Semi-Structured Qualitative Studies

Abstract

HCI addresses problems of interaction design: delivering novel designs, evaluating existing designs, and understanding user needs for future designs. Qualitative methods have an essential role to play in this enterprise, particularly in understanding user needs and behaviours and evaluating situated use of technology. There are, however, a huge number of qualitative methods, often minor variants of each other, and it can seem difficult to choose (or design) an appropriate method for a particular study. The focus of this chapter is on semi-structured qualitative studies, which occupy a space between ethnography and surveys, typically involving observations, interviews and similar methods for data gathering, and methods for analysis based on systematic coding of data. This chapter is pragmatic, focusing on principles for designing, conducting and reporting on a qualitative study and conversely, as a reader, assessing a study. The starting premise is that all studies have a purpose, and that methods need to address the purpose, taking into account practical considerations. The chapter closes with a checklist of questions to consider when designing and reporting studies.

1 Introduction

HCI has a focus (the design of interactive systems), but exploits methods from various disciplines. One growing trend is the application of qualitative methods to better understand the use of technology in context. While such methods are well established within the social sciences, their use in HCI is less mature, and there is still controversy and uncertainty about when and how to apply such methods, and how to report the findings (e.g. Crabtree et al., 2009).

This chapter takes a high-level view on how to design, conduct and report semi-structured qualitative studies (SSQSs). Its perspective is complementary to most existing resources (e.g. Adams et al., 2008; Charmaz, 2006; Lazar et al., 2010; Smith, 2008; http://onlineqda.hud.ac.uk/), which focus on method and principles rather than basic practicalities. Because 'method' is not a particularly trendy topic in HCI, I draw on the methods literature from psychology and the social sciences as well as HCI. Rather than starting with a particular method and how to apply it, I start from the purpose of a study and the practical resources and constraints within which the study must be conducted.

I do not subscribe to the view that there is a single right way to conduct any study: that there is a minimum or maximum number of participants; that there is only one

way to gather or analyse data; or that validation has to be achieved in a particular way. As Willig (2008, p.22) notes, "Strictly speaking, there are no 'right' or 'wrong' methods. Rather, methods of data collection and analysis can be more or less appropriate to our research question." Woolrych et al. (2011) draw an analogy with ingredients and recipes: the art of conducting an effective study is in pulling together appropriate ingredients to construct a recipe that is right for the occasion – i.e., addresses the purpose of the study while working with available resources.

The aim of this chapter is to present an overview of how to design, conduct and report on SSQSs. The chapter reviews methodological literature from HCI and the social and life sciences, and also draws on lessons learnt through the design, conduct and reporting of various SSQSs. The chapter does not present any method in detail, but presents a way of thinking about SSQSs in order to study users' needs and situated practices with interactive technologies.

The basic premise is that, starting with the purpose of a study, the challenge is to work with the available resources to complete the best possible study, and to report it in such a way that its strengths and limitations can be inspected, so that others can build on it appropriately. The chapter summarises and provides pointers to literature that can help in research, and at the end a checklist of questions to consider when designing, conducting, reporting on, and reviewing SSQSs. The aim is to deliver a reference text for HCI researchers planning semi-structured qualitative studies.

1.1   What is an SSQS?

The term 'semi-structured qualitative study' (SSQS) is used here to refer to qualitative approaches, typically involving interviews and observations, that have some explicit structure to them, in terms of theory or method, but are not completely structured. Such studies typically involve systematic, iterative coding of verbal data, often supplemented by data in other modalities.

Some such methods are positivist, assuming an independent reality that can be investigated and agreed upon by multiple researchers; others are constructivist, or interpretivist, assuming that reality is not 'out there', but is constructed through the interpretations of researchers, study participants, and even readers. In the former case, it is important that agreement between researchers can be achieved. In the latter case, it is important that others are able to inspect the methods and interpretations so that they can comprehend the journey from an initial question to a conclusion, assess its validity and generalizability, and build on the research in an informed way.

In this chapter, we focus on SSQSs addressing exploratory, open-ended questions, rather than qualitative data that is incorporated into hypothetico-deductive research designs. Kidder and Fine (1987, p.59) define the former as "big Q" and the latter as "small q", where "big Q" refers to "unstructured research, inductive work, hypothesis generation, and the development of 'grounded theory'". Their big Q encompasses ethnography (section 1.4) as well as the SSQSs that are the focus here; the important

point is that SSQSs focus on addressing questions rather than testing hypotheses: they are concerned with developing understanding in an exploratory way.

One challenge with qualitative research methods in HCI is that there are many possible variants of them and few names to describe them. If every one were to be classed as a 'method' there would be an infinite number of methods. However, starting with named methods leaves many holes in the space of possible approaches to data gathering and analysis. There are many potential methods that have no name and appear in no textbooks, and yet are potentially valid and valuable for addressing HCI problems.

This contrasts with quantitative research. Within quantitative research traditions – exemplified by, but not limited to, controlled experiments – there are well-established ways of describing the research method, such that a suitably knowledgeable reader can assess the validity of the claims being made with reasonable certainty, for example, hypