied dataset to a new value, but they do not make any changes to the original dataset. Transactions allow the user to follow the audit trail and see how values have been changed. The transactions are saved in a transaction dataset, which records the table, variable, serial number and related key variables, the old value and the new value, of the record which has been changed.

Much of the writing which I carried out for the site was HTML code and JavaScript, none of which is visible when it is read by a web browser. I did write some text which was visible on the site, primarily instructions for use of the site: e.g. how to use the styles in the BAD. The most substantial piece of writing which I did that is visible on the site is the help file and, in writing this, I adopted the tone of the Guide, see Figure 7.6.
Using the questionnaire documentation

Description

The documentation of the questionnaire is split into three separate pages, each covering the parallel blocks: household schedule, benefit schedule, and assets block.

Within each parallel block page, the documentation is organized according to blocks of questions which correspond to the way the interview program is divided up. Each block has a name, a shorthand version of the content of the block. These block names are listed on the left hand side of the page. Clicking on these names links you to the relevant part of the questionnaire.

Within each question block, the documentation is organized around question names.

The documentation text can be hidden or shown in two ways:

- Using global instructions to hide away all types of, e.g. condition text.
- Managing the text within individual question blocks or around individual question names.

Global instructions

At the top of each parallel block page, there are four global instructions given in blue text. A single click on the blue text performs a specific function on the whole documentation text.

- Show and hide conditions - controls whether or not condition text is shown.
- Show and hide question text - controls whether or not question text is shown.
- Show and hide answer type - controls whether or not answers are shown.
- Show and hide all - reduces documentation to a list of question blocks, or refreshes the page to show all text.

Managing individual blocks

The text of the questionnaire can also be reduced and expanded within individual blocks.

A variety of factors, therefore, contributed to the anonymous 'voice' of the text of the site. Firstly, a large proportion of the text that appeared on the site was automatically generated (from the BAD or the Excel metadata file). I adapted the text for a number of pages from internal documentation such as the Guide, and where I did
write text, I modelled it on the style of this existing documentation, adopting its objective and distanced tone.

Why did I choose to write the text in this way rather than adopting a different style? The tone was certainly compelling; in fact, I continued to use it well into writing this thesis: the original version of chapter 6 read like a standard scientific report, using the passive voice throughout. In addition, I had a strong sense of the people who would be reading and using the documentation, i.e. I had a strong awareness of its intended audience. The people with whom I was working were fluent in a very specialized and technical language (not just as professional statisticians but, more specifically, with the detail of the FRS). Adopting the 'voice' of the language which they themselves used was, perhaps, more likely to give the texts that I was writing credibility; adopting this tone provided a degree of legitimacy for the work that I was doing; I 'learnt the language' in order to participate successfully in the community.

Having described how I took text directly from internal sources and presented it on the site, I now want to look at more specific examples in which there was collaboration between myself and other staff members in producing the content and form of the documentation website.

**Collaborative authorship**

If my own authorship of the site was anonymized, so too were the contributions of other members of ASD, whose expertise and knowledge I drew upon at various stages throughout the development. I have already given an example of text on the site which was provided by another, anonymous author (the text from the Guide which I put online); I now want to look at two examples from the development of the website where other members of the department became involved, who they were, and what changes I made to the documentation as a result.

My first example of collaborative authorship is concerned with the production of the search facilities for the variable metadata. In my discussion, I will draw on Latour's (1990) idea of 'inscriptions' as a means of understanding how data is represented and reproduced. In my second example of collaborative authorship, I describe a point in the development where I faced a design choice which had the
potential to alter the form of the site significantly. The HTML manuals I was using suggested one course of action; the users with whom I was speaking suggested another direction: this was therefore an example of choosing between two different authorities on the best form for the site to take.

*Collaborative authorship 1: writing search facilities*

Latour and others have stressed the importance of 'inscriptions' (e.g. types of data representation such as graphs, tables, diagrams) in the construction of scientific facts and as a means by which scientific knowledge-claims are constructed, revised, and contested. These kinds of inscription are a powerful tool in stabilizing facts, and allowing them to be reproduced and transported. In the case of the FRS, the hierarchical tables, variables, and metadata can be viewed as inscriptions to organize the data. In this section, I want to look first at how a set of inscriptions was re-inscribed onto the documentation website and how this was related to a particularly immutable network. I will then show how a new classification system was devised to organize data on the site, and how it was connected to this other, earlier classification system, i.e. how a 'cascade of inscriptions' was created.

The Excel variable metadata file had a classification system for variables in the FRS. Variables were classified in four ways: key variables (variables used to locate individual responses - the same across all tables); data variables (which hold data information); derived variables (variables which are created in the process of using the questionnaire, calculated from one or more other variables); and system variables (which store system-related information). When I first designed the format for the pages for each table on the site, I did not organize the list of variables according to the classification system. Figure 7.7 shows the original layout for ACCOUNTS table web page in the pilot documentation. In discussions with the FRS project manager, I was asked to include the classification system as part of the format of the page, because users were used to it, and already classified the variables that way. Figure 7.8 shows the revised format for the table web pages, with the classification in place. (This was particularly significant for those table pages with a much longer list of variables than the ACCOUNTS table.)
The classification system was therefore re-inscribed from the Excel file onto the documentation website and, given that I used the classifications of key, data, derived, and system to form links around the table pages, it became a navigational feature of the site. The site was organized to encourage users to 'read' it in a certain way, i.e. according to this classification system. As I noted above, the Excel file was generated automatically by staff at ONS from the BLAISE program in which the FRS questionnaire is written. This was a particularly immutable network of organizational and technical actors (e.g. the CAPI program) which affected the organization of the documentation website in a significant way.

Figure 7.7 ACCOUNTS table web page for the pilot documentation
Latour (1990) uses the idea of 'cascades of inscription' to describe the process whereby inscriptions are translated into other inscriptions. I now want to describe the process of developing the search facilities for the variable metadata in these terms, to show how an already existing system of classification was transformed into a new system of classification. The FRS project manager, working from her own knowledge
of the variable metadata and how it was used by the team and her predictions of how they would want to use such a search facility, constructed this classification system, which I then used as the basis for coding the search facilities for the variable metadata. I discussed these search facilities in detail in chapter 6 (pp. 162-169); to summarize: the FRS project manager's classification system organized the variables according to twenty first-level topic classifications, some of which had subcategories. (Table 6.4 gives the full topic classification list, showing both first- and second-level classifications.) Based on this classification system, I devised two ways of searching the metadata by topic: I coded a search engine into which users could enter topics to search for the relevant variables; and I coded a search tool using drop-down boxes to search according to the first-then second-level classifications.

I have described the coding that I did to develop these tools in detail in chapter 6; what I want to show here is how closely I relied upon the classification system which was devised by the FRS project manager. I shall focus on the first-level classification which she devised: 'Assets and savings'. Table 7.1 shows this first-level classification, and the second-level classifications associated with it. (Compare this with the full list of classifications given in Table 6.4 above.)

Table 7.1 Second-level classifications of first-level classification Assets and savings

<table>
<thead>
<tr>
<th>First-level classification</th>
<th>Second-level classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assets and savings</td>
<td>Accounts and investments</td>
</tr>
<tr>
<td></td>
<td>held</td>
</tr>
<tr>
<td></td>
<td>Capital value</td>
</tr>
<tr>
<td></td>
<td>Interest and dividends</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
</tbody>
</table>

The FRS project manager provided me with an Excel file which classified every variable on the 1998-1999 FRS according to the first-levels and (where applicable) second-levels she had devised. I sorted the Excel file according to each first-level classification, and worked from these documents (twenty in all - one for each first-level classification), to develop the search facilities. Table 7.2 shows the variables classified under 'Assets and savings'.
Table 7.2 Variables classified under Assets and savings

