



The Magenta Book

Guidance Notes for Policy Evaluation and Analysis

Chapter 8: Qualitative Research and Evaluation

How Do You Know Why (and How) Something Works?

Published: September 2004

Government Chief Social Researcher's Office
Prime Minister's Strategy Unit
Cabinet Office
Admiralty Arch
The Mall
London SW1A 2WH

Contact: phil.davies@cabinet-office.x.gsi.gov.uk

Telephone: 020 7272 1862



Chapter 8: Qualitative Research and Evaluation

How Do You Know Why (and How) Something Works?

8.1 Introduction – Qualitative Research: What Is It and Where Does It Come From?

Qualitative research encompasses a range of different approaches, disciplines and data collection methods. It includes research methods that capture naturally occurring data in their real-life context (principally participant and non-participant observation, documentary analysis, discourse and conversation analysis), and those that generate their own data through a reconstruction or retelling of views or behaviours (principally different forms of interviews, and focus groups).

Qualitative social research has its roots in the late nineteenth-century reaction against ‘positivist’ methods of social science and the attempts of the latter to emulate the principles and procedures of natural science. Rather than examine relationships in quantitative terms, or test hypotheses using statistical techniques, the interpretivist tradition argues that social research should provide in-depth, qualitative understanding of the subjective meanings of social life. In other words, it grew out of a concern to understand social life as it is experienced by people and as they make sense of it, rather than to derive laws or theories about it.

The natural sciences typically look for law-like patterns to social life – laws or theories which can be said to apply invariably. Interpretivist social science, on the other hand, looks to understand social phenomena from the viewpoint of the individuals and groups who experience these phenomena in specific (and not necessarily reproducible) social contexts. For interpretivists, social research is more akin to history than to natural science; i.e. it involves understanding, interpreting and explaining events, contexts and people in terms of personal and shared meanings, rather than looking for invariant laws detached from human agency and experience. Although qualitative research is usually

associated with the interpretivist tradition and quantitative research with positivism, some commentators argue that this overplays distinctions in the ways in which they are used, particularly in multidisciplinary research, and that each can contain at least some elements of both positivism and interpretivism.

8.2 Key Features of Qualitative Research

Qualitative social research is diverse. As a research approach, it draws on a number of different philosophical schools of thought and academic disciplines, particularly sociology, anthropology, philosophy, linguistics and psychology. Qualitative research has also developed a range of research methods and techniques (for example, in-depth interviews and focus groups). Qualitative methods are usually employed in *naturalistic* settings, rather than the contrived environments of controlled trials or social surveys. The goal of most qualitative social research is to capture as closely as possible the understandings, interpretations and experiences of ordinary people in their everyday lives and environments.

Qualitative social research is also used with quantitative methods, and is undertaken in less than wholly naturalistic settings (e.g. as part of a social survey or a controlled evaluation). One of the major challenges of qualitative social research is to capture and explain social life as authentically as possible without affecting, manipulating or changing it. This is often very difficult. Qualitative social research, like most other types of social research, brings to the topic that it is investigating theories, background knowledge and the researchers' own experiences and meanings, all of which influence, if not bias, the topic or subject of inquiry. This raises important methodological problems about the status of the knowledge and data gathered by qualitative social research. Questions about the validity, reliability and generalisability of qualitative data are just as important as they are for quantitative and experimental research, though these issues are generally discussed by qualitative researchers using a different vocabulary (e.g. credibility, defensibility, external validity – see Spencer *et al.*, 2004).

Given these different origins and approaches, it is unsurprising that there are few shared definitions of qualitative research. There is, however, broad agreement about its key characteristics, although the emphasis given to individual features will vary (Spencer *et al.*, 2004, drawing on Bryman, 2001; Denzin and Lincoln, 2000; Hammersley and Atkinson, 1995; Mason, 2002; Patton, 2002). Generally, the key characteristics are seen as being:

- in terms of approaches
 - a concern with *meanings*, especially the subjective meanings of participants;
 - a commitment to viewing (and sometimes explaining) phenomena *from the perspective of those being studied*;
 - an awareness and consideration of *the researcher's role and perspective*;
 - *naturalistic inquiry* in the 'real world' rather than in experimental or manipulated settings;
 - a concern with *micro*-social processes (i.e. their manifestation at the level of individuals, groups or organisations);
 - a mainly *inductive* rather than deductive analytical process (i.e. broadly, deriving theories or findings from empirical research data, rather than deducing a hypothesis *a priori* which is then tested by empirical research).
- in terms of research conduct
 - prolonged *immersion* in, or *contact* with, the research setting;
 - the absence of methodological orthodoxy and the use of a *flexible (emergent) research strategy*;
 - the use of *non-standardised, semi-structured or unstructured methods of data collection* which are sensitive to the *social context* of the study;
 - the capture of data which are *detailed, rich and complex*;
 - the collection and analysis of data that are mainly (although not exclusively) in the form of *words and images* rather than numbers.
- in terms of the use and presentation of research data
 - the setting of data in *context*;

- a commitment to retaining *diversity and complexity* in the analysis;
- a respect for the *uniqueness* of each case as well as themes and patterns across cases;
- attention paid to *categories and theories which emerge from data* rather than sole reliance on *a priori* concepts and ideas;
- *explanations offered at the level of meaning* (i.e. the individual and shared meanings that things hold for people) or in terms of local ‘causality’ (why certain interactions do or do not take place in individual cases) rather than context-free laws of general application.

8.3 Applications of Qualitative Research to Evaluation

These key features of qualitative research mean that it makes a distinctive contribution to policy evaluations, particularly because of its ability to explore issues in depth and capture diversity, its concern with context, and its focus on exploring meanings. This means that it can bring real depth to the understanding of the contexts in which policies operate and their implementation, processes and outcomes.

Personal, social, structural and environmental contexts can all be important in policy evaluation. To illustrate, in an evaluation of a welfare to work policy aimed at disabled people, qualitative research would provide understanding of how disabled people understand and experience disability, the values associated with work and social security benefits, and feelings about and experiences of working. These individual contexts would provide critical understanding of how people relate to or think about the policy or service. Qualitative research would also illuminate issues such as how employers’ attitudes or policies might act as barriers to disabled people who want to work, how the social security system can complicate or hinder the move from benefits to work, or how the culture of an organisation delivering the service shapes its approaches to implementing it. Understanding these broader aspects of the setting in which the policy works is also critical. The ability of qualitative research to illuminate contexts means it plays an important role in the development of policy. By understanding how people view an aspect of their world, and how they behave or make decisions within it, qualitative

research can help to identify possible policy approaches and strategies, and to enhance the likely success of their implementation and delivery.

In terms of implementation, qualitative research can be used to look at issues such as what and who is involved in the process of implementation, the steps and processes involved, barriers and facilitators, decisions made, whether the policy is implemented as envisaged, and the reasons for deviation from the original design.

Qualitative research also lends itself to evaluations which require an understanding of processes. It can, for example, generate a detailed description of what interventions are involved in a service or policy, who provides them, what form they take, how they are delivered, and how they are experienced by participants and by those who deliver them. It can provide an in-depth understanding of the decisions, choices and judgments involved, how they are made and what shapes this. This is particularly important where the policy or intervention is itself highly process-orientated, where the intention is to effect change through interactions (for example, between several members of staff and a client) rather than through a one-off event or input.

Qualitative research also plays a key (although sometimes neglected) role in understanding impacts and outcomes. Rather than providing quantitative measurements of gross or net impact, it can answer more detailed questions which might be summarised as *'how, under what circumstances, in what ways and for which types of people is the policy working ... and what do we mean by "working" anyway?'* It can tell us about the range and types of impacts a policy has, giving a voice to outcomes that were not anticipated or intended and which an evaluator might not have thought to consider. In this respect, qualitative research and evaluation has much in common with the theories of change and realist approaches to policy analysis (see Chapter 3).

Qualitative research can illuminate the meaning that outcomes hold, or how they are perceived and experienced, by the people involved. Building up from detailed individual accounts, it tells us how outcomes are arrived at – which elements of the programme, at

which stages, experienced in what ways, and with what relationship to changes in a client's wider life, contribute to the impacts experienced. It explains what might act as facilitators or barriers to the desired impacts, how they act in this way, and how barriers can be overcome and facilitators harnessed. And it explains how and why the impacts vary.

Example: Fathers in Sure Start Local Programmes (Lloyd *et al.*, 2003)

(www.surestart.gov.uk)

The Sure Start programme provides a good example of how qualitative research is used alongside quantitative evaluation to provide a rounded understanding of how this programme is working. In addition to a national survey of 250 Sure Start areas and 50 Sure-Start-to-be areas in order to assess the impact of the Sure Start programme, the evaluation team is also undertaking a series of themed studies and local context analyses. By using in-depth interviews, informal observations and documentary sources, qualitative research identified, for example, that fathers were a hard-to-read group for Sure Start and other community-based programmes and that their needs were not always recognised or accommodated by these government programmes. It was also able to identify successful strategies for engaging fathers in Sure Start programmes.

Qualitative research is particularly valuable in certain evaluations, for example (Davies, 2000; Patton, 2002; Popay and Williams, 1998; Williams, 1986):

- where a policy or social context is not well understood, and the evaluation questions, issues or criteria are not immediately obvious;
- where 'insider' values and perspectives are particularly important as well as the 'official' perspective;
- where diversity in how the policy operates across different sites or services needs to be understood;

It can

- provide new insights into the implementation or experience of the policy;
- check for unintended or perverse consequences of the policy or project;
- explore the complexity of what goes on, in its natural settings;
- explore ‘taken for granted’ practices, lay behaviour and organisational cultures.

Implications for government research, evaluation and policy making

Qualitative research will make an important contribution to any policy evaluation where what is required is more than a quantification of users, costs, outcomes or estimated impact. Given that qualitative research is particularly helpful in understanding *why*, *how*, and *under what conditions* policies, programmes and projects work or fail to work, it helps policy making by identifying:

- the context in which the policy operates, and what this means for the design, development and likely success of the policy;
- how the policy is delivered and experienced on the ground;
- what impacts it has and which aspects of the policy contribute to them;

and it will provide a critical explanation of why the policy works for some people, or in some circumstances, but not for, or in, others.

It is invaluable in framing policy questions in ways that are meaningful to ordinary people, and in eliciting their perceptions, understandings and experiences of policy interventions. It is probably not too grandiose to suggest that qualitative research helps enhance the democratic process of government by introducing the citizens’ (or ‘users’) perspective in rigorous, systematic and non-anecdotal ways.

8.4 Qualitative Research Methods

This section discusses the main qualitative research methods used in evaluation: in-depth interviews; focus groups and consultative and deliberative methods; participant observation and ethnography; documentary analysis; narrative and biographical approaches, discourse and conversation analysis and case studies.

8.4.1 In-depth Interviews

What are they?

In-depth interviews (also called unstructured interviews) are probably the most frequently used form of qualitative research in government evaluation (Spencer *et al.*, 2004). Personal spoken accounts are seen as having central importance in social research because of their power to illuminate meaning (Hammersley and Atkinson, 1995). Individual, personal accounts display the language that people use, the emphases they give, and allow people to give explicit explanations for their actions and decisions. The in-depth interview is sometimes described as being akin to a conversation, or a '*conversation with a purpose*' (Webb and Webb, 1932: 130). However, although a good in-depth interview will appear fluent, spontaneous and naturalistic, it is in fact a much more one-sided and carefully managed event than an everyday conversation.

In-depth interviews are a fairly lengthy, focused and usually private type of interaction. As such, they provide an opportunity to collect rich and detailed data, with the research interviewer 'mining' the subject and encouraging the participant to give more and more depth to their responses. They are ideal for an in-depth exploration of a subject which gives the researcher a detailed insight into the participant's own world, their beliefs, experiences and feelings, and the explanations they provide for their own actions or beliefs. Such interviews also lend themselves well to exploring complex processes or unpacking decision making. A good research interviewer will build a rapport with the participant which particularly aids the exploration of sensitive, painful or difficult subjects. Biographical research approaches use particular forms of in-depth interviews and are discussed in Section 8.4.7.

The degree of structure in interviews varies. A key feature of qualitative interviews is that questions are not pre-formulated: there is scope for the participant to influence the direction and coverage of the interview. In some studies, perhaps particularly where the purpose is to reveal to the researcher a social world that is unfamiliar to them, the interview may be relatively unstructured, with the researcher asking very broad questions

and the participant shaping the account. In others, the researcher will have a stronger sense of the issues that need to be explored and will play a more active role in directing coverage. The terms ‘unstructured’ and ‘semi-structured’ are sometimes used to denote different degrees to which the research directs the interview, although there is not always consistency in the ways they are applied.

How are they conducted?

The process of generating data through in-depth interviews is both systematic and flexible. It is systematic in that there needs to be careful and detailed thought initially about the type of data required and how to generate or collect it, and some consistency between interviews in the issues covered. However, data collection also needs to be flexible to reflect the uniqueness of each individual case, to explore what is of particular relevance to it, and to allow the formulation of the research questions to develop and sharpen as the study proceeds.

Some form of instrument – usually described as a topic guide, interview guide or interview schedule – is required. These act as an *aide mémoire* in the field, and help to ensure that there are no gaps in interview coverage. They are also an important aspect of accountability to those funding, commissioning or steering research, and provide a public document of an aspect of the research process that it is otherwise difficult to describe. Research reports usually show the topic guide used – for examples, see some of the studies quoted in this chapter. However, topic guides should be seen as the starting point for data collection, not as a straitjacket. They need to be used flexibly if the aim is in-depth exploration, and it is good practice to review and make any necessary amendments to the guide after the first few interviews.

Good interviewing involves asking a range of different types of questions. Wide, ‘mapping’ questions are generally used to open a research topic, but the participants’ initial responses are followed up with further questions and probes to ‘mine’ the topic in depth. A variety of follow-up questions will be needed for each topic to amplify, clarify, and seek explanation, and to stimulate, challenge or encourage the participant to review

the topic from other perspectives. For example, a ‘mapping’ question might be along the lines of ‘How did you come to decide to use the service?’ The respondent might reply that they thought it sounded as if it might be helpful. ‘Mining’ questions would then be asked to find out in what way the respondent thought it might be helpful, what they knew about it, and what sort of help they wanted from it, with probes such as ‘why was that?’, ‘in what way?’, ‘how did you know that?’ and so on used to ensure each issue is explored in depth. Although broad and open questions are critical, good in-depth interviewing also involves a high level of specificity to get an in-depth understanding of the participants’ attitudes, experiences or behaviour. What is always important is that questions are clear and not leading: questions are formulated as neutrally as possible so as not to imply that a particular response is expected or required.

In-depth interviews are usually tape-recorded. Note-taking is not sufficient to capture data in enough depth: the nuances and details which constitute the richness of interview data are lost, and the effort required in keeping pace with the respondent impedes the interview interaction. There is a potential disadvantage that being tape-recorded changes what a participant says, or makes the interaction more formal and charged. In practice, people appear quickly to get used to the presence of the tape-recorder – they seem surprised, for example, when the machine clicks off after one side of tape. Using field notes to record what is heard or seen outside the main research interaction can also be useful.

Implications for government research, evaluation and policy making

In-depth interviews are a very rich resource for researchers and policy makers. They provide detailed personal accounts of beliefs or experiences. They display the language, constructions and emphases of people in ways that are very revealing. And they allow participants to give explicit explanations for their views, decisions or actions, describing what has shaped them.

There are some things that interviews cannot do. The emphasis is on the participant’s own interpretation, so they may not be the optimal research method where people are

unable or unwilling to give an open account of themselves, or where they find introspection and self-questioning particularly difficult. Where it is difficult for people to reconstruct and retell something or where they are less likely to do so honestly or accurately, other methods of data collection, such as observation, would be more appropriate. There will also be limitations on the extent to which people are able to pinpoint what has led to a particular behaviour, or what difference an intervention has made, or what would have happened in its absence. Asking people about their *perceptions* of these things will always be important and illuminating, but their perceptions may not always fully reflect what would have happened if they had, or had not, experienced the policy. Surveys collecting quantifiable data and using experimental designs such as control groups or randomised control trials are therefore needed to give an accurate measurement of net impacts, although qualitative research works effectively alongside such methods to provide explanations for the impacts found.

But interviews have many uses in evaluations. For example, they can be used to find out:

- how a service was set up and delivered by the management team;
- how staff work, the decisions and judgments they make, what they think works well and less well;
- how clients experience a service;
- within the detailed context of an individual case, what impact a service has, and how.

It can be difficult to assess whether the interview process is being well conducted. Ultimately, the proof is in the pudding – whether the data generated is rich, in-depth and illuminating. But research managers and policy makers should expect to see:

- a topic guide or interview schedule which comprehensively covers the ground but clearly allows for flexibility;
- reflective preparation for fieldwork by the team;
- tape-recording and verbatim transcriptions of interviews;

- an account, if requested, of how the interviews were conducted, what emerged spontaneously, how issues were raised and pursued by the research team;
- evidence of depth and insight in the data collected, demonstrated in the way it is analysed and presented.

8.4.2 Focus Groups

What are they?

Focus groups or group discussions usually consist of around four to eight people, who may be acquainted with each other or may be strangers, brought together to discuss a particular topic or set of topics. The discussions typically last around an hour and a half, although this is certainly not fixed (see below). The group is moderated or facilitated by a researcher. Although focus groups have acquired a somewhat dubious image, they are a well-established and rigorous method of social research and evaluation.

In focus groups, data are shaped and refined through the group interaction. Hearing from other participants stimulates further thought, encouraging people to reflect on their own views or behaviour and triggering further material. Focus groups are synergistic (Stewart and Shamdasi, 1990) in the sense that the group works together, and the group forum is used explicitly to generate data and insights (Morgan, 1997). They also provide a strong social context to the discussion. This may be a natural social context if those in the group already know each other (for example, colleagues). But even if the group members are strangers brought together for the research, there will be more spontaneity than in an individual interview. People's social frames of reference will be more on display, there will be insights into how ideas and language are shaped by the social context, and social constructions – normative influences, collective as well as individual self-identities and shared meanings – will be illuminated (Bloor et al, 2001; Finch and Lewis, 2003; Krueger and Casey, 2000).

Focus groups have an application in any study where what is sought is refined and reflective discussion, or the social context made visible. The data they generate is in depth, not at the individual level as with interviews, but because it is the result of

listening, and thinking further. They provide opportunities for creative thinking, for projective or enabling techniques, for group work and for giving information, for example, on technical subjects. They can work very well in tackling abstract or conceptual topics, whereas on a one-to-one basis a participant may 'dry up'. They can also be used for sensitive subjects, provided there is enough similarity between participants in their social characteristics and their connection with the research subject to create an environment that feels safe.

Focus groups work well in combination with interviews or other research methods. For example, at the beginning of a study they can be used to map out the territory, to give early insight into how people approach, discuss and construct a subject. At the end of a study, they offer a deliberative forum for refining understanding of an underlying theme, exploring causes or origins of problems, examining implications for service delivery or policy development, or generating or prioritising solutions.

Example: Attitudes and Aspirations of Older People (Hayden *et al.*, 1999)

<http://www.dwp.gov.uk/asd/asd5/102summ.asp>

This study formed part of a programme of research on older people by the Department of Work and Pensions. The qualitative study was preceded by a literature review on attitudes and aspirations of older people.

This research used 15 focus groups of between six and eight older people aged 50 and above across the UK. 20 follow-up in-depth interviews were conducted to develop case study data. The aim was to explore the views older people have about their lives and their expectations and wishes for the future and to improve the understanding of the factors influencing the attitudes and aspirations of older people.

The focus groups were designed to cover diversity in gender, ethnicity, dependency and location. The focus groups explored themes around employment, work and volunteering; around retirement and activities in retirement; on issues of health and social care; and on services, consultation and citizenship. The research confirmed that attitudes and views of

older people are influenced by a variety of socio-economic factors and different life experiences. It found considerable agreement among all participants in their desire to be as active and independent as possible and identified a number of barriers to active ageing.

How are they conducted?

The role of the researcher or facilitator is key to making focus groups effective (Finch and Lewis, 2003). As with in-depth interviews, the amount of structure will vary. The researchers will, at least once the group is underway, want as much as possible of the discussion to emerge from the group itself. But they will also help the group to focus and structure their discussion, bring discussion back or move it on, widen the discussion to include everyone, and ensure a balance between participants. They guide and pace the discussion to ensure all the issues are covered, and they probe individuals and the group as a whole to encourage in-depth exploration. They are also alert to non-verbal language and to the dynamic of the discussion, and they need to challenge or stimulate the group if what is said seems too readily to reflect social norms or apparent consensus.

Careful consideration needs to be paid to the composition of a focus group. This includes how many focus groups need to be convened to cover an issue adequately (see Sections 8.5 and 8.6), and which combination of individuals in each focus group will work best. For example, it may be difficult for individuals in an organisation to express their views openly in a focus group setting if more senior members of staff from the organisation are also in the group. Similarly, it can be difficult to discuss certain topics in mixed gender groups or mixed age groups.

Practical arrangements are also important – the location and timing of the group's meeting can make an enormous difference to who is willing to attend, and the research team needs to take steps to make participation easy for people and to encourage and reassure them about attending.

Implications for government research, evaluation and policy making

Focus groups are very useful for any evaluation question where deliberation, discussion and the stimulation of other people are likely to bring depth to the topic. They are not appropriate where the presence of others is likely to inhibit or constrain people, or influence them in ways that make their accounts less open or less reliable. They provide much less depth at the individual level, and so are less appropriate where the personal context is critical, or for very detailed or complex subject matter. As with in-depth interviews, the emphasis is on the person's own interpretation, which may not always be what is required. But they are ideal where hearing what others say helps people to sharpen their understanding and articulation of their own views or experiences – for topics that are a little abstract or not at the front of people's consciousness, for example, for illuminating different views or beliefs, or for highlighting variation in behaviours or decision making.

As with interviews, it can be difficult to assess the conduct of fieldwork, but the research manager or policy maker can expect the research team to be able to describe the approach they have taken, or plan to take, to carrying out the groups. They should be alert to signs in the presentation of the data that the group dynamic inhibited open discussion and should expect to see diversity, depth – and the unexpected.

8.4.3 Consultative and Deliberative Methods

What are they?