<table>
<thead>
<tr>
<th>Table</th>
<th>Variable</th>
<th>Label</th>
<th>First level</th>
<th>Second level</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCOUNTS</td>
<td>ACCOUNTS</td>
<td>Account Type</td>
<td>Assets and savings</td>
<td>Interest and dividends</td>
</tr>
<tr>
<td>ACCOUNTS</td>
<td>ACCTAX</td>
<td>Is that interest before or after tax</td>
<td>Assets and savings</td>
<td>Interest and dividends</td>
</tr>
<tr>
<td>ACCOUNTS</td>
<td>ADCH</td>
<td>Whether Adult or Child Account</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>ACCOUNTS</td>
<td>NSAMT</td>
<td>Value of National Savings Investment</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>ADULT</td>
<td>ACCOUNTS</td>
<td>Whether any accounts</td>
<td>Assets and savings</td>
<td>Accounts and investments</td>
</tr>
<tr>
<td>ADULT</td>
<td>ANYMON</td>
<td>Any money left in current account</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>ADULT</td>
<td>INVESTS</td>
<td>Whether any investments</td>
<td>Assets and savings</td>
<td>Accounts and investments</td>
</tr>
<tr>
<td>ADULT</td>
<td>OTINVA</td>
<td>Whether any NSC-type Investments</td>
<td>Assets and savings</td>
<td>Accounts and investments</td>
</tr>
<tr>
<td>ASSETS</td>
<td>ADCH</td>
<td>Whether Adult or Child Assets</td>
<td>Assets and savings</td>
<td></td>
</tr>
<tr>
<td>ASSETS</td>
<td>AMOUNT</td>
<td>SAYE: amount regularly paid</td>
<td>Assets and savings</td>
<td></td>
</tr>
<tr>
<td>ASSETS</td>
<td>ASSETYPE</td>
<td>Asset Type</td>
<td>Assets and savings</td>
<td></td>
</tr>
<tr>
<td>ASSETS</td>
<td>HOWMANY</td>
<td>Number of shares/bonds/units held</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>ASSETS</td>
<td>HOWMUCH</td>
<td>Value of asset (respondent's estimate)</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>ASSETS</td>
<td>ISSDATE</td>
<td>Date NSC issued</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>ASSETS</td>
<td>ISSVAL</td>
<td>Value of NSC</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>ASSETS</td>
<td>PD</td>
<td>Poode: SAYE: amount regularly paid</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>BENUNIT</td>
<td>TOTSAV</td>
<td>Estimated value of accounts/investments</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>CHILD</td>
<td>ANYMON</td>
<td>Any money left in current account</td>
<td>Assets and savings</td>
<td>Capital value</td>
</tr>
<tr>
<td>CHILD</td>
<td>CHINC</td>
<td>Whether any income-producing investments</td>
<td>Assets and savings</td>
<td>Accounts and investments</td>
</tr>
<tr>
<td>CHILD</td>
<td>OTINVC</td>
<td>Whether any NSC-type investments</td>
<td>Assets and savings</td>
<td>Accounts and investments</td>
</tr>
<tr>
<td>CHILD</td>
<td>TOTSAVE</td>
<td>Estimated value of accounts/investments</td>
<td>Assets and savings</td>
<td></td>
</tr>
</tbody>
</table>

Note the first- and second-level classifications from the FRS project manager's classification system. 'Variable' is the variable name (each one of which had an associated web page); 'Table' is the table in the hierarchical dataset with which each variable is associated; 'Label' gives more information about the variable.

I then coded this information into the JavaScript for the two search facilities. Table 7.3 shows the JavaScript for the search engine, for two variables, ACCINT and NSAMT. Lines beginning 'title' show the search terms entered for each variable (compare with the topic classification system for 'Assets and Savings'); lines marked 'desc' give the variable description (the label from the metadata); lines marked 'links' point to the variable HTML file. The search terms that I entered into the 'title' lines were taken directly from the first- and second-level classifications, and I also included the variable name.
Table 7.3 JavaScript showing how search terms are entered for variables in the topic search

```javascript
    title[2] = "assets savings interest dividends accint"
    desc[2] = "Interest received"
    links[2] = "accint.htm"
    matched[2] = 0

    title[6] = "assets savings capital value national nsamt"
    desc[6] = "Value of National Savings Investments"
    links[6] = "nsamt.htm"
    matched[6] = 0
```

For the drop-down search, I once again coded in the first- and second-level classifications devised by the FRS project manager. Figure 7.9 shows this text.

Figure 7.9 Classification search for Assets and savings

![Classification search for Assets and savings](image)

Table 7.4 shows the JavaScript which I wrote to produce this function for the two levels of classification. All first-level classifications are shown in the first section of this code; the second section of the code shows second-level classifications for
'Assets and savings': again, the code I wrote drew directly from the classifications devised by the FRS project manager.

Table 7.4 HTML and JavaScript code showing how search terms are entered for variables in the classification search

```
<p class="prose">Select first level:
<select onchange="firstLevelItemChanged(this.selectedIndex)"
<option>1. Assets and savings
<option>2. Care
<option>3. Consumer durables
<option>4. Council tax
<option>5. Demographic characteristics
<option>6. Employment
<option>7. Health
<option>8. Housing benefit
<option>9. Housing costs
<option>10. Insurance policies
<option>11. Maintenance
<option>12. NHS services
<option>13. Non-state pensions
<option>14. Other income
<option>15. Other social security benefits/tax credits
<option>16. Qualifications
<option>17. Tenure
<option>18. Travel to work
</select>
</p>
<p class="prose">Second level:</p>
<div id="div1" class="prose">
<ol>
<li><a href="topicsearchassets1.htm">Accounts and investments held</a></li>
<li><a href="topicsearchassets2.htm">Capital value</a></li>
<li><a href="topicsearchassets3.htm">Interest and dividends</a></li>
<li><a href="topicsearchassets4.htm">Other</a></li>
</ol>
</div>

The behind-the-scenes code that I wrote for these two search facilities therefore drew directly on the detailed classification system devised by the FRS project manager (and the automatically generated Excel file which contained the metadata). In Figure 7.10, I show the instructions which I wrote for using the search facilities and which I put into the help file: note the objective tone of the language used, and the passive voice in sentences such as: "Links are given to lists of variables, and from there to individual variable pages." A collaborative process between myself and the FRS project manager re-inscribed an existing classification system onto the website, and had the end result of an anonymous piece of text.
The search facilities which I coded for the FRS documentation website were therefore extremely constrained by existing classifications; a 'cascade of inscriptions' was created, whereby the information from the Excel metadata file was reclassified and this two-level classification system coded into the JavaScript for the search facilities. I have described above the organizational and technical actors which gave rise to the Excel metadata file; once again, this proved to be an immutable network that was not displaced as the FRS team manager and I developed the classification systems for the search facilities. This example also shows the limited scope for authorship. Not only did I work collaboratively with the FRS team manager, but the classification system and the search facilities which were produced were very closely tied to already existing material.
Collaborative authorship 2: changing the form of the site

In this section, I want to look at collaborative authorship in terms of the disparate influences I took into account when making a particular design decision which had the potential significantly to alter the form of the site. I had the choice between following the advice given in HTML manuals, or the opinions expressed by FRS team members in interviews I conducted with them about using the documentation website. This example from the development of the site can therefore be seen as being concerned with authority: I had to decide which would carry more weight - the advice given in the HTML manuals or the expectations which users expressed in interviews.

The design decision surrounded whether or not to use the HTML attribute \texttt{target="_blank"}. This attribute and its value had the potential to alter the form of the website by allowing multiple windows to open. The \texttt{target="_blank"} attribute (as discussed briefly in chapter 6, p. 188), when used in conjunction with the anchor \texttt{<a>} tag (and hence as part of a hyperlink), opens up a new window when a link is followed, rather than moving the user on to the requested page within the same window. Table 7.5 shows the HTML I used to produce a link from the ADULT table page \texttt{(adult.htm)} to the variable page for the variable \texttt{ABSINO [absino.htm]}. In the first row of the table, the \texttt{<a>} tag has no attributes; this would open the \texttt{ABSINO} variable page in the same window. In the second row, the \texttt{<a>} tag contains the \texttt{target="_blank"} attribute, which will open the \texttt{ABSINO} variable page in a new window, leaving the ADULT table page window open. (Figure 7.11 shows the ADULT table page with the link to \texttt{ABSINO} highlighted in red, i.e. how this HTML appears in a web browser.)

Table 7.5 Comparative HTML for opening variable pages

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;tr&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;/tr&gt;</td>
</tr>
</tbody>
</table>

| 2. |   |
|   | <tr> |
|   |   |<td><a href="absino.htm" target="_blank">ABSINO</a></td> |
|   |   |<td>Weeks away during this spell of absence</td> |
|   | </tr> |

Extensive use of this attribute is considered to be bad style in HTML manuals. Niederst (1999) writes:
'If you set a link on your page to target a _blank window, every link will launch a new window, potentially leaving your user with a mess of open windows' (p. 137).

This can lead to precisely the lack of structure against which Girill and Luk (1992) warn; other manuals I consulted encouraged the sparing use of this attribute.

Figure 7.11 Linking to new windows
As the content of the site expanded, I could see two places where the \texttt{target=_blank} attribute might in fact make the site more user-friendly.

- \textit{The help file}. I had provided a link to the help file from all pages, and written the help to be context-sensitive, i.e. the user would be linked directly to the relevant part of the help file rather than to the top of the page. I thought that it might be more useful to have the help file open as a separate page, since users would then be able to read and act on the help information whilst still being able to see the page to which it referred.

- \textit{Opening variable metadata pages}. Implementing this was more likely to lead to a 'mess of open windows', since there were 1440 such pages. However, I also thought that it was possible that users would prefer to be able to compare the information on several metadata pages simultaneously. In the interviews which I conducted in the final phase of the project, I asked users about opening new windows in two places on the site:

  - When linking to variable pages from table pages (as shown in the example above). I coded this option on the ADULT table, so that users could compare it with the links from the other tables, where new windows did not open.
  - When linking back from the questionnaire to variable pages, so that users could compare the route through the questionnaire to a particular variable alongside information about the variable itself.

When I conducted the interviews, I specifically asked users whether opening new windows was confusing or useful. Despite the advice from the manuals, the people that I interviewed showed a distinct preference for having new windows open in each of these cases, as shown in these statements made in interviews:

'I\textit{f you're doing more than one thing, you don't want to keep on going back... you can have more than window open at one time, and you can compare them... if I wanted to compare the differences between

231
variables, it's better to have two windows open, rather than having to move backwards and forwards to check.\textsuperscript{79}

And:

\textit{What I tend to do when navigating a website is to open things in new windows; that way I know if I get lost in its structure, I've got all the windows open at the bottom, I can work out where it is.}\textsuperscript{80}

Rather than making moving around the site confusing, then, the users of the documentation that I interviewed thought that opening new windows in these cases was not just an active aid to navigating the site, but also more appropriate for the material that was being examined. As a result, I implemented the \texttt{target=_blank} attribute throughout the site.

My decision to use the \texttt{target=_blank} attribute can be seen as one where I chose the tacit, informal expertise of users over the conflicting 'authority' of the written form in the HTML manuals which I was consulting. One would, perhaps expect to attribute more significance to the written word. What factors might then have disposed me towards accepting tacit knowledge over the written advice of the manuals? Perhaps the 'genre' of computer manuals lends itself less to treating the contents as authoritative, and more to viewing them as 'advice' rather than as rules which should not be broken. My own sense is that, as the project progressed, I had developed an increasing awareness of the importance of the users of the site. As I documented in chapter 6 (p. 171), the 'story' of the project turned from a design-driven pilot to a user-driven prototype. There were reasons internal to the progress of the project for me to have an increasingly compelling sense of the importance of the users, and to give more weight to their opinions on how the site should look. The 'audience' for the site had, by this stage (close to the end of the project), changed significantly from the perceived academic audiences which I had been addressing in the pilot stage. Nevertheless, my design decisions were still similarly motivated by a need to satisfy that audience's requirement; they still needed justification, and I

\textsuperscript{79} Interview with LC, 22/07/02.
\textsuperscript{80} Interview with JS, 22/07/02.
supported my design decisions with what I had learnt about the users' expertise. This user expertise derives from two different sources. Firstly, users had prior knowledge of the structure of the hierarchical dataset (so, *specific* knowledge related to the FRS). Secondly, users had general, prior knowledge of how to use or read Web pages (so, *general* knowledge related to increasing Web-literacy). I shall discuss each of these in turn.

The information which the variable metadata windows contained had been highly hierarchically pre-structured. As I discussed in section 3, I had conceived of the variable metadata documentation specifically on the lines of Hunter's 'type' of topic-driven hypermedia texts, i.e. a multi-dimensional filing system. Users access these pages via a structure which mirrors the hierarchical dataset - from the hierarchical dataset front page, to table pages, to variable pages (see Figure 7.3). As Hunter says:

'[S]ubstantial categorising and hierarchising had already taken place, which could not be disrupted without unhelpfully disordering the expectations of users' (1999: 113).

Users were already familiar with the structure of the pages, and this was because I had derived it from the structure of the FRS dataset.

Users were familiar too with how to use Web pages generally: when I asked in interview about the layered structure of the site, and the navigational links on the left-hand side of the pages, one respondent commented:

'I like websites where there's a button that will take you back to the front page... I like them hierarchical, with sections and pages within that section... This follows that principle. I like websites that are hierarchical. I know how easy it is to get lost sometimes. With a book, you know you read from left to right and if you want the index you flick to the back; similarly, with a website, it's not something you think of anymore, you know the menu's on the left. Users of the FRS will be fairly web-savy. "