Since the latter part of the twentieth century there has been a considerable expansion in the range of methods used for consultative purposes, particularly within local government. The boundaries between consultation and qualitative research are not absolutely clear cut, and some consultation methods involve the application of established research methods in a more dialogic and deliberative process. Lowndes *et al.* (1998) found an increasing repertoire of methods being used in local government, with a growing interest in innovative methods as well as, or in place of, traditional ones. They list 19 forms of public participation used, including meetings, committees and forums;

interactive websites; surveys, polls and referenda; and citizens' panels and citizens' juries. More innovative forms of consultation are:

- the Delphi method (Adler and Ziglio, 1996; Cantrill *et al.*, 1996; Critcher and Gladstone, 1998). This is an iterative process particularly aimed at forecasting, in which a group of experts are asked to respond individually to questions (either in a survey or using qualitative research). Initial responses are then circulated among panel members who are asked to review their own responses, with further rounds of circulating and refining responses until consensus, or agreed difference, is reached. The group does not meet physically.
- the Nominal Group Technique. A variant of the Delphi method, this follows a similar pattern to the Delphi method in terms of eliciting initial responses from panel members. After the first round of elicitations, however, further iterations are conducted using an interactive group format somewhat similar to a focus group. The aim of the group is to reach a consensus in areas of both agreement and disagreement. A good facilitator is vital to ensure that all participants in the group are able to contribute to the discussion, and that the group works in a disciplined manner to reach agreement and agreed difference.
- citizens' juries (Coote and Lenaghan, 1997; Davies *et al.*, 1998; White *et al.*, 1999). A group of between 12 and 20 people are brought together over the course of several days to hear from and put questions to expert 'witnesses', deliberate and discuss among themselves, and make recommendations about courses of action, which may or may not be consensual.
- deliberative polls (Fishkin, 1995; Park *et al.*, 1999). This is a set of activities with a focus on exploring how public views change when the public has an opportunity to become well informed on an issue. A survey is conducted to 'benchmark' public opinion. Participants attend a joint event, usually over a weekend, which involves discussion in small groups, expert plenary sessions and political sessions in which party spokespeople respond to questions. The survey is repeated to measure the direction and level of change in views.

- consensus conferences or workshops (Seargeant and Steele, 1998). These follow a model developed in Denmark in which a panel of around 16 people work together to define the questions they wish to address within a particular subject, question experts, receive information, deliberate, and aim to reach consensus. The panel produces its own report, which is presented to an open conference and discussed further.
- ‘participatory appraisal’. This has historically been used in overseas development work but is now more common in the UK. It is designed to involve people, particularly from socially excluded communities, in decisions that affect their lives. It combines a variety of visual (mapping) tools with group discussions and semi-structured interviews (Oxfam UK, 2003).
- ‘planning for real’. This is a community consultation process where models are used to encourage residents to explore and prioritise options for action (Gibson, 1998).

Implications for government research, evaluation and policy making

There is a wide range of consultation methods available to the government researcher and policy maker. This raises the question of when different methods should be used, and how the choice between different methods should be made.

Consultation methods have the potential to enhance policy formulation, evaluation and development, as well as to contribute to cultural and personal change within the organisations using them. Researchers and policy makers will want to look to consultation methods, rather than research methods, when they want to move beyond exploring people’s views and behaviours, to getting them to come up with, or to appraise, solutions and strategies. The group orientation of most consultation methods means that they are particularly useful when the issues involved are technical, complex and require consensus, and where additional information needs to be fed into the process of deliberation.

Effective consultation requires:

- clarity of purpose;
- clarity about who should be consulted and what they are expected to contribute;
- organisational capacity and skills;
- careful selection of a consultation method which fits the question and is feasible given time and resources;
- careful planning and management;
- the desire and competence to make use of the outputs;
- evaluation of the method used (Audit Commission, 2003; Lowndes *et al.*, 1998 Seargeant and Steele, 1998).

Consultation generally involves intensive exercises with relatively small groups, and thus raises questions about value for money and representativeness. Well-conducted consultation will help to highlight and explain areas of difference, as well as agreement, among participants. A careful balance needs to be struck between the need for consultation to point to an agreed way forward and the danger that it produces an artificial consensus.

8.4.4 Participant Observation and Ethnography

What are they?

One of the major ways in which social research can understand an activity, group or process is to get as close as possible to them without disturbing its 'natural' operations (Hammersley and Atkinson, 1995). This can be done at one extreme by being a wholly detached observer of a social situation, working as unobtrusively as possible, making observations, listening, and remembering details. At the other extreme, one can join the group or activity in question and participate in it as a member in order to learn about it from the inside out. This may or may not involve 'going native', i.e. becoming so closely involved in the group, activity or processes that one loses one's detachment and outsider status. Clearly, there are positions mid-way between these extremes; one can function as an observer-participant or as a participant-observer, the difference being the degree of detachment and involvement that is possible for the social researcher (Figure 8.1)

Figure 8.1

The observer–participant spectrum



Social research has a fine tradition of using participant-observation/observer-participation to better understand gangs, criminal behaviour, drug use/misuse, school participation and achievement, health and illness behaviour and many other substantive topics. Two of the ‘classics’ of participant observation are W.F. Whyte’s (1955) *Street Corner Society* and Elliot Liebow’s (1967) *Tally’s Corner*. Erving Goffman’s work, *Asylums* (1961), *Stigma* (1968), *Presentation of Self in Everyday Life* (1959) and *Behaviour and Public Places* (1963) stands in the pantheon of participant-observational research.

Ethnography

Ethnography is a method used by anthropologists which has been adopted by social researchers more generally. The term ‘ethnography’ means the description (‘*graphy*’) of a people or culture (‘*ethno*’). More precisely, then, ethnography is the *detailed description of a culture, group or society, and of the social rules, mores and patterns around which that culture, group or society are based*. Ethnographic inquiry is able to elicit the *cultural knowledge* of a group or society; i.e. what one needs to know to be a fully competent member of that society. Ethnography also involves the detailed investigation of the *patterns of interaction* of a group or society in order to understand the values, processes and structures (including language) of that group.

As with other types of qualitative social research, ethnography studies social groups and social activity in as ‘natural’ a way as is possible, i.e. without trying to affect or manipulate them in any way. Ethnography also shares with other types of social research an interest in capturing the ‘first-order’ meanings of the people it is studying, and their relationship to the ‘second-order’ categories and constructs that are used by researchers. Ethnography is a social research method that allows researchers to ‘get inside’ a group, culture or society and understand it from the inside out (Chambers, 2000).

How are they conducted?

Observation, listening, remembering and detailed note-taking are key techniques for social researchers using participant-observation and ethnographic methods of inquiry. A good participant-observer or ethnographer will learn to observe what is going on in a group, or a social situation, and to identify verbal and non-verbal communication, patterns of interaction, people's responses to verbal and non-verbal activity, body language, people's demeanour and deference to others, status hierarchies, and the like.

There is some debate among participant-observers and ethnographers as to whether one should take notes and make recordings of what is observed *in situ*, or whether one should refrain from doing so until the activity being observed has finished. If notes are not taken during observation, the observer/participant will make detailed field notes and recordings of what they have observed as soon as possible after the observation has been concluded, usually in the privacy of a private room or study.

In the past two decades or so, audio- and video-recording has become more readily available to social researchers, and is becoming less intrusive and more discrete. This has allowed participant-observers and ethnographers to collect audio- and video-taped data on everyday social settings such as classrooms, doctors' surgeries, courtrooms, and office life. Audio- and video-taped data has the advantage of allowing extensive (some say exhaustive) analysis of naturally occurring social activity, and for these activities to be retrieved for further analysis, verification and challenge by others. This can enhance the reliability and validity of qualitative analysis, and provide more transparency of qualitative social research.

Against these advantages is the potential disadvantage that audio- and video-recording disturbs the 'natural' social activity one is trying to capture and understand. To some extent this can be minimised over time as people's awareness of the presence of audio microphones or video cameras usually diminishes as they get used to the presence of these devices. In the short term, however, and possibly beyond, there is the possibility of

some Hawthorne¹ effect of openly using audio- or video-recording equipment. An alternative to openly using audio- or video-recording devices would be to use them surreptitiously, without the awareness or consent of the people being studied, but this is generally considered unprofessional and unethical. There may, however, be scope to use recordings which have been made as part of everyday professional practice (e.g. in police interrogation rooms, or some psychiatric consulting rooms) if permission to use them for research purposes is subsequently sought.

Triangulation (see Section 8.5.2) is an important analytical technique of participant observation and ethnography.

Implications for government research, evaluation and policy making

Government researchers and policy makers may well ask what all this has to do with policy making, policy implementation and policy evaluation. There are a number of ways in which participant-observation and ethnographic inquiry can help develop, implement and evaluate policy. They can:

- Provide robust evidence on the processes by which front-line agencies work and how these might operate to promote successful implementation, or militate against successful implementation.
- Provide a ‘window’ through which the dynamics and decision-making processes of organisations can be observed first-hand.
- Identify variations in the social and cultural environment within which policies, programmes and projects are expected to work.
- Identify real-life drivers of policy success and failure.
- Identify real-life pinch points, barriers or vulnerability factors which might undermine the successful delivery of policies, programmes and projects.

¹ Scriven (1991, p. 186) defines Hawthorne effects as: 'The tendency of a person or group being investigated, or experimented on, or evaluated, to react positively or negatively to the fact that they are being investigated/evaluated, and hence to perform better (or worse) than they would in the absence of the investigation, thereby making it difficult to identify any effects due to the treatment itself.'

- Identify key personnel who might operate as ‘product champions’ for policies, programmes and projects.
- Identify key personnel, groups and values that might stand in the way of successful policy implementation and delivery.
- Identify process and outcome measures that have external and ecological validity (i.e. are meaningful and relevant to the communities that Government serves).
- Identify ways in which different policy-evaluation methods might be possible or unlikely to succeed.

8.4.5 Documentary research

What is it?

Most evaluations are likely to use documents as part of an evaluation. In some instances, documents are the source for ‘reconstructing’ a baseline for evaluation, recording what a project intended to achieve, the conditions it developed from, how the programme developed, what changes were made and how they were implemented. For example, documents can help us understand how professionals present themselves to service users, how they establish ‘authority’, and how this might make it difficult for a service user to make the most effective use of the service. In other instances, the documentary analysis accompanies other data collection methods, for example, in-depth interviews or surveys.

Most social researchers would agree that documentary data, like all other forms of data, are socially produced, that is, they are produced on ‘the basis of certain ideas, theories, ... principles,’ and written for specific purposes and audiences, which in turn shaped their content and form (Macdonald and Tipton, 1993, p.188). Documents and records are never simple facts, but are mediated by the social context in which they were produced. Any documentary analysis needs to take this on board. The distinguishing feature of documents is that they have a historical dimension – that they are separate from the ‘author’, the ‘producer’ and ‘user’ by space and time. This may pose a specific challenge to the users of evaluation research, because there may be no running commentary available to provide additional information about the production and intentions of the documents. However, it means that documents can provide another ‘reading’ or

perspective on an event or process, one that does not rely on a retelling or narration of it, with all the biases or inaccuracies that might bring.

Documents used in research and evaluation include public records, for example, legislation, parliamentary papers, administrative and historical records, annual and financial reports, minutes of meetings, strategy plans, policy papers; private papers, such as letters, diaries and notebooks; and other, non-text documents such as photographs, maps and plans, and even buildings (Mason, 2002).² They can also include media sources, such as newspapers, magazines and leaflets. As new technologies are proliferating, new forms of public and private documents are developing. They include video and DVD records, records resulting from the use of new web technology (Internet discussion sites, interactive websites, etc.), all of which have the potential to make a valuable contribution to evaluation research. This section will limit itself to discussing text-based sources.