81 Interview with JS, 22/07/02.
Again, then, users drew on tacit knowledge - here, how to use a website - to enable them to make sense of the information before them. There were user expectations not just as to the content of the website but, based on experience of other sites, there were also expectations as to how this site would and should behave. As users said in the interviews which I carried out, this combination of knowledge about how to use the Web, and knowledge about what to expect from the organization of the site was, I decided, sufficient grounds for me not to follow the advice given in the HTML manuals to avoid using the target="_blank" attribute.

It could be argued, perhaps, that the feedback which I received from users on the use of multiple windows is more comparable to, for example, editorial comment. I want to characterize it as authorship because of the significant impact it had on the form of the site, and also because the advice I received ran counter to the 'authority' of the HTML manuals which I was consulting. I made a design decision which weighed in favour of the experience and 'savvy' of department members in the face of the authority of HTML manuals instructing me about the best way to write a site. As I have indicated earlier in this section, I believe that this decision primarily arose from the direction in which the project had been going towards a user-driven prototype. At this point in the development - the writing - of the site, it was feedback from its users which had the greatest influence upon it.

Was this a radical move allowing innovation in the face of technical authority? Or was I raising the informal experience of users to the status of knowledge; as Porter (1995) puts it, did formalizing the tacit expertise of the users into the structure of the site 'produce knowledge independent of the particular people that make it (p. xi)'? Where one stands on this question depends, I suspect, on whether one views the associations of the actor-network of government social survey research primarily as having much in common with a wider network of governance; or whether one views the objectivity aimed for and achieved by civil servants as differentiating it significantly from that political context. Certainly some commentators on the interaction of scientific method and bureaucratic culture are scathing of its effects in masking sources of authority; Porter (1995) in particular fulminates that '[o]bjectivity lends authority to officials who have very little of their own' (p. 8)? My own preference is to follow the lead of researchers such as Haggerty (2001) and take practitioners seriously on their own terms. As Haggerty (2001) writes in a study of the Canadian Centre for Justice Statistics:
'Centre employees are keenly aware that the validity of their numbers is underwritten by public trust and that insinuations of political partisanship can undermine that trust' (2001: 270).

The objectivity of tone for the civil servants in Haggerty's study was seen on their own part as an essential aspect of their professional identity that allowed them to maintain a necessary distance from the political 'masters' that they were serving; part of the boundary work done to enable them to differentiate their professional expertise from the political purposes which it was drawn upon to serve.

As I have tried to show in this section, the objectivity of tone with which the FRS documentation website was authored derived from a collaborative and anonymous writing process. The highly technical content of the subject matter gave rise to at least some of the anonymity of the text; as did the way in which the technical information was re-inscribed upon the site from, for example, the Excel metadata file. I would argue too, however, that it was also a consequence of the institutional setting in which I was producing the texts. Bringing on other staff members to collaborate in the production of the site also contributed to the anonymization of the site text.

Haggerty (2001) writes that 'measurement is an institutional accomplishment' (p. 706); the same holds for the documents and texts that are achieved in these institutions by means of a collaborative and anonymous writing process. The objectivity of tone, derived from both a bureaucratic and a scientific ethos, may well impersonalize the official; it simultaneously works towards securing his or her trustworthiness and non-partisanship - it is an important part of professional identity.

4 Summary

In this chapter, I have used detailed examples from the development of the FRS documentation website to illustrate the social processes of the development, and to explore the writing of a bureaucratic text. I summarize these now.

The project emerged within the context of a general shift to computerization in survey research, and also within the context of specific policy initiatives within UK government which aim to increase the dissemination of information online. Information society narratives were therefore an influence on the text at the earliest
stage; for example, informing the decision to put the documentation online to such a degree that there was little perceived alternative. Authorship of the FRS documentation website was an anonymous and collaborative process. It took place within the context of strong technological and organizational networks which constrained the scope for authorship on the site: I drew heavily upon automatically generated texts and prior classification systems when writing the documentation site.

At the pilot stage of the documentation, I was primarily concerned with formulating academic questions related to the nature of hypertext; I drew extensively upon academic accounts of hypertext in making my design choices. At the second stage of the project, I was concerned with what was wanted by users of the site. When making one design decision at this stage of the development, I was sufficiently guided by perceived user expectations to choose their tacit expertise over the written advice of manuals. At the pilot stage, I was concerned with justifying my design decisions to a perceived academic audience. As the project went on, my sense of the site's audience shifted, and I became more guided in my development choices by the users of the site. In general, however, my decisions were primarily guided by a sense of the site's audience, and my design choices can be characterized as a series of rhetorical moves made to justify the decisions to these perceived and changing audiences.

In this chapter, I drew upon specific and detailed examples of design decisions that I made during the course of the development of the site, in order to illustrate two themes: influences upon the text and authorship in a bureaucratic setting. My focus so far has, therefore, been mainly on the social processes behind the development of the FRS documentation website (i.e. with the social shaping of the site). In the next chapter, I want to look at how the project as a whole underwent several 'translations'; how the 'story' of the project changed, and why (what I characterized in chapter 3 (p. 57) as the social construction of the site; again, note that this is a loose conceptual distinction that I am drawing primarily to help me organize this discussion).
8 Translating technology

1 Introduction

In this chapter, I shall look at my own role as the 'translator' of the project. In section 2, I shall examine more closely two points in the development where the meaning of the project was in flux. I examine how and why the project was translated into different terms; i.e. I shall look at how I went about re-negotiating the project at these key points in the development. In particular, I want to look more closely at the justifications I deployed to make these decisions; i.e. the rhetorical work I carried out to support my choices during the development process, to ensure the continued support of those most closely associated with the project (i.e. the FRS project manager, the IT staff manager, users of the site, and myself). My 'translations' of the project entailed me drawing upon a variety of textual (and other) resources to make the project meaningful for all involved.

In the last section of this chapter, I consider the final translation of the work I carried out, reflecting upon the process of producing a coherent academic narrative of this project, and using Law's (1999) account of academic 'storytelling' to consider a different model of writing narrative. I reflect upon what some of the ramifications of describing oneself as a 'writer of narrative' are for the social scientist, particularly in terms of 'boundary work', as I discussed in chapter 2 (p. 23).

2 Changing perceptions of audience

Akrich (1992) studied a process of technology transfer between a Swedish company in possession of a machine to convert waste material into briquettes, and the Nicaraguan government (Nicaragua being short of fuel). Akrich examines the series of negotiations by which the technology was transferred (e.g. experimental setbacks, storage issues, finding a market for the machine's products). Her key point is that during this process the machine itself began to change as it was transferred - translated

---

82 I have, in effect, turned from focusing on the social shaping of the documentation to considering the social construction; see chapter 3 (p. 57) for more on this conceptual distinction, which I am making primarily in order to organize my material.
- from Sweden to Nicaragua. The project to develop the FRS documentation website was not a project of technology transfer, but it did involve negotiations between actors during the course of which the technology changed and took on different meanings: the project went through a series of translations. In this section, I want to look at two particular points in the development of the FRS documentation where the meaning of the project was 'under negotiation'; where decisions were taken to follow a particular route in the development; how these decisions were made; and how the website changed as a result of these negotiations.

Throughout this section, I shall continue to highlight, as I have done throughout this chapter, the significance of the audience for the site. By the end of the project, my perception of the site's audience had shifted markedly from a perceived academic audience to the potential users of the site within the DWP. 'User expectations' gained increasing significance for several decisions which I took, including my decision not to use certain standards for online documentation, and also my construction of the project in relation to the TADEQ project (see also section 3, above, where I discussed how I chose user expertise over the advice in manuals in my decision to use the target="_blank" attribute). This transition in my perception of the audience for the site was complete by the end of the development work, by which time I was defining the project in terms of it being a 'user-driven prototype' (see my discussion below).