One way of distinguishing documents is by asking who produced them, why and for whom. Ricoeur points out that all texts are written 'to do something' (Ricoeur, 1996). Another approach involves analysing the 'type' of record under investigation. Lincoln and Guba (1985) distinguish between 'records' where formal transactions are documented (for example, driving licenses document the passing of a driving test), and 'documents', which are produced for more personal reasons, for example, memoranda or notes. Transactions and records offer different types of information. Others have categorised textual documents according to their closeness to speech. E-mails used in an organisation, for example, would be closer to speech than Acts of Parliament, which are stylised in form and content and utilise a 'full state technology of power' (Hodder, 2000, p. 703).

² The use of documents in social research and evaluation has a long tradition, going back to the earliest days of social science research. Founders of social science methods, such as Max Weber in his study of bureaucracy and Emile Durkheim, in his study on suicide are based on the analysis of public documents (Macdonald and Tipton, 1993). And the Chicago School used documentary analysis in some of its classic studies (Hammersley and Atkinson, 1995).

Evaluators are not restricted to using existing documents; documents can be produced as part of an evaluation exercise. For example, documentary evidence can be solicited by asking participants to keep diaries in order to produce time series of events. Still, the same questions about the conditions of its production apply, albeit in the context of the evaluation itself.

How are they used?

Documentary analysis can make a contribution to both formative and summative evaluation. In summative evaluations, documentary analysis can provide evidence on, for example, whether a policy has achieved its objectives and what the changes are. In formative evaluation, documents can provide evidence on the process for instance, of how a new service is developing over time; highlighting where barriers continue to be encountered. Other uses include the construction of different/alternative perspectives to those produced through other data source. Hammersley and Atkinson, for example, talk of documentary analysis as ‘giving voice’ to muted and suppressed groups, which might otherwise not be heard in the course of the evaluation exercise (1995). Two ways in which documentary research is used:

- as an independent dataset
 - Documents can be a historical ‘audit’, providing or supplementing information that would otherwise not be traceable; for example, it could be used to fill gaps when central stakeholders or staff cannot be interviewed as part of the evaluation, either because there are issues of confidentiality or because they have left a service.
 - Documentary analysis can track processes over a period of time (for example, the way a policy process emerged over time).
 - Documents can ‘unlock’ the otherwise ‘hidden’ history of a process or organisation and fill in ‘gaps’ in personal memories of interviewees and in what they are be prepared to say.
- as part of a larger research design

- Documentary analysis can inform the design of research questions and other data collection methods (for example: carefully chosen documents could be used in focus groups to explore an issue further; findings of documentary analysis can inform the questionnaires and topic guides; a written record of the policy objectives can be used to further probe interview participants or focus groups; a leaflet explaining the key components of a service could suggest additional lines of inquiry in a survey).
- Documentary analysis could be used to look at differences between documents and between documents and interview accounts in order to explain how differences in perceptions of a service can arise.
- It can be used for the triangulation (see Section 8.5.2).

Example: Hospital Trust Mergers in London (Fulop *et al.*, 2002)

<http://bmj.bmjournals.com/cgi/content/abridged/325/7358/246>

This research project studied the processes and outcomes of mergers and reconfigurations of hospital trusts in London between April 1998 and 1999. The study employed a mixed-method approach. It included a management cost analysis, a cross sectional study of nine mergers between 1998 and 1999, and six case studies of trust mergers at two intervals. The findings of the study assessed the outcomes of hospital merger in terms of service delivery; whether trust size matters; on management structures and organisational culture; and on staffing and financial issues.

Documentary analysis formed an important dimension of the research:

In the cross sectional study, public consultation documents, outlining the reasons for the proposed mergers, were analysed to identify drivers informing the merger plans and the favoured organisational structure. Drivers identified included the need to make savings in management costs and invest them into patient services, to safeguard and develop specialist clinical services and units and to address staff retention and advancement issues.

Findings from the documentary analysis were used to inform interviews with key informants in the cross sectional study and in the case studies. For example, in the interviews with key informants views on the relative importance of stated drivers in the merger process were explored; informants were asked about additional reasons for the merger (unstated drivers) and whether objectives had been achieved or were likely to be achieved.

Different techniques in documentary analysis have been developed and continue to develop, drawing on different disciplines and expertise, for example, on literary, reflexive and interpretative techniques (Mason, 2002). They include:

- *Theoretical analysis*, where documents are studied to as to whether they support a pre-defined theory or explanation. For example, are the stakeholders' theories about a chain of events or change in a service reflected in the paper trail of memos, reports, etc.? Do the documents reveal other factors that need further exploration?
- *Structural analysis*, where the structure of the document is studied in detail, i.e. how it is constructed; which context it is set in; how it conveys its messages. For example, what can the structure, content and format of a particular health information leaflet tell us about the type of message, the target group and the expectations about behaviour change? Does this match what interviewees have told us about these issues in interviews? What do users of the service make of the leaflet?
- *Content analysis*, where the information in particular types of documents is studied and compared. For example, evaluators could compare tender, contract and training documents to 'chart' the changes in thinking about the service from its initial inception to its implementation and to identify significant turning points in the process of developing the service. This could then be used to inform the questions for stakeholder interviews, develop the sample of stakeholders and identify further areas of investigation.

Less technically, documents may be analysed to bring context to the evaluation and to triangulate findings with other data sources (see Section 8.52). It would be advisable to

seek expert advice on the specific requirements and skills for each technique that is to be used.

Implications for government research, evaluation and policy making

Being clear about what the document can and cannot deliver, what role it will have in the evaluation and how the claims made about a document can be substantiated will have a bearing on the inclusion of documentary analysis in the evaluation. Studies that use documentary analysis should have a clear description of:

- which documents were used and why;
- how they were analysed; i.e. the study should present an analysis framework for the documentary analysis;
- how the documentary analysis relates to other aspects of the evaluation.

One of the problems for including documentary analysis in the research is access to documents. Not all documents that are required may be readily available – many government documents are restricted; documents may be lost. The use of private documents obviously raises issues of confidentiality.

8.4.6 Conversation Analysis and Discourse Analysis

What are they?

Conversation analysis and discourse analysis place a particular emphasis on talk and text as data sources. They focus on how the spoken and written word are used in everyday settings and how they structure communication and understanding. Discourse analysis examines the ways in which social events are reported and represented by people with different interests, and it uses this to identify the underlying values and agendas of different social groups.

How are they used?

Conversation analysis studies naturally occurring talk in order to identify the underlying rules of social interaction and what makes everyday social life ordered.

This requires data-gathering methods using audio- and video-recording, detailed note-taking and textual sources such as official documents, policy statements, briefings and media articles. These ‘naturally occurring’ data are then subjected to detailed analysis and interpretation using turn-taking analysis (the detailed analysis of a dialogue between two or more speakers), speech-act analysis (the study of meanings, structure and symbolism embedded in speech and dialogue) and other sociolinguistic methods.

Implications for government research, evaluation and policy making

While conversation analysis and discourse analysis may not initially be seen as a method of policy evaluation, they have been used by social researchers to study the ways in which politicians speak and the rhetorical devices they use to influence people’s understanding of politics, policy and society (Atkinson, 1984; Brown, 1987; Wilson, 1990; Schaffner, 1997). They have also been used to help identify implicit assumptions and agendas of policy makers and of policy statements. Such approaches may have much to offer qualitative methods of policy evaluation, especially at a time of wide suspicion and concern about political ‘spin’ and the manipulation of social and political reality. They are important in understanding how professionals/staff interact with service users. For example, they can be used in looking at how work-focused interviews are conducted – who raises the issue of work under what circumstances, how far different staff challenge people’s stated reasons for not being able to work, which ways of doing this are most effective/least threatening. Robert Dingwall (1997) has used it to look at how far mediators in family breakdown really are unbiased or neutral.

8.4.7 Biographical Approaches, Life Histories and Narratives

What are they?

This group of approaches involves collecting detailed descriptions of a person’s whole life, or of a key event or period within it. They use a particularly intensive form of data collection and analysis. For example, researchers will often go back to an informant repeatedly to collect further information or conduct a series of interviews. Or they may follow an individual over a period of time, collecting detailed data on developments in

their lives. They may study, for example, the perspective of a family by interviewing different family members, or the members of specific communities. The approaches have been utilised, for example:

- to study the recovery from alcohol and drug dependency;
- to collect (auto)biographies in mining communities;
- to record narratives of organisations and of historical events and periods;
- to give (authentic) voice to and contextualise the experience of homelessness.

Biographical, life history and narrative approaches draw from a range of interpretist schools in the social sciences. At their core lies an understanding of social actors as participants in their social environment. People are perceived as being able to make sense of and convey experiences through their accounts.³ Personal accounts and shared narratives are a source for tapping into social patterns, rules, experiences and for exploring individual and collective meanings – they are another way into ‘reading social, cultural and economic history’ (Mason, 2002, p. 57). Biographical and narrative approaches are also able to engage with sensitive and emotional experiences.

How are they conducted?

Biographical, life history and narrative research can use accounts generated specifically for the research project, but existing accounts, such as verbal accounts, stories, written biographies and video accounts have also been used.

Techniques are based on in-depth interviewing (see Section 8.4.1). Some approaches have adapted methods used in psychotherapy and counselling to encourage people to speak about their experiences and lives (Wengraf, 2001). Some researchers will return repeatedly to informants to attempt to understand a specific narrative – turning the interview process into a dialogue between researcher and researched. Some approaches perceive of the interview process as following a dynamic process of narration and

memory, which can be guided by the interviewer through the use of specific interactive techniques (Wengraf, 2001). In this version of narrative interviewing, highly structured techniques of analysis are employed. More indirect ways of collecting data include, for example, storytelling, and the keeping of diaries and other records by informants. (for an example see also Schuller et al., cited on p.48 of this chapter)

Implications for government research, evaluation and policy making

Currently, biographical and narrative approaches seem to be underused in policy research. This is somewhat surprising, given that they are a unique source for gaining access to the processes and impacts of interventions at an individual level. Biographical and life history approaches can personalise and bring to life consequences of interventions. Narratives and biographies are also a means of fixing otherwise elusive processes, such as cultural changes. For example, narratives and biographical approaches contribute to the evaluation by:

- illustrating the personal consequences of an intervention;
- tracking the longer-term and complex outcomes of change at individual and group level;
- integrating the personal and emotional with the social and structural dimensions of peoples lives, values and perspectives;
- providing a voice to otherwise neglected perspectives and groups (Shaw, 1999).

For example, asking a disabled person to narrate and describe key events and periods of their life would help others understand to what extent previous and current interventions have shaped their life chances, the views and perspectives. They could help in examining the ‘fit’ between an intervention and the material, personal and emotional circumstances of a person’s life, and also, through a personal ‘narrative’, the impact of an intervention could be exemplified. Biographical, life history and narrative approaches could make a critical contribution to an understanding of a fluid process such as social exclusion,

³ Narrative and biographical approaches have diverging views about how accounts relate to the actual lived experience and to what degree social actors understand their own complex life actions. For a discussion of

because it would help to identify the periods in someone's life when they felt more or less included/excluded and to understand the structural and personal circumstances which were relevant at those times and which might have contributed.

The approaches can be used in a variety of ways, for example:

- as illustrative case studies, i.e. thoroughly researched 'individual accounts' showing the longer-term impact of an intervention;
- as 'histories' of individuals and groups;
- as particular themes within the evaluation;
- as test cases for findings reached.