Law (1999), writing about Akrich's study, notes that the work 'shows that translation implies both similarity and difference. Similarity, for there is some sense in which it is possible to say that the briquette machine in Nicaragua is "the same" as the briquette machine in Sweden; but also difference, because by the time it has been located in its new Nicaraguan environment it has undergone many changes' (pp. 3-4). Of the two examples from the development of the documentation website that I shall now discuss, one is primarily concerned with translating the site into terms which differentiate it from other actors in the networks in which it is embedded; the other is primarily concerned with the translation of the project into terms of similarity of purpose between three players involved with the project (i.e. with finding consensus between the three players). The first example concerns the work which I carried out on the TADEQ project; the second concerns the shift towards describing the project as a user-driven prototype. (I return to the final translation which took place - from website to thesis - at the end of this chapter.)
In chapter 6 (pp. 173-174), I discussed my involvement with testing the software developed in the TADEQ project. The focus of my discussion there was to indicate how I drew upon aspects of the TADEQ tool for the design of the questionnaire in my own project. I have noted already how my work on testing the TADEQ tool brought my own project into connection with a wider network of practitioners involved in developing methods of documenting electronic questionnaires. I want to show now how I translated my own project into terms such that it was differentiated from the TADEQ project; this process of negotiation is also related to reasons why I chose not to use either of the markup standards for electronic questionnaire documentation which were under development at the time (see chapter 6, pp. 171-172). This was an extremely significant series of decisions for me in relation to the work that I was carrying out. If my development work followed too closely what was being done on the TADEQ project then, I believed, it would be insufficiently differentiated to be valuable to either myself (as developer of the site and for my thesis), or to the DWP (as the sponsors and intended users of the project). My attempts to differentiate my own work from the TADEQ project therefore were to do with providing legitimacy or justification for the project within the broader context of work already being done by organizations external to the DWP who had been involved in the FRS largely since its inception.

The TADEQ project was a collaborative R&D project funded by the EU ESPRIT Programme (TADEQ: a Tool for the Analysis and Documentation of Electronic Questionnaires). The project was aimed at developing a tool to make a human-readable presentation (on paper or electronically in hypertext format) of the electronic questionnaire. The project focused on the development of an XML markup standard for electronic questionnaire documentation: the Questionnaire Definition Language or QDL (see chapter 4, pp. 108-110, for a more detailed discussion of markup standards). The project involved national statistical institutes, research institutes, and commercial marketing research organizations. One partner in the project was Statistics Netherlands (developers of the CAPI program BLAISE, in which the FRS is programmed); another partner was the Social Survey Division of the ONS (with NatCen, ONS currently hold the contract for conducting the FRS).

As I tested the TADEQ software, I became increasingly aware of this large degree of overlap between it and my own project. The purposes of the TADEQ project (to provide online documentation for electronic questionnaires) mirrored the
purposes of my own project (to provide online documentation for the FRS), and the
FRS was one of the questionnaires which the TADEQ tool was documenting. My
work on the TADEQ project specifically involved testing the tool using the FRS. As I
worked on the testing, I began to look for ways in which it did not answer specific
requirements for the FRS, so that I would not be repeating work that was already
done: that is, I became involved in what can be characterized as a process of
negotiating difference between my own project and the TADEQ project. As I stated
above, this was done substantially to allay my own concerns that I could easily find
myself repeating work that was already being carried out by actors that had been
involved in the FRS for a long time and that had more experience and resources (and,
therefore, more legitimacy). This process of negotiating difference was, therefore,
concerned with providing sufficient justification for my own project in such a way
that both myself (as the person carrying out the work) and the FRS team (sponsoring
the work) would be satisfied with the documentation website alongside a project like
TADEQ.

The resolution of this situation came when I considered the expectations of
users of TADEQ and the FRS documentation. Although there was overlap, the
intended users of TADEQ did not map precisely onto the people at the DSS/DWP
using the FRS documentation. As I noted in chapter 6 (pp. 172-173), the TADEQ tool
led to some loss of detail familiar to users of the FRS. For example, in the metadata
documentation, the terminology used for the answer types by TADEQ differed from
that used in the Excel metadata documentation produced specifically for the FRS.
TADEQ groups variables according to six types (numeric, real, numeric integer, open,
closed, date, and time variables). The Excel metadata file generated from BLAISE for
the FRS organizes this information differently: eleven types of variable are listed in
this documentation (categorical, date, frequency, key, monetary, period code,
quantitative, string, date components, weekly, and system variables). In addition, the
variables are categorized according to usage. These differences were related to
TADEQ's standardization process. In order to be useful for a variety of electronic
questionnaires (not just the FRS), developers had come up with a solution which
applied generally to the broad range of surveys, but which led to the loss of survey-
specific information. I was therefore able to use how these differences might be
inconvenient for users of the FRS documentation to differentiate my own work from
TADEQ (i.e. to give it meaning beyond the TADEQ project) and, also, to justify several design decisions.

For example, I took the decision not to use the XML standard which the TADEQ tool had developed, in favour of documenting the FRS in such a way that it kept this survey-specific information. As a further result of this, I chose not to use the other XML documentation standard that was in existence, the Data Documentation Initiative Document Type Definition (DDI DTD; see chapter 6, pp. 171-172). I was now able to position my project to myself and to the FRS team manager as providing solutions which were specifically for the FRS documentation and its users. This satisfied my own need for a distinctive project, and the sponsors for a useful piece of work. As Law (2000) writes, 'all entities [ANT says] achieve their significance by being in relation to other entities' (p. 4). At this point in the development, I translated my project into terms which would distinguish it from the TADEQ tool in order to give it significance and meaning in relation to the TADEQ project. This negotiation of difference between the projects also contributed to my decision not to use available standards. In retrospect, this also marked the point where the audience of the site began to shift from the perceived academic audience that I discussed in chapter 7, and towards the users of the site within the DSS/DWP. Users of the site became the most significant element in the social construction of the FRS documentation site; I shall now go on to describe this process.

In chapter 6, I discussed how the project can be seen as falling into two phases, each on either side of a major point of change for the FRS, when the long-term project manager moved on and another project manager was appointed from within the FRS team. There was a short hiatus in the development work I was doing at this time. Prior to this break in the project, my work on the site had focused mainly on putting online the existing documentation, i.e. paper-based documentation and the variable metadata information, and developing search facilities for the metadata. When I began work on the site again, I was working closely with the new FRS project manager, and a member of staff from the IT department.

During the course of a series of meetings involving the three of us, we discussed what each of us wanted the project to achieve, and how this translated into work to be done on the site. The FRS project manager was concerned with making sure that the project was useful from the point of view of the FRS team. The member of the IT department was most interested in receiving user feedback (which time
constraints did not usually allow technical staff to get). My specific interests at this point were two-fold. Firstly, I had practical concerns, i.e. I was aware that none of the questionnaire documentation itself was yet online, and I wanted to work out a set of specific tasks related to this that would be most useful for the department and could be achieved within the time remaining on the project. Secondly, as I described in detail in chapter 7 (section 2), my theoretical interests had begun to move away from the design of the website towards its use (i.e. from writing the site to exploring how it might be 'read'). As I have described, I was in the process of shifting away from Hunter's (1999) classification of hypertext applications as the 'best fit' for my own project.

The FRS project manager, the IT staff member and I negotiated three specific outcomes from these discussions, all of which focused on the people who would be using the documentation. Firstly, we began an ongoing process of revising the layout of the site, such as on the individual variable pages, and the organization of the site as a whole. Secondly, we conceived of the process of putting online the questionnaire documentation as one of providing a user-driven prototype. This would give users in the department a representation of how online documentation might look, rather than my attempting to provide specific technical solutions, which was within the IT department's area of expertise, rather than my own. Thirdly, as a corollary of this, and so that I could obtain feedback from users, I would make a version of the online documentation available to users within the department, and carry out interviews with them to get feedback on the site's layout and functionality for the IT department, and also to pose some questions that were of interest to me. In the course of a series of discussions, then, the FRS project manager, the IT staff member and I redefined the project in terms of its users, and this underpinned the work that was done for the remainder of the project (building an online prototype of the questionnaire documentation, reworking the layout and structure of the site, and conducting interviews with users). The project was translated into terms about which the three people most closely involved could agree.

In part, I would attribute this to the time constraints which were becoming more significant - the period of time for which the DWP was funding my project was, by this point, more than half over. This sense that only a certain amount of time was left for the project helped focus all three of us on defining what we all believed would be a useful project which could be completed in the remaining time available. Again,
however, I would note the importance of users in our redefinition of the project; all three of us were very conscious of the eventual audience of the site. Furthermore, the meaning that was fixed at this point in the project in turn became a constitutive element in influencing the form the documentation finally took: the feedback which I received from users was the most significant influence upon my decision to allow multiple windows to open (see chapter 7, section 3). From a hermeneutic perspective, this kind of ongoing interaction and exchange between perspectives can be understood in terms of the 'fusion of horizons': the emergence of a new view of the project which drew upon older conceptions and about which there could be consensus from all participants.

In this section, then, I have shown how the meaning of project went through two significant translations: one differentiated the project from the TADEQ project; the other shifted the project away from its focus on a perceived academic audience towards the users of the site. In the last section of this chapter, I want to discuss the final translation that the project underwent: back into an academic narrative.

3 From bureaucratic text to academic narrative

There was, of course, at least one other translation of this project to be made - into academic terms. In this section, I shall begin by describing the difficulties I had in writing an academic account of this study, and suggest some reasons as to why this might have been the case. I shall then conclude by reflecting more generally upon the writing of academic narrative, and I shall organize my discussion largely in response to two articles by Law (1999; 2000). In the first, he writes about actor-network theory as a kind of academic storytelling; in the second, he presents a 'story' about working in a bureaucratic context. These articles address several themes which I want to reflect upon, such as what kinds of narratives academics might choose write; the position of the academic working in a bureaucratic context; and some of the ramifications of describing academic writing as 'storytelling'.

When I first attempted to write an academic account of the FRS documentation project, I automatically reached for a very specific kind of language. The result of this was that I wrote the earliest version of chapter 6 (i.e. my description of the project) almost entirely in the passive voice and in the tone of a scientific
report. Much of this language can still be seen in the way I have structured the chapter, in terms of 'tasks', 'outcomes', 'solutions', and my presentation of the project as having a coherent narrative of the development progressing from a pilot stage to a later stage. Throughout the development of the FRS website and whilst I was working with the various team members, it was certainly easiest for me to think of myself in terms of being a part-time member of staff at the department, and to construct the work in terms of 'tasks', 'outcomes' and 'solutions', i.e. to adopt the standard language of project management and report-writing. This language served, however, to obscure the social processes of the development, and when I came to write an academic account of the project, I wrote it first in precisely the kind of anonymous language which I have described in chapter 7. Translating the project from this into an academic narrative was a difficult process for me.

Why might this have been the case? Firstly, the work which I was carrying out was not immediately recognizable as 'field research'. Apart from the very few interviews which I conducted towards the end of the project, my work on the documentation was entirely practical, developing the website; and a consequence of this was that there was no obvious academic narrative model for me to follow when I came to translate the project into academic terms. Another reason might be the strength of the networks within which the documentation project was embedded; I have discussed this already (chapter 7, pp. 216-220), where I indicated that I had a compelling sense of the tone in which I felt I should be writing the text of the website. I also described in chapter 7 the strength of technological actants such as the BAD and the Excel metadata, many aspects of which were re-inscribed into the documentation site. In addition, as I have emphasized throughout this chapter, my focus shifted perceptibly in the second stage of the project away from academic interests and towards practical ones. It is difficult to determine whether this shift contributed towards my difficulties in writing an academic account or whether it was a result of the changes in academic focus that I described in section 2 (and a consequent weakening of the influence of academic priorities upon the work I was carrying out); nevertheless, by the end of the project, the scientific and bureaucratic tone that I had used in writing the site had a 'stickiness' which made it difficult for me to make a transition to writing an academic account of the project.

I now want to reflect in more general terms on academic writing; in particular, I want to reflect upon two articles by Law (1999; 2000) on actor-network theory, and
on working in a bureaucratic context. I am drawing upon these two studies in order to write some kind of narrative about my own study and complete my own translation of it into academic terms. In the first article, Law (1999) positions ANT as a form of storytelling, and emphasizes what he calls 'diasporic' character:

'is there such a thing as "actor-network theory" at all? Answer: yes. We can certainly make a story that tells of unity. But the answer is also no, for it is just as easy to tell tales of a kind of diaspora, of interaction with other "theories", of confusion, or if you prefer, of complexity, overlap and partial connections' (p. 4; his emphasis).

I shall return below to the ramifications of presenting academic writing as storytelling; for the moment I want to examine a specific story of working in a bureaucratic context that Law tells in his second article, and relate this to my own experience of working in the DWP. In the second article, Law (2000) presents a story about a study he conducted of an unsuccessful British attempt to build a nuclear tactical strike and reconnaissance aircraft. He discusses (pp. 5-6) how he interviewed senior members of the RAF, politicians, senior civil servants, and executives and engineers of the aerospace industry. He describes how during the course of the interviews he became 'deeply uncomfortable' (p. 6) as he realized that these people were hoping he would document the failures of the project and that his study would be useful so that they could learn lessons for the future. He describes his discomfort as being two-fold: that he was 'colluding' in the process of military procurement in the possibility that he might teach the military anything useful, but also that the terms used by those he was studying were, more or less, the same as he was using to analyze the project himself. Here, Law is concerned with the performative rather than the descriptive aspects of analysis; that every description, as he goes on:

'helps to bring into being what it describes... we are... [t]ending to bring some relations into being, while pushing others out of being. We are always, then, in the business of making a difference - or, to put it differently and more negatively, we are always at risk of collusion' (p. 6; my emphasis).
He concludes:

'if we write as network analysts what we may be doing, what we're often doing, is buying into and adding strength to a functional version of relationality. One that is, to say it quickly, managerialist... All that is solid - human and non-human - melts into air in the face of the need to create a coherent, ordering, and functioning heroic or bureaucratic actor' (pp. 6-7; his emphasis).

The particular subject matter of Law's study no doubt invites such a dramatic language of 'collusion'. Nonetheless, one risk of language such as this, I would suggest, is that it produces a script for academic analysis which, in trying to unwrite the antagonistic figure of the 'functioning heroic or bureaucratic actor' produces as a mirror image of the academic as critical protagonist. And, as a consequence, it also runs the risk of conflating the role of the academic with that of the critic. In chapter 2 (pp. 35-36), I drew upon Pleasants' (1999) account of the conflation of theoretical representation and social critique, and I agreed with his contention that 'the presentation of a new "ontological picture" consisting of universally possessed tacit knowledge and a transcendental order of rules, does not ipso facto constitute a critical perspective on, nor an emancipatory intervention in, social life itself' (p. 77). Law (2000) writes of actor-network theory that 'built into its vocabulary is the idea of translation - and the idea that translation (the attempt to render equivalent) is also betrayal' (p. 7). But description and narrative are not necessarily forms of collusion and betrayal; representation is not, necessarily, critique. And writing the academic as critical protagonist risks narrowing the roles available to the academic; it makes the range of stances available less heterogeneous.

The language of 'collusion' and 'betrayal' does not sit easily with any narrative I would want to write of my own project. In the latter stages of my project, in particular, my interactions with DWP staff were marked by our attempts to achieve consensus over the project's aims in order to complete it within the time available, and to the satisfaction of each of us. Was it possible that a lack of a sense of 'collusion' on my own part led to my difficulties in translating the story of the project into an academic narrative? Again, I think that the 'story' of my own project is not suited to telling in such terms; conflict between protagonist and antagonist is not the only
option for narrative. Le Guin (1998) writes insightfully about the obscuring effects of the focus on conflict in the writing of narrative:

'Modernist manuals of writing often conflate story with conflict. This reductionism reflects a culture that inflates aggression and competition while cultivating ignorance of other behavioral options. No narrative of any complexity can be built on or reduced to a single element. Conflict is one kind of behavior. There are others, equally important in any human life, such as relating, finding, losing, bearing, discovering, parting, changing' (p. 146).

The story of my intervention to produce the online documentation for the FRS website allows me neither to come in praise of 'bureaucratic actors' nor to bury them. Sometimes stories, including academic stories, can be ordinary ones.