The analysis and writing up of live histories, narratives and biographies tend to be time-consuming and complex. There is usually a great deal of redrafting, working with research participants and other researchers necessary to develop the material to its full potential. This in turn requires a high degree of specialist skill and experience on the part of the analysts.

8.5 Design

Having described the key data collection methods used in qualitative research, we now go on to look at important aspects of the design, sampling, analysis and presentation of qualitative research.

8.5.1 Key principles

The key principles of qualitative research design share much with those relevant to quantitative research. In either case, there must be a clear set of research questions that are sufficiently specific to be actionable in the research. Research questions that are broad or vague can easily lead to unsatisfactory studies that simply do not produce new insights. More time spent early on sharpening up what needs to be known will help to prevent this. However, additional questions often arise during the research process, which

this, see Wengraf, 2001.

means the initial questions may need to be revisited and amended in the light of these developments.

There needs to be coherence between the research questions and the populations studied. These should be the populations that are going to give the most direct and insightful information on the subject matter. There needs to be some thought as to which subsets of these populations are particularly critical to include, or which should be excluded. For example, does the evaluation require information from policy developers or from staff delivering the service and, if so, in what roles? From their managers? If so, at what levels? From stakeholders, from current participants, eligible non-participants, past participants, etc.? In each case, which types of members of these groups are likely to have different but important things to say?

Although it is unusual to have formal 'control' groups in qualitative research designs (since these are usually associated with the measurement of difference), building comparison into research designs can be very helpful and can lead to more in-depth understanding. So, for example, a study looking at a particular phenomenon among lone parents (such as attitudes to work) might be enhanced by including couple parents. Comparing the responses of the two groups will help with understanding what is a function of being a lone parent, as opposed to just being a parent.

There should also be coherence between the research questions and the settings studied. Qualitative fieldwork is usually focused on a small number of different sites or areas, carefully selected to provide coverage of different types of contexts. These might be different institutions, organisations, labour markets, or types of urban or rural setting, for example.

There should also be a logic between the research questions and the data collection methods used. For instance, are naturally occurring data needed because, say, what is being researched is best illuminated by observing behaviour or interaction, or by reviewing existing documents? Or is what is sought people's own accounts, explanations

and interpretations, in which case should this be through individual interviews or group discussion?

And there needs to be a logic to the timing of episodes of data collection. This involves some careful thought about which perspective, or perspectives, on what is being researched will be most illuminating. Is it the early delivery of the service that should be researched, or later stages when delivery incorporates lessons learnt? Is it the experiences of participants as they first contact the service, or as they use it, or after they have stopped using it, that are needed? In practice, a thorough evaluation design will often use more than one point of data collection and will involve both repeat cross-sectional and panel designs:

- Repeat cross-sectional design is where different people are selected at different points in time. For example, one set of staff might be interviewed when implementation is being planned, another set in the early days when the service is ‘bedding down’, and a new set of people in later stages when relationships have been formed and ways of working established.
- In a panel design, the same people are interviewed more than once. For example, clients might be interviewed in the early days of their use of the service, later while they are still in contact, and then some time later when they have left the service but when more than the immediate impacts have been experienced.

The feasibility and appropriateness of a proposed design and approach within the actual research setting – the service or policy being evaluated – is of critical importance. Where what is being researched is relatively new or has not been the subject of much scrutiny, early field visits and discussions with those involved in the design, delivery and use of a service can be invaluable in informing the research design.

A further issue of key importance in evaluation is to give early thought to the criteria against which a policy or service is to be evaluated, and how these are to be generated. Qualitative research provides an opportunity to generate evaluative criteria from the

viewpoint of clients, staff or stakeholders, as well as using criteria which are established *a priori* by policy makers.

One of the key advantages of qualitative research is that it is flexible and can, more easily than quantitative research, be adapted. As more is learnt about the research phenomenon and setting, it is not uncommon for there to be a need to modify the study sample, the research questions being asked and how they are formulated in the study, or the fieldwork approaches. Changes might also be made because the information requirements of Government change – new issues are thrown up by what is learnt early on, or by changes in the broader policy context. Although this is not an excuse for poor planning, the process of design in qualitative evaluation is a continuous one and the scope for modification an advantage.

8.5.2 Triangulation

A good research design will often include elements of triangulation. Triangulation means bringing together different types of data, or sometimes different ways of looking at data, to answer the research questions.

Denzin (1989) describes four types of triangulation: *methodological triangulation*, *data triangulation*, *investigator triangulation* and *theory triangulation*, of which the first two are most commonly used in government evaluations.

- Methodological triangulation means combining different research methods. Denzin distinguishes between ‘within method’ triangulation (where, for example, a range of different lines of questioning might be used to approach the same issue), and ‘between method’ triangulation (where different data collection methods are combined). For example, observation of client–staff interactions might be combined with analysis of documents (e.g. case notes or guidance notes), or with interviews (e.g. to ask the participants about their experience of the interaction), or with group discussions (e.g. to ask staff how they vary their approach in different cases). Since

each of these methods has its own strengths and weaknesses, combining them provides a more rounded picture of what is being researched.

Methodological triangulation might also involve combining qualitative with quantitative data. Some people would criticise this approach on the grounds that the fundamental philosophical beliefs that underpin the two methods are so different that they can never be combined meaningfully. It is certainly true that qualitative and quantitative research address different research questions and generate very different types of data. However, many evaluation designs combine the methods to very useful effect. Qualitative research can be used:

- *before* quantitative research, for example, to map and clarify issues for coverage, generate hypotheses and define terminology and concepts;
- *concurrently*, to explore in depth issues measured by quantitative research;
- *after*, to provide further explanation, particularly of unexpected findings for research among key subgroups identified by quantitative research.

For example, evaluation designs typically use surveys to measure levels of satisfaction with the service and outcomes for different groups, and qualitative research to look at how the service is delivered and experienced, and to understand what works for individual people in different personal circumstances.

- Data triangulation means combining data from more than one source, say, from a number of settings, points in time or types or groups of people. For example, a range of different service delivery locations might be selected or the study might be conducted with staff, clients and non-users of the service to explore differences in their experiences and perceptions.
- Investigator or analyst triangulation is slightly different. It involves more than one researcher looking at the data so that they can either check or challenge each other's interpretation or deliberately approach the data from different angles.
- Theory triangulation means looking at the data from different theoretical positions in order to explore the fit of different theories to the data and to understand how looking at the data from different assumptions affects how it is interpreted. In the evaluation context, for example, it might involve looking at data from a goals-based and a goals-

free perspective to understand the differences in how it meets those different sets of criteria (see Chapter 1).

Case studies

Method and data triangulation come together in case studies. The term ‘case studies’ is used in different ways, sometimes to imply a focus on a single case, but in this chapter we use it to mean bringing together different perspectives to understand a context, or a set of contexts, in more detail (Patton, 2002; Robson, 2002; Yin, 1993 and 1994). The different perspectives might come from the use of qualitative and quantitative research, from the use of different qualitative methods, or from drawing together the accounts of different players or groups – such as staff, managers, clients and other stakeholders. The context might be:

- a relationship – between a couple, for example, or a professional and their client;
- an organisational entity – such as a local education authority, or a school, or a class;
- a process – for example, the legal resolution of relationship breakdown, or the determination of an application for a social security benefit.

Case study designs are used where what is required is very detailed in-depth understanding that is holistic, comprehensive and contextualised. They allow comparisons to be made between different actors within a single case, between cases, and between groups across cases. So, for example, in the context of schools-based research, one could look at how different people within a school have different understandings of a new educational initiative, or at how different schools have implemented it differently, or at how head teachers view the initiative compared with teachers or pupils.

Case study designs are intensive and thus can be expensive, time-consuming and complex, but they can bring very powerful understanding to policy evaluation. Observational and ethnographic data can be triangulated by other people who took part in the observed activity, by other researchers (especially where audio- or video-recordings

can retrieve what took place), and by documentary analysis (e.g. of case notes, office files, report, minutes, memoranda).

Making sense of triangulated data

There is some debate about whether the purpose of triangulation is to verify (that is, establish the truth) or to broaden and deepen understanding. The idea of using different approaches to verify understanding is increasingly challenged, particularly on the grounds that there is no single version of ‘truth’ that can be captured, and that different methods or approaches inevitably produce different types of data which are unlikely to be concordant (Brannen, 1992; Denzin, 1989, 1997; Fielding and Fielding, 1986; Flick, 1992; Hammersley and Atkinson, 1995).

The purpose of triangulation is more often understood to be to add richness, depth and breadth to a study. Different methods, or different populations, are unlikely to yield data that calibrate precisely. But drawing together different perspectives and types of information provides a more rounded understanding. Where inconsistencies are found, they may be a key finding in themselves, for example, highlighting that staff and clients have different understandings of why people use the service, or that what staff describe as ‘client-focused and responsive’ practice is not experienced as such by clients. Or they may prompt further examination of datasets to see if explanations for the inconsistency can be found. But the approach is not generally used to mediate between different methods or sources and to say which is ‘right’. This type of in-depth analysis requires iteration and time, and the value of a study using triangulation can be lost if time is not allowed for it.

Example: Evaluation of New Deal for Lone Parents (Lewis *et al.*, 2000)

<http://www.dwp.gov.uk/asd/asd5/rport122/main.pdf>

This study was commissioned by DWP as part of the evaluation of New Deal for Lone Parents in its earlier stages. The study used in-depth interviews with lone parents who had used NDLP, followed by focus groups with personal advisers who deliver the

service. The qualitative research with lone parents showed how their work experiences and broader aspects of their lives gave rise to very different needs of the service. It also highlighted diversity in their experiences of the service in terms of the intensity of contact, the breadth and depth of support given, the allocation of responsibility for action, the pace and degree of work-focus, the amount of personal support and attention to underlying issues and whether or not lone parents found work after using it. These differences were not fully explained by lone parents' different circumstances.

The research with personal advisers used vignettes – short examples of cases – drawn from the lone parents interviewed and more general discussion of approaches and practices. This highlighted, and provided reasons for, differences in the way in which personal advisers approach their jobs, which appeared to be important explanation for the diversity of experiences found among lone parents. The research highlighted that these differences in personal advisers' practices were influenced by a range of factors, which shaped advisers' perceptions of the needs of clients, the scope of the service to meet them, and views about appropriate ways of working in different cases.

Example: Experiences of and Attitudes to Disability (Grewal *et al.*, 2002)

<http://www.dwp.gov.uk/asd/asd5/rrep173.asp>

This study, funded by DWP, involved in-depth interviews and focus groups with disabled people, focus groups with non-disabled people and a large-scale survey of disabled and non-disabled people. The qualitative work preceded the survey and was used to shape the survey coverage and develop specific questions.

For example, the qualitative research showed that very different meanings are attached to the concepts of 'prejudice' and 'discrimination'. The former is seen as an expression of perceptions of difference, sometimes unintentional, such as staring or asking inappropriate questions. The latter is seen as more serious, involving organisations or their representatives preventing someone from having an equal role in society. The qualitative research highlighted many examples of each in the experiences of disabled

people, and provided understanding of their impacts, their perceived causes, and views about what might be done to tackle them. The survey was then used to look at the extent to which disabled people had encountered different forms of discrimination and prejudice, deriving examples from the accounts of those who had participated in the qualitative research.

8.5.4 Validation

The concept of validation is sometimes linked with triangulation. Validation means bringing to bear on research findings other forms of knowledge or information. Again, there is some debate about whether this is to verify the findings (i.e. to establish their 'truth' or their credibility) or more generally to bring another perspective. The most common forms of validation are:

- *participant or member validation*, which involves taking findings back to research participants to see whether the researcher's interpretation of data accords with the perspectives and intentions of participants. This can be a useful check on bias and the quality of research, although, of course, research participants may not necessarily have a neutral or objective take on their own behaviour. It is also a chance to fill gaps in data or in explanations, and it is an important aspect of participatory and emancipatory evaluation.
- *peer or expert validation*, which involves taking findings to a wider group who know the research phenomenon or population well, to see how the findings chime with their knowledge and experiences.
- *validation through corroboration with other research*, where other research is used to help to assess the credibility of the findings. Again, neither set of findings can necessarily be privileged, but one would expect to find robust explanations for and checks on findings which were out of line with other research.

Implications for government research, evaluation and policy making

Qualitative research design is a complex art, and it is a process – a series of decisions which need to be reviewed and revisited – rather than a one-off event. For government evaluators it is critical that:

- Research questions are clear, sharp and understood.
- There is a coherence between the questions and the research design in terms of the populations, settings, data collection methods and timing of data collection.
- Thought has been given to whether comparison needs to be built into the design to sharpen its focus.
- Different data methods and sources are used in combination to maximise the potential to provide full, insightful and credible answers to the research questions.

8.6 Sampling in Qualitative Research

Key principles

Qualitative research sampling has a quite different logic from that of quantitative research. The objective is to select the individual cases (which might be people, documents, visual images, events, settings, etc.) that will provide the most illuminating and useful data addressing the research questions. The intention is therefore not to provide a precise statistical representation of the researched population, but to reflect aspects of its diversity that are expected to generate insight.

There are two main approaches to sampling. In *purposive sampling* (Arber, 2002; Patton, 2002; Ritchie and Lewis, 2003) sample cases are chosen deliberately to represent characteristics known or suspected to be of key relevance to the research questions. The selection criteria are prescribed at the first stage of sample design, based on a review of existing research or information, discussions with people with expertise or experience in the research area, or on the researcher's hypotheses. The required composition and size of the sample is then determined and individual cases selected to fit the required composition. Purposive sampling is also called 'judgement sampling' (Burgess, 1984; Honigmann, 1982) or 'criterion sampling' (LeCompte and Preissle, 1993).

In *theoretical sampling* (initially Glaser and Strauss, 1967; Strauss, 1987; Strauss and Corbin, 1998; see also Bryman, 2001; Finch and Mason, 1990; Seale, 1999), the researcher makes decisions, as the study proceeds, on the basis of emergent theory from their analysis of initial data, about the type of data to collect and participants to involve next. Theoretical sampling is strongly associated with grounded theory – an approach that focuses on generating theory inductively from data. Here, the sample evolves during the course of the study as theoretical thinking develops. The process is an iterative one: an initial sample is selected, the data analysed and an emergent theory developed; a further sample is then selected to help refine an aspect of the emerging theory and the data analysed; this leads to further development of theory, more case selection and fieldwork, and so on. The process continues until ‘data saturation’ is reached – the point where it is judged that no new insights would be obtained from further cases.

Theoretical sampling is likely to be more appropriate where theory development is a central objective of the study, or for a research subject or setting about which too little is known to allow selection criteria to be prescribed before sampling and fieldwork begin. However, there is much less clarity about funding and time requirements than with purposive sampling, and purposive sampling is more often used in government evaluations, although with the flexibility to modify sampling and data collection as appropriate.

It is important to understand that the deliberate, non-random selection of cases in qualitative research is its strength, not a weakness. Sample design and selection is systematic and theoretically based, but follows a different logic to quantitative sampling. Some data collection methods and some circumstances require more informal approaches. For example, in ethnographic and participant-observation research there will be less scope to select cases for inclusion at a distance from the research setting and a greater need to take advantage of opportunities as they emerge. Nevertheless, the researcher is making rational and defensible decisions that have a coherence with the research questions.

Qualitative samples need to be large enough to include key subgroups and to reflect diversity. The emphasis is on mapping and understanding issues, rather than counting or numerical representativeness. In fact, large samples are a positive hindrance. The data generated in qualitative research are rich and intensive. Depth lies in the quality of data collection and analysis, not in quantity. The appropriate size of a sample will vary and is always a matter for judgement, but it also needs to be reviewed during fieldwork and as fieldwork draws to a close so that gaps in sample coverage can be filled. The same principles apply for group data collection methods, such as focus groups.

Finally, the sample frames used in qualitative research are varied, as in quantitative research. Broadly, they may be existing data sources, such as survey samples, administrative records, registers or databases, or sources which are generated specifically for the research.

This latter group would include, for example, household screens (where households are approached, usually without prior selection, and a short questionnaire used to ascertain whether the households include anyone eligible to participate). Networking through organisations is also sometimes necessary to generate a sample, where organisations are asked to put potential participants in touch with the research team. For some studies it would be appropriate to use a flow population as a sample source: approaching people in a particular location or setting that is relevant to the research, such as a GP's waiting room or a Jobcentre Plus office. Sometimes the only feasible approach is to use snowballing, where research participants are asked to introduce the research team to another person in the required study group – a method primarily used to find small and dispersed populations where the relevant sample characteristics are unlikely to be widely known to others outside the group (classically, sexuality).

Implications for government research, evaluation and policy making

In assessing the quality of sampling in a qualitative research study, government researchers and policy makers will want to be concerned with:

- whether the study group has been defined and classified in a way that is meaningful, given the diversity present in the relevant population;
- whether the rationale for case selection is logical and likely to include all key subgroups;
- whether the sample size is sufficient to map the required sample with adequate depth in key subgroups, but not so large as to overwhelm the research team or be unfeasible with regard to the time and resources available;
- whether the sample frame used is inclusive and comprehensive and does not build bias or gaps into the sample.

8.7 Analysis of Qualitative Data

Key principles

Approaches to the analysis of qualitative data vary, particularly reflecting different assumptions about knowledge, social existence and the nature of qualitative research. Some distinctive theoretical approaches to qualitative research offer frameworks for analyzing qualitative data, such as phenomenological analysis, grounded theory, qualitative comparative analysis and analytic induction (Patton, 2002). More generally, approaches to analysis vary, for example, according to:

- the status of the data – whether it is treated as a representation of ‘reality’ or as social constructions;
- how the data are condensed or reduced;
- how concepts are derived from or applied to data, and the level of abstraction involved;
- the extent to which data are retained in context;
- the way analysed data are accessed and displayed;
- the place of the researcher in the analytical account.

In thinking about analysis, it is helpful to distinguish between data management (when data are labelled, ordered or summarised) and data interpretation (which involves conveying or displaying data and finding meaning within data).

One distinction between methods for data management is whether they are paper-based or computer-assisted. Paper-based methods differ in terms of whether they keep cases separate or bring them together in a single framework, and whether they involve working with raw data or summarising data. They include:

- case summaries in which individual summaries are constructed for each case;
- thematic data ordering in which raw data from different cases are brought together under thematic headings;
- matrix methods within which each individual case is summarised thematically within a single framework;
- mapping methods where thematic or cognitive maps are designed based on constructions and linkages within the data.

Computer-assisted methods for qualitative data analysis (of which NUD*IST, Ethnograph, NVivo, winMAX and ATLAS/ti are the main packages) vary in terms of: how data are entered and stored; approaches to coding; data-linking mechanisms; mechanisms for search, retrieval and display; mechanisms for attaching memos or notes to codes; and mechanisms for tracking and recording the analysis approach (Fielding and Lee, 1998). A distinction is also drawn between text retrieval, code and retrieve approaches, and theory-building or conceptual network building, although the more popular packages fulfil most or all of these functions.

There are no rules about whether or when paper-based or computer-assisted methods are superior, or which approaches within each are to be preferred. What is more important is that the selected approach is appropriate to the key features and objectives of qualitative research and will aid the later stage of interpretation. Important considerations in the data management process are (Spencer *et al.*, 2004):

- remaining grounded in the data – using a structure that allows analytical ideas and concepts to emerge from data rather imposing them;

- capturing the synthesis of data – so that the link between raw and summarised data is preserved and can be reviewed;
- permitting searches within cases and between cases – so that the integrity of individual cases is retained as well as facilitating thematic analysis and comparisons between cases;
- systematic and comprehensive coverage of the dataset – rather than a selective or biased focus on themes, issues or cases;
- flexibility – so that the structure within which data is managed can be modified if necessary;
- transparency – so that the approach can be described and made accessible to others.

Whatever the data management method used, the process of interpretation is an intellectual one in which the researcher must draw on their own cognitive and conceptual skills. It involves ‘creativity, intellectual discipline, analytical rigor, and a great deal of hard work’ (Patton, 2002, p. 442).

The outputs of this involve *descriptive analyses*, the identification of patterns and associations in data, and interpretive and explanatory accounts. In descriptive analyses, the researcher is concerned with the substantive content of their data. They seek to identify key dimensions of phenomena, experiences, attitudes or behaviours, and to construct, categorise and display them in ways that illuminate the data. In speech- and text-based methods they are concerned with the language used by respondents, in observation with their detailed behaviours, and they seek to display the richness, colour and texture of the original data. Lofland (1971, p. 17) describes this as ‘documenting in loving detail the range of things that exist’. In evaluative research, this might involve describing, for example, the way in which a service is organised and delivered; the types of support provided; the circumstances of service users; or the range of (intended and unintended) outcomes experienced.

In *associative analysis*, the researcher looks for patterns, replication and linkages in the dataset. These might be associations within the data, such as linkages between attitudes,

or between attitudes and behaviours, or between circumstances and needs. Or they might be patterns in the location of a phenomenon within the data (which types of people held a particular view, for example), or differences in how it is manifested among different cases. The purpose here is not to display differences or associations quantitatively, but to use the associations or patterns found in the data to enrich understanding of the phenomenon in question, and to prompt further searching of the dataset to understand why the association or pattern exists. In evaluative research, this might involve, for example, looking at which providers deliver the service in different ways; which service users experience positive outcomes; linkages between service delivery and outcomes; or how requirements of the service are influenced by circumstances.

In *interpretive and explanatory analysis*, the researcher builds explanation from or finds explanation within the data for the views, behaviours or accounts described, and for the associations and patterns found. These are not narrow deterministic explanations. They are sometimes described as explanations at the level of meaning rather than at the level of cause (Hughes and Sharrock, 1997), that is, explanations that reflect the social construction of phenomena and the meaning attributed to them by research participants. Where ‘causes’ are offered, they are not based on mechanistic linkages between isolated variables, but on an examination of the way in which different meanings and understandings come together to influence an outcome, behaviour or decision (Patton, 2002), and on identifying the nature of, and relationships between, different influences or contributory factors. They might be, for example, explanations for why a service is delivered differently by different providers; why some users experience positive outcomes and others do not; how perverse or unwanted consequences arise; or how and why the needs of different service users vary.

The ‘evidence’ on which explanations are based may be (Ritchie et al, 2003):

- explicit statements made within an individual account;
- constructed by the researcher, based on an underlying logic inferred from the data – this may be underpinned, for example, by the juxtaposition or interweaving of

themes, the absence of something in one case which is found in another, or the repeated coexistence of phenomena;

- everyday or ‘common sense’ assumptions;
- explanatory meta-concepts developed from the study;
- concepts or explanations drawn from other studies;
- existing established theoretical or conceptual frameworks.