What purposes are served by characterizing academic writing as a form of storytelling? And what are the ramifications of describing academic writing in this way? Characterizing social scientific discourse in this way is certainly 'boundary work' (Gieryn, 1999), but does it limit the boundaries of social science? Why limit the role of the social scientific actor in this way? Why can social scientists not be heterogeneous in the roles that they perform? As a student sponsored, in part, by the DWP (writing both online documentation and this thesis), my own position was certainly a multiple one: both bureaucratic and academic actor, as well as designer of the website and author of a bureaucratic and collaborative text. Adopting a 'critical' strategy for the narrative of my project might have lent strength to the story I was writing, but in the case of this project it would have set up a false conflict between academic protagonist and bureaucratic antagonist.

---

83 Le Guin, for example, whom I cited above, is both a storyteller and an anthropologist, and tells anthropological stories through her writing. Making the claim that academic writing is a kind of storytelling implies that it has some access to 'extra-methodic' truth (see chapter 2, p. 32), but there is not a resultant need to construct a conflict between the social scientific actor and the natural scientific actor, rather than acknowledging affinities and family resemblances. There is no need to make Taylor's (2002) formal break between the natural and social sciences (see chapter 2, p. 10). See my discussion at the start of chapter 2 (pp. 11-12) of Geertz's response to Taylor's description of a 'watershed' in the intellectual world: does characterizing social scientific discourse as storytelling keep the meanings but, to lapse for a moment into antagonistic terms, surrender the mechanisms?
4 Summary

In this chapter, I described the main 'translations' made during the course of this project. I summarize these now.

In the first section, I described two key points at which the project was redefine: the first was a process of differentiation between my own project and the TADEQ project; the second was a process of negotiation between myself, the project manager and the IT staff member to find consensus on project aims. Throughout this section, I continued to emphasize the importance of the audience for the site, and the justifications I was making to support my design choices, i.e. the rhetorical work I carried out during development.

The final, key translation that I made in this project was from the website to the thesis. As I have described in this chapter, it proved difficult for me to make this transition to writing an academic account of the project. I attribute this to two main factors. Firstly, I found it difficult to write an account of the project that did not use the anonymized tone of the site. This resulted from several issues including: the strength of the organizational and technological networks in which I was embedded; the project's progress from being concerned with design and academic issues in the earlier stage towards the users of the site in the second stage; and my position as, in effect, a part-time member of staff. Secondly, I had difficulty finding an appropriate academic account to emulate. Because the work I had carried out was not immediately recognizable as 'field research', there was not an obvious academic narrative model for me to follow. In addition, I found that 'critical' academic narratives did not suit, in part because my interactions with staff at the DSS/DWP had been marked by our attempts to achieve consensus over project aims in the time available.
9 Concluding remarks

In this chapter, I shall summarize the main findings of the thesis, and then go on to reflect upon the significance of the project and suggest areas for future research.

This thesis has presented a critical investigation of the implementation of online documentation for the Family Resources Survey (FRS). I conceptualized the work that I did in terms of it being both a technological application and the writing of a bureaucratic text.

I gave a critical account of this text. I described the background in which it can be located, giving an overview of the history and organization of government statistics, with particular reference to the departments and bodies relevant to the FRS, the computerization of the survey research process, and those shifts in policy which have increased the online dissemination of survey research material.

I documented the development of the site and described the project itself, drawing attention to the two stages of the project, which fell on either side of a major change for the FRS (the appointment of a new project manager for the survey).

I made a specific analysis of the text which examined the social process of the site's development, the choices made during this process, and the reasons for these choices. I drew upon hermeneutic and actor-network perspectives to consider three main themes: 'authorship' of the site; 'influences' upon the text; and my own role at various points as 'translator' of the site.

I used examples from the development work to characterize authorship of the site as a collaborative process. This process involved myself, FRS team members, perceptions of user needs and conventions of use. This collaboration contributed to the anonymization of the site, as did the fact that much of the content of the site was automatically generated. I examined points in the project where the classification systems present in earlier documentation were re-inscribed onto the documentation that I was developing, e.g. in the development of the search facilities.

I highlighted two significant influences upon the text: firstly, academic models of hypertext, which I drew upon extensively in the first stage of the project; and, secondly, perceived user expectations, which became increasingly significant as the project went on. In general, however, my design decisions were primarily guided by a sense of the site's audience: initially, a putative academic audience and, latterly, the
site's users. I characterized my design choices as a series of rhetorical moves made to justify the decisions to these perceived and changing audiences.

I reflected upon two key points of 'translation' in the course of the project. The first I characterized as a process of differentiating the site from associated projects such as TADEQ, which had legitimacy in the environment in which I was working; the second was a process of negotiating common goals for the project between myself, the FRS team manager, and the IT staff member which resulted in us characterizing the project as a user-driven prototype.

Finally, I reflected upon some of the difficulties which I had in making a third and final translation, from bureaucratic to academic text - from the website to this thesis - and suggested that my difficulties arose from two sources. Firstly, the strength of the technological and organizational networks in which the project was embedded made the anonymous 'voice' in which I had been writing the site difficult to move away from. Secondly, I had difficulty in finding an academic narrative model which suited my project: the project had involved a very limited amount of recognizable 'field research', and I was also not satisfied that 'conflict'-orientated narrative models, which posited an academic protagonist and a bureaucratic antagonist, best described the social process of this project.

As a piece of work located within the social studies of science, this thesis has documented a detailed case study of the implementation of a technological application within a government department. One way of characterizing this case study, perhaps, would be in a Kuhnian sense, as the description and analysis of 'normal social science' in action; i.e. a description and analysis of the routine processes of producing a bureaucratic text in a government department involved in the production of social statistics.

On a theoretical level, the thesis has worked with conceptions of technology as text to map concepts of the social shaping and social construction of technology onto the production and reception of text. In addition, I have attempted to engage with actor-network theory, and respond to a shift in that body of literature which posits an 'after' actor-network stage and suggests interaction between ANT and other theoretical approaches (see Law, 1999). To that end, I attempted to bring my hermeneutic orientation into dialogue with ANT, using conceptions from both to conceptualize aspects of the development of the site; for example, characterizing narratives of the information society as 'actants' upon the documentation. Adopting tools from ANT
restored to my approach a focus on authorship which is often lacking from textually-orientated analyses, where the focus is primarily on resistant reading and the appropriation of texts by readers. At the same time, my hermeneutic orientation allowed me to respond critically to academic models in the ANT literature which were reliant upon a 'conflict'-orientated model of storytelling, and to be able to reject this in favour of an account which better suited this project.

The literature on theoretical models of hypertext turned out to be of limited value; as the project went on, I found that I was not able to sustain an investigation of my earliest questions surrounding 'best fit' models of hypertext in relation to this particular case or, perhaps, I discovered that this was an unprofitable way of formulating my research questions. Whilst I was able to offer a characterization of the construction of those theoretical models of hypertext which have emerged from the humanities and qualitative social science (as a 'defence of context'), I was not, on the strength of this study alone, in a position to attempt a similar characterization of uses of hypertext beyond that literature. However, alongside the extensive discussion of CAPI in the literature, there is now also considerable interest in Web-based survey methods. I believe that examination of this new 'genre' in the literature on survey research methods might be of interest; for example, a comparison of rhetorical devices in this literature with those in the literature concerned with the introduction of the laptop into survey research. A very brief review which I made in this thesis of the literature on Web survey methods (chapter 4, pp. 106-107) showed the topics and goals in common with the earlier literature on the introduction of the laptop, and this literature may be open to further examination of the rhetorical use of technology within survey research.

Also within the context of the literature on survey research methods, this thesis reviewed the shift in UK government statistics towards the computerization of the survey research process, and has documented policy initiatives which have been increasingly moving survey research material towards online dissemination. The project itself was no more than a small-scale technological achievement; however, studying how the prototype created in this project was used in developing an automated version of the FRS questionnaire documentation would be of interest. For example, the final version of the questionnaire documentation which I developed made extensive use of <div> tags so that the IT staff would have a means of creating a database format for the documentation. These <div> tags used attributes which were
derived from the BAD. One area of study might be to examine whether the automation of the questionnaire documentation led to a further cascade of inscriptions from the BAD; another area of analysis might track the course of other inscriptions into the automated questionnaire form - for example, the classification system which the project manager devised for the search facilities.

My detailed description of the points of difference between my own project and the closely associated TADEQ project raises an interesting question concerned with technological convergence. The TADEQ project attempted to provide an automated documentation tool for a number of surveys (of which the FRS is only one, and the project was also an international one) and, as I described, this led to a loss of detail which was particular to the FRS. I described the extent of the legitimacy which the TADEQ project had (project partners included many of the bodies concerned with the production of the FRS) and how I used the fact that TADEQ lost survey-specific information to provide legitimacy for my own project. One interesting question surrounds whether a bespoke version of the online FRS documentation would continue to be developed, or whether the more general solution which the TADEQ project supplied would replace it entirely. In the latter case, an examination of the processes whereby information specific to the FRS was translated, re-inscribed, or lost would be of interest.
References


258


Appendix 1: Summary of contents of CD-ROM

Overview

The documentation for the Family Resources Survey is presented online on the DWP intranet and was authored in HTML and JavaScript. The CD-ROM appended to this thesis includes a version of the website for examination in conjunction with the text of the thesis.

File names related to this CD-ROM are given in square brackets and blue text, directing the reader towards example pages from the site relevant to the points under discussion in the text. All material on the site and the CD-ROM is Crown copyright.

Front page

The website's front page is accessed from the file: docufront.htm. The front page provides links to the two main sections of the site: general documentation (i.e. information which is related to each survey year), and version-specific documentation (i.e. information related to specific survey years). A help file can be accessed from each page on the site.

Please note that because this is a demonstration version of the documentation which is not linked to the DWP intranet, the links on the left-hand side of the front page are not active.

General documentation

This provides links to pages that gather together information which holds for each year. This includes:

- Background: survey information; overview and response rates [frsback.htm].
- Structure: where the datasets are stored within the department; full description of both the hierarchical and flatfile datasets [frsstruct.htm].
- Programming examples: SAS examples for hierarchical and flatfile datasets [frsprog.htm].
• Imputation: details on how missing values are computed [frsimpl.htm].
• Style guide: how to navigate the paper-based BLAISE automatic documentation using the styles function in Word [frstyles.htm].
• Questionnaire: description and instructions for using the BLAISE questionnaire [frquse.htm].

Version-specific documentation

This forms the bulk of the online documentation. The project provided documentation for one survey year (1998-1999) for the database, the metadata, and the questionnaire. This is presented here.

Database documentation

The database documentation covers information on each table in the dataset; provides links to detailed metadata related to each variable in the dataset; and links to search facilities:

• Table descriptions: a single page gathers together links to pages for each table in the hierarchical dataset [frshd.htm]. Each table page [e.g. adult.htm] consists of links to pages for each individual variable in that table.
• Variable metadata: each page contains metadata related specifically to the variable [e.g. adch.htm].
• Three variable search facilities have been provided [frshdtopicsearch.htm]:

270
**Metadata documentation**

These pages collect supporting metadata documentation which holds for multiple variables, and so is linked to from most variable pages. This includes information on:

- Benefit key: code numbers mapped alongside the benefits they represent [frshdbenkey.htm].
- Benefit map: mapping of benefit code numbers alongside relevant questions [frshdbenmap.htm].
- Period codes: this page maps the period codes alongside the period they represent [frshdpcode.htm].
- Usage: details of variable type (e.g. categorical, etc.) [frshdusage.htm].

**Questionnaire documentation**

The documentation of the questionnaire is split across three separate pages, each covering the parallel blocks: household schedule, benefit schedule, and assets block. Each of these pages is linked to from a single questionnaire documentation front page [frquest.htm]. For the purposes of this demonstration version, only the assets block for this survey is presented here [frshdassets.htm].
Appendix 2: Sample of the BLAISE Automatic Documentation

FRS35.QRenting

Questions about renters

Ask if: QAccomdat.Tenure IN [Part .. Squatting]

Landlord

QRenting

Who is your landlord?

HELP <F9>

(1) The local authority/council/New Town development/Scottish Homes
(2) A housing association or co-operative or charitable trust
(3) Employer (organisation) of a household member
(4) Another organisation
(5) Relative/friend (before you lived here) of household member
(6) Employer (individual) of a household member
(7) Another individual private landlord

Warn if: QAccomdat.Tenure IN [Part .. Squatting]

Landlord = RESPONSE

This is a 'Key Question': it is VERY IMPORTANT to get an answer here if possible. If you cannot do so (either now, or later) please make a Note about the circumstances.

Compute if: QAccomdat.Tenure IN [Part .. Squatting]

Housing_Benefit := 'Housing Benefit/ rent rebate/ allowance'

Ask if: QAccomdat.Tenure IN [Part .. Squatting]

Furnish

QRenting

Is this accommodation provided...

(1) furnished,
(2) partly furnished (eg. curtains and carpets only),
(3) or unfurnished?
**ResLL**

QRenting

Does the landlord live in the building?

(1) Yes
(2) No

**ResLL2**

QRenting

Does the landlord live in the same flat as you or not?

(1) Yes
(2) No

**YStart**

QRenting

In which year did you first become a tenant of this accommodation?

INTERVIEWER: 'YOU'=PERSON(S) NAMED AT 'HHolder', THAT IS... ^RentName.

(1) 1988 or earlier
(2) From 1989 to February 1997
(3) March 1997 or later
ASK IF: QAccomdat.Tenure IN [Part .. Squatting]
AND: Landlord IN [FrndRel .. OthIndiv]
AND: (ResLL = No) OR (ResLL2 = No)
AND: YStart IN [ToFeb97 .. AftMar97]

Ctract
QRenting

When you started to rent this accommodation ...READ OUT (RUNNING PROMPT)...

(1) ...did you and the landlord sign a written agreement,
(2) ...did you have a written agreement which you didn't sign,
(3) ...or did you just have an unwritten agreement?

ASK IF: QAccomdat.Tenure IN [Part .. Squatting]
AND: Landlord IN [FrndRel .. OthIndiv]
AND: (ResLL = No) OR (ResLL2 = No)
AND: YStart IN [ToFeb97 .. AftMar97]
AND: Ctract IN [Signed .. NotSign]
AND: YStart = ToFeb97

Short1
QRenting

There is a form of tenancy called a shorthold. It is for a fixed period and you had to be given a notice in writing by the landlord that told you it was a shorthold tenancy agreement. Here is an example of a notice to a tenant saying that the agreement is an assured shorthold. SHOW EXAMPLE OF NOTICE.

Does your agreement or notice state that it is a assured shorthold or not?

(1) Yes, an assured shorthold
(2) Other agreement
Most tenancies are assured shortholds which are for a fixed period. There are others, just called 'assured', which are not for a fixed period. For these you have to be given a notice in writing by the landlord that tells you it is NOT an assured shorthold tenancy agreement. SHOW EXAMPLE OF NOTICE.

Does your agreement or notice state that it is NOT an assured shorthold?

(1) Not an assured shorthold
(2) Other agreement

Most rents are agreed privately between landlord and tenant. Sometimes the tenant can apply to the local rent officer or rent assessment committee to decide on a fair rent which is then registered. Has your rent for this accommodation been registered as a fair rent in this way, or not?

(1) Yes
(2) No
There are various ways in which landlords can let accommodation. Will you please look at this card and tell me if your letting is one of these?

SHOW CARD C

CODE FIRST THAT APPLIES.

(1) Company licence
(2) College licence
(3) Non-exclusive occupancy agreement
(4) Holiday let
(5) Low season let
(6) None of these
Ask if: QAccomdat.Tenure IN [Part .. Squatting]
    AND: AccJob = Yes

AccJobPer

QRenting

Who is that?

CODE ALL THAT APPLY.

SET [14] OF
    (1) 'DMName[1]
    (2) 'DMName[2]
    (3) 'DMName[3]
    (4) 'DMName[4]
    (5) 'DMName[5]
    (6) 'DMName[6]
    (7) 'DMName[7]
    (8) 'DMName[8]
    (9) 'DMName[9]
    (10) 'DMName[10]
    (11) 'DMName[11]
    (12) 'DMName[12]
    (13) 'DMName[13]
    (14) 'DMName[14]

Check if: QAccomdat.Tenure IN [Part .. Squatting]
    AND: AccJob = Yes
    AND: In loop FOR Index := 1 TO 14
    AND: Index IN AccJobPer
    PRec[]Depend[Index] = Adult

Code Index is not valid for this question.

Compute if: QAccomdat.Tenure IN [Part .. Squatting]
    AND: QAccomdat.HHStat = Shared

es_household := ('you, that is, just ' + HoHNames + ',')

Compute if: QAccomdat.Tenure IN [Part .. Squatting]
    AND: QAccomdat.HHStat = Shared

IsAre := 'Are'

Compute if: QAccomdat.Tenure IN [Part .. Squatting]
    AND: NOT (QAccomdat.HHStat = Shared)

es_household := 'es your household'
**Compute If**:

\[ Q\text{Accomdat.Tenure IN [Part .. Squatting]} \]
\[ \text{AND: NOT (QAccomdat.HHStat = Shared)} \]

\[ \text{IsAre := 'Is'} \]

**Ask If**:

\[ Q\text{Accomdat.Tenure IN [Part .. Squatting]} \]
\[ \text{AND: PTenure IN [Rents, Part]} \]

**RentDoc**

\[ Q\text{Renting} \]

Do you have a rent book, rent card, Housing Benefit statement or some other rent document that you could consult?

**IF HB STATEMENT AVAILABLE PLEASE CONSULT THIS.**

1. Housing Benefit Statement
2. Some other document
3. None

**Compute If**:

\[ Q\text{Accomdat.Tenure IN [Part .. Squatting]} \]
\[ \text{AND: PTenure IN [Rents, Part]} \]
\[ \text{AND: RentDoc IN [HBStmt, Oth]} \]

\[ \text{Consult the document := 'PLEASE CONSULT THE DOCUMENT.'} \]

**Compute If**:

\[ Q\text{Accomdat.Tenure IN [Part .. Squatting]} \]
\[ \text{AND: PTenure IN [Rents, Part]} \]
\[ \text{AND: NOT (RentDoc IN [HBStmt, Oth])} \]

\[ \text{Consult the document := ''} \]

**Ask If**:

\[ Q\text{Accomdat.Tenure IN [Part .. Squatting]} \]
\[ \text{AND: PTenure IN [Rents, Part]} \]

**Rent**

\[ Q\text{Renting} \]

How much rent does household currently pay?

**HELP <F9>**

0.00..999997.00
WARN IF: QAccomdat.Tenure IN [Part .. Squatting]
     AND: PTenure IN [Rents, Part]
     Rent = RESPONSE

This is a 'Key Question': it is VERY IMPORTANT to get an answer here if possible. If you cannot do so (either now, or later) please make a Note about the circumstances.

ASK IF: QAccomdat.Tenure IN [Part .. Squatting]
     AND: PTenure IN [Rents, Part]
     AND: Rent > 0

RentPd

QRenting

How long does this cover?

(1) One week
(2) Two weeks
(3) Three weeks
(4) Four weeks
(5) Calendar month
(7) Two Calendar months
(8) Eight times a year
(9) Nine times a year
(10) Ten times a year
(13) Three months/13 weeks
(26) Six months/26 weeks
(52) One Year/12 months/52 weeks
(90) Less than one week
(95) One off/lump sum
(97) None of these (EXPLAIN IN A NOTE)