For all three types of analysis – descriptive, associative and explanatory – the process involves reviewing and interrogating the dataset. This interrogation will be prompted by the research questions, by the researcher’s own hunches and by other research or theoretical frameworks. It involves moving between different levels of abstraction and conceptualisation and carrying out both detailed study within individual cases and comparisons between cases. The search is systematic and comprehensive and the researcher looks for convergence and divergence between cases, for typical and outlying cases, always testing the fit of an association or explanation across the dataset, expecting to find multiplicity and looking for rival explanations.

The quality of qualitative analysis comes from the creativity, flair and insight of the questions the researcher asks of the dataset. It is underpinned too by the analytical conscience they show in systematically reviewing the fit of apparent findings across the dataset, looking for diversity and multiplicity and refusing to be content with single stories or explanations. Qualitative analysis requires considerable intellectual transformative work with the data, which is why some commentators have referred to qualitative analysis as being akin to the artistic process. It is certainly a time-consuming activity and should be allocated enough space in the overall research process.

Example: A qualitative study of mentoring interventions with young people (Philip *et al.*, 2004)

<http://www.jrf.org.uk/knowledge/findings/socialpolicy/324.asp>

This study, funded by the Joseph Rowntree Foundation, explored how young people interpret mentoring in relation to other social and professional relationships, and how their own backgrounds and experiences interact with mentoring. It involved a literature review, analysis of documents, observation of mentoring relationships, two phases of interviews with young people in mentoring relationships and interviews with mentors and other stakeholders. The researchers make clear that they draw particularly on the research with young people to capture the transitions and changes in young people's lives which form the context of mentoring. The use of in-depth interviews in this study seems particularly suitable in addressing subtle and changing relationships. Early in the report individual in-depth examples are used to give a rounded and holistic picture of their lives and of mentoring within them.

The report looks at the underlying processes involved in mentoring – how mentoring relationships become significant, how they can be used by young people to test out ideas and identities, and the negotiations and processes involved in the mentoring relationship. The report also looks at how young people use and interpret mentoring, for example how they use skills they see as acquired through mentoring to negotiate with parents and friends; how they see them helping in rebuilding relationships; and how they help in reflecting on significant life events.

Example: Learning in Adult Life (Schuller *et al.*, 2002)

<http://www.learningbenefits.net/>

This research study, conducted by a centre funded by the Department for Education and Skills, involved over 140 biographical interviews with adults involved in a range of different learning contexts. The research explored interactions between learning and life, and the effects of learning on people's lives.

The study explores these issues in particular depth. They look at how learning shapes identities which link with civically aware behaviour, and at the influence of personal and external contexts on this process. Individual biographies are used as examples to illustrate people's pathways into and around learning and the interactions between health, family and civil participation effects.

For example, in highlighting that a growth in self-confidence was the most fundamental and pervasive positive impact, the researchers identified 15 different benefits which respondents described as flowing from increased self-confidence. The analysis also explored in detail the direct and indirect effects of learning on social capital and social cohesion: impacts included the acquisition of civic competences by individuals and providing opportunities for civic engagement. The research analysed the impact on learning on social networks and the flow of social capital between groups, identifying four mechanisms of transmission.

Implications for government research, evaluation and policy making

There are no standardised procedures for qualitative data analysis, but this does not mean that analysis should not be systematic, grounded in the data and defensible. Flair and creativity are undoubtedly important aspects of qualitative data analysis – and indeed of quantitative data analysis too. But the analysis process needs to be rigorous, comprehensive and involve very close attention to detail. Government researchers and users of research should expect to see:

- transparency about how analysis has been carried out, that is, in what form the data was used, how analytical categories were constructed and applied, how the data were searched and reviewed;
- analytical outputs that provide an accessible, structured and nuanced window on the data and which are oriented clearly to the research questions;
- multiplicity and diversity – it should be clear that the researchers have not stopped at the first and most obvious explanations or the most recurrent views or stories. For every finding, they should expect to see discussion of atypical or divergent cases and evidence that the researchers have looked for alternative explanations.

8.8 Presenting and Using Qualitative Research

Qualitative research evaluations – like those using experimental or quantitative methods – will rarely point directly to a single course of policy action. This is partly because a good piece of qualitative research will make a research user more, rather than less, aware

of the complexity and diversity of their subject; and partly because good policy development often involves drawing on more than one piece of research or information. The worlds of research and policy making overlap but they are not identical in terms of the expertises and knowledge they draw on. Developing policy out of research is a creative and reflective process of translation and not simply a matter of direct application.

Nevertheless, there clearly are circumstances which should make research users in Government more or less secure about using a piece of qualitative research in the formulation and development of policy and in making decisions about moving from piloting to rolling out policies. Central here are the robustness and credibility of the research.

Reports of qualitative research evaluations should provide a detailed explanation of the design and conduct of the research, outlining key decisions made about: the choice of research methods; the selection of sites, populations and samples; the design and focus of data collection instruments; the fieldwork strategy and how it was implemented; and the approach to analysis. Reflections on limitations that flow from the design and conduct or caveats in using the research also add credibility and reassurance.

We looked in the previous section at features of analysis that lend credibility, particularly whether it is clear that the data have been explored systematically and comprehensively. Triangulation, validation and looking at how the findings relate to existing research can also be helpful.

In considering the findings, users should also look for evidence of a clear link between data and conclusions, that is, the building blocks or conceptual steps which led from data to conclusion should be displayed. Research reports should clearly address the original research questions and demonstrate how and to what extent they were answered in the research. The report should convey the depth and richness of the data. The reader should feel they have stepped into the research participants' shoes and that they have been given a chance to understand their world from within. The report should convey the complexity

and subtlety of the research phenomenon, but in a way that is structured and makes it more, not less, clear. In evaluations, it should also be clear what criteria have been used in the appraisal of the policy or service and where these came from.

Determining what reliance can be placed on the findings and how they can be used also raises the issue of whether and how the findings can be generalised. i.e. whether they have a wider relevance beyond the specific study sample. The considerations involved will be slightly different depending on the relationship between what has been researched and the setting to which the findings are to be applied – whether they are the same setting, or whether what is envisaged is the application of findings from one setting to another. These require what are essentially different forms of generalisation, the former being *representational generalisation* (generalising from the study sample to the population it represents) and the latter *inferential generalisation* (extrapolating from the study setting or population to another setting or population) (Ritchie and Lewis, 2003).

In *representational generalisation*, the concern will be with how well the setting and population were represented in the research study design and sample, the quality of data collection and the credibility of the claims made. In inferential generalisation, an additional consideration will be the degree of congruence between the ‘sending’ context from which the findings derive, and the ‘receiving’ context to which their application is being considered (Lincoln and Guba, 1985). Assessing this requires ‘thick description’ (Geertz, 1993) of research settings and observations, so that they can be understood in depth by the research user, and similarities with and differences from the context to which the findings are to be applied can be understood.

Assessing how a piece of qualitative research can be used involves forming a view about its quality. Good quality standards and criteria for research are as important for qualitative research as they are for quantitative investigations. Evaluators, commissioners, policy makers and funders need to be able to make assessments about the quality of research studies they are using or intend to use. For example, research commissioners need to assess the quality of a proposal; policy makers need to make an

assessment of the quality of research findings to include them in the development of new policies. Qualitative research studies have to be of a high enough standard to instill trust in the findings they develop, the arguments they present and the case they are trying to make.

Qualitative research has used peer review as a means of assessing quality – that is, research is assessed by others working in the field – for example, in appraising journal articles and in proposal assessments. While it is an accepted mode of quality assessment, it has its pitfalls in that it is often applied inconsistently and inadequately (Boaz and Ashby, 2001).⁴ As qualitative research has expanded and found entry into different areas of policy evaluation, the notion of a quality framework of qualitative research has been discussed as a means of promoting high-quality research standards in qualitative research. A number of quality criteria and frameworks of assessing quality have been developed in different research fields (Spencer *et al.*, 2004).

However, the development of quality standards is not universally accepted and remains a contested area. Critics have pointed out that the diversity of philosophical and theoretical positions informing qualitative research makes it difficult to derive at common standards of quality. They point to different traditions of design, data collection, analysis and reporting/writing-up styles, which are difficult to bring under one common understanding of quality. Some commentators reject the idea of common quality standards for political reasons: they feel that the diversity of the field is a strength, and that standards would bring a push towards uniformity and would lead to preferential use of some techniques over others. Against this, the movement towards setting criteria for qualitative research argues that despite the great diversity and multiplicity of qualitative approaches, qualitative research and evaluation adheres to some core of quality principles. Developing frameworks injects openness, clarity and transparency into the process. Used with care, quality frameworks can champion innovative and diverse qualitative research approaches.

⁴ Of course, this is also true of peer review in quantitative research.

Quality in Qualitative Evaluation: A framework for assessing research evidence has attempted to stir a middle pass between developing quality criteria that qualitative research should adhere to and encompassing a wide range of qualitative research. It is based on four guiding principles for assessing quality, addressed through a total of 18 appraisal questions, which reflect the key features and processes involved in qualitative inquiry (Spencer *et al.*, 2004, p. 20).

The guiding principles are:

- Research should contribute to knowledge and understanding.
- Research should be defensible in design.
- Research should be rigorous in conduct.
- Research should be credible in claim.

The 18 appraisal questions cover findings, design, sampling, data collection, analysis, reporting and other aspects of research conduct. The result of this is a detailed assessment tool, which is sensitive to the diversity of qualitative research but provides guidance for the appraisal of individual research studies.

Implications for government research, evaluation and policy making

No single piece of research, whether qualitative or quantitative, will point unconditionally towards a single policy action. But qualitative research, when properly conducted and credibly reported, will provide insight into the policy, how it operates, what makes it successful in different circumstances, and how it sits and is experienced within the social worlds of those it affects. Users of research need to appraise critically the quality of qualitative research evidence and reporting. There are no standardised rules or procedures for doing this, but the user of research can expect a report to give them:

- a clear explanation of the research design and an account of its conduct, where the decisions made by the research team and their implications for the robustness of the data are discussed;

- a sense that the data have been critically and thoroughly analysed, with different data methods and sources used where appropriate. The reader should be left with a more nuanced and deeper understanding of the research topic, with new insights and conceptions of the topic;
- findings which are credible and plausible, rooted in the data, with the links between data and conclusions apparent;
- guidance as to the relevance the findings have beyond the study sample and setting, and enough description of the setting and findings for users to make their own judgements about whether the findings can be extrapolated to a different setting.

Recommended further reading

Bryman, R. (2001), *Social Research Methods*, Oxford: Oxford University Press.

Hammersley, M. and Atkinson, P. (1995), *Ethnography: Principles and Practice*, London: Routledge.

Mason, J. (2002), *Qualitative Researching*, London: Sage.

Patton, M. Q. (2002), *Qualitative Research and Evaluation Methods*, 3rd edition, Thousand Oaks: Sage.

Ritchie, J. and Lewis, J. (2003) (eds), *Qualitative Research Practice*, London: Sage.

Silverman, D. (2000), *Doing Qualitative Research: A Practical Handbook*, London: Sage.

References

Adler, M. and Ziglio, E. (1996), *Gazing Into The Oracle: The Delphi Method and Its Application to Social Policy and Public Health*, London: Jessica Kingsley Publishers.

Arber, S. (2002), 'Designing Samples' in Gilbert, N. (ed.), *Researching Social Life*, 2nd edition, London: Sage.

Atkinson, J. M. (1984), *Our Master's Voices: The Language and Body Language of Politics*, London: Routledge.

Audit Commission (2003), *Connecting with Users and Citizens*, London: Audit Commission.

Boaz, A. and Ashby, D. (2001) *Fit for purpose? Assessing research quality for evidence based policy and practice* ESRC, Working Paper 11, London: UK Centre for Evidence Based Policy and Practice.

Bloor, M., Frankland, J., Thomas, M. and Robson, K. (2001), *Focus Groups in Social Research*, London: Sage.

Brannen, J. (1992) (ed.), *Mixing Methods: Qualitative and Quantitative Research*, Aldershot: Gower.

Brown, R. H. (1987), *Society as Text: Essays on Rhetoric, Reason and Reality*, Chicago: University of Chicago Press.

Bryman, R. (2001), *Social Research Methods*, Oxford: Oxford University Press.

Burgess, R. G. (1984), *In the Field: An Introduction to Field Research*, London: Allen & Unwin.

Cantrill, J. A., Sibbald, B. and Buetow, S. (1996), 'The Delphi and Nominal Group Techniques in Health Services Research', *The International Journal of Pharmacy Practice*, No. 4, p. 67–74.

Chambers, E. (2000), 'Applied Ethnography', in N.K. Denzin and S. Lincoln (Eds.), *Handbook of Qualitative Research*, 2nd Edition, Thousand Oaks: Sage, pp. 851-869.

Coote, A. and Lenaghan, J. (1997), *Citizens' Juries: Theory Into Practice*, London: Institute for Public Policy Research.

Critcher, C. and Gladstone, B. (1998), 'Utilizing the Delphi Technique in Policy Discussion: A Case Study of a Privatized Utility in Britain', *Public Administration*, Vol. 76, pp. 431–449.

Davies, P. T. (2000), 'Contributions from Qualitative Research' in Davies, H. T. O., Nutley, S. M. and Smith, P. C. (eds), *What Works: Evidence-Based Policy and Practice in Public Services*, Bristol: The Policy Press.

Davies, S., Elizabeth, S., Hanley, B., New, B. and Sang, B. (1998), *Ordinary Wisdom: Reflections on an Experiment in Citizenship and Health*, London: Kings Fund.

Denzin, N. K. (1989), *The Research Act: A Theoretical Introduction to Sociological Methods*, 3rd edition, Englewood Cliffs, NJ: Prentice Hall.

Denzin, N. K. (1997), *Interpretive Ethnography: Ethnographic Practices for the 21st Century*, Thousand Oaks, CA: Sage.

Denzin, N. K. and Lincoln, Y. S. (2000), *Handbook of Qualitative Research*, 2nd edition, Thousand Oaks, CA: Sage.

- Dingwall, R. (1997), 'Accounts, Interviews and Observations' in G. Miller and R. Dingwall (eds), *Context and Method in Qualitative Research*, London: Sage.
- Fielding, N. and Fielding, J. (1986), *Linking Data*, London: Sage.
- Fielding, N. and Lee, R. M. (1998) (eds), *Using Computers in Qualitative Research*, London: Sage.
- Finch, H. and Lewis, J. (2003), 'Focus Groups' in J. Ritchie and J. Lewis (eds), *Qualitative Research Practice*, London: Sage.
- Finch, J. and Mason, J. (1990), 'Decision Taking in the Fieldwork Process: Theoretical Sampling and Collaborative Working' in R. G. Burgess (ed.), *Studies in Qualitative Methodology*, No. 2, pp. 25–50.
- Fishkin, J. (1995), *The Voice of the People*, Yale: Yale University Press.
- Flick, U. (1992), 'Triangulation Revisited: Strategy of Validation or Alternative?', *Journal for the Theory of Social Behaviour*, Vol. 22, No. 2, pp. 175–97.
- Fulop, N., Protopsaltis, G., Hutchings, A., King, A., Allen, P., Normand, C. and Walters, R. (2002), 'The Process and Impact of NHS Trust Mergers: A Qualitative Study and Management Cost Analysis', *British Medical Journal*, pp. 325–246 (3 August).
- Geertz, C. (1993, first published 1973), *The Interpretation of Cultures: Selected Essays*, New York: Basic Books.
- Gibson, T. (1998), *The Do-ers' Guide to Planning for Real*, Neighbourhood Initiatives Foundation.
- Gilbert, N. (1993), *Researching Social Life*, London: Sage.

Glaser, B. G. and Strauss, A. L. (1967), *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago: Aldine de Gruyter.

Goffman, E., (1963), *Stigma: Notes on the Management of Spoiled Identity*, New York: Simon & Schuster.

Goffman, E., (1961), *Asylums. Essays on the Social Situation of Mental Patients and Other Inmates*, Garden City, NY: Doubleday Anchor.

Goffman, E., (1959), *The Presentation of Self in Everyday Life*, Garden City, NY: Doubleday Anchor.

Goffman, E. (1963) *Behavior in Public Places: Notes on the Social Organization of Gatherings*, Glencoe: Free Press.

Grewal, I., Joy, S., Lewis, J., Swales, K., Woodfield, K. with Bailey M. (2002) *Disabled for Life? Attitudes Towards, and Experiences of, Disability in Britain*, London: Corporate Document Services.

Hammersley, M. and Atkinson, P. (1995), *Ethnography: Principles and Practice*, London: Routledge.

Hayden, C., Boaz, A. and Taylor, F. (1999) *Attitudes and Aspirations of Older People: A Qualitative Study*, DSS Research Report No. 102, London: Department for Social Security.

Hodder, I. (2000), 'The Interpretation of Documents and Material Culture', in N. K. Denzin and Y. S. Lincoln, *Handbook of Qualitative Research*, 2nd edition, London: Sage, pp. 703–716.

Honigmann, J. J. (1982), 'Sampling in Ethnographic Fieldwork', in R. G. Burgess (ed.) *Field Research: A Source Book and Field Manual*, London: Allen & Unwin.

Hughes, J. and Sharrock, W. (1997), *The Philosophy of Social Research*, London: Longman.

Krueger, R. A. and Casey, M. A. (2000), *Focus Groups: A Practical Guide for Applied Research*, Thousand Oaks and London: Sage.

LeCompte, M. D. and Preissle, J., with Tesch, R. (1993), *Ethnography and Qualitative Design in Educational Research*, 2nd edition, Chicago: Academic Press.

Lewis, J., Mitchell, L., Sanderson, T., O'Connor, W. and Clayden, M. (2000), *Lone Parents and Personal Advisers: Roles and Responsibilities*, London: Corporate Document Services.

Liebow, E. (1967), *Tally's Corner: A Study of Negro Streetcorner Men*, Boston: Little, Brown and Company.

Lincoln, Y. and Guba, E. (1985), *Naturalistic Inquiry*, Beverley Hills, CA: Sage.

Lincoln, Y. and Guba, E. (1986), 'Research, Evaluation and Policy Analysis: Heuristics and Disciplined Inquiry', *Policy Studies Review*, Vol. 5, No. 3, pp. 546–565.

Lofland, J. (1971). *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*. Belmont, CA: Wadsworth.

Lowndes, V., Stoker, G., Pratchett, L., Wilson, D., Leach, S. and Wingfield, M. (1998), *Enhancing Public Participation in Local Government*, London: Department of the Environment, Transport and the Regions.

Mason, J., (2002), *Qualitative Researching*, London: Sage.

MacDonald, K. and Tipton, C. (2002) 'Using documents', in Gilbert, N. (ed.) *Researching Social Life*, London: Sage, pp. 187-200.

Morgan, D. L. (1997), *Focus Groups as Qualitative Research*, 2nd edition, Thousand Oaks and London: Sage .

Oxfam UK Poverty Programme (2003), *Have you been PA'd? Using Participatory Appraisal to Shape Local Services*, Glasgow: Oxfam GB.

Park, A., Jowell, R. and McPherson, S. (1999), *The Future of the National Health Service: Results from a Deliberative Poll*, London: Kings Fund.

Patton, M. Q. (2002), *Qualitative Research and Evaluation Methods*, 3rd edition, Thousand Oaks: Sage.

Philip, K., Shucksmith, J. and King, C. (2004) *'Sharing a Laugh? A Qualitative Study of Mentoring Interventions with Young People*, York: Joseph Rowntree Foundation.

Popay, J. and Williams, G. (1998), 'Qualitative Research and Evidence-based Healthcare', *Journal of the Royal Society of Medicine*, Vol. 191, No. 35, pp. 32–37.

Ricoeur, P. (1996), *The Hermeneutics of Action*, edited by Richard Kearney, London: Sage.

Ritchie, J. and Lewis, J. (2003), 'Designing and Selecting Samples', in J. Ritchie and J. Lewis (eds.), *Qualitative Research Practice*, London: Sage.

Ritchie, J., Spencer, L. and O'Connor, W. (2003), 'Carrying out Qualitative Analysis' in J. Ritchie and J. Lewis (eds.), *Qualitative Research Practice*, London: Sage.

Robson, C., (2002), *Real World Research*, 2nd edition, Oxford: Blackwell.

Schaffner, C. (1997) (ed.), *Analysing Political Speeches*, Clevedon: Multilingual Matters.

Schuller, T., Brassett-Grundy, A., Green, A., Hammond, C. and Preston, J. (2002), *Learning, Continuity and Change in Adult Life*, Wider Benefits of Learning Research Report No. 3, London: Institute of Education.

Scriven, M. (1991), *Evaluation Thesaurus*, 4th edition, Newbury Park CA: Sage.

Seale, C. (1999), *The Quality of Qualitative Research*, Oxford: Blackwell.

Sargeant, J. and Steele, J. (1998), *Consulting the Public: Guidelines and Good Practice*, London: Policy Studies Institute.

Shaw, I. (1999), *Qualitative Evaluation*, London: Sage.

Spencer, L., Ritchie, J., Lewis, J. and Dillon, L. (2004), *Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence*, 2nd edition, London: Cabinet Office.

Stewart, D. W. and Shamdasi, P. N. (1990), *Focus Groups: Theory and Practice*, London: Sage.

Strauss, A. L. (1987), *Qualitative Analysis for Social Scientists*, Cambridge: Cambridge University Press.

Strauss, A. L. and Corbin, J. (1998), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, 2nd edition, Thousand Oaks, CA: Sage.

Webb, B. and Webb, S. (1932), *Methods of Social Study*, London: Longmans Green.

Wengraf, T. (2001), *Qualitative Research Interviewing*, London: Sage.

White, C., Elam, G. and Lewis, J. (1999), *Citizens' Juries: An Appraisal of their Role*, London: Cabinet Office.

Whyte, W. F. (1943) *Street Corner Society*, 2nd edition, Chicago: University of Chicago Press.

Williams, D. D. (1986), *Naturalistic Evaluation: (Programme Evaluation Series 30)*, San Francisco: Jossey-Bass Wiley.

Wilson, J. (1990), *Politically Speaking: The Pragmatic Analysis of Political Language*, Oxford: Basil Blackwell.

Yin, R. K. (1994), *Case Study Research: Design and Methods*, 2nd edition, Beverley Hills, CA: Sage.

Yin, R. K. (1993), *Applications of Case Study Research*, Newbury Park, CA: Sage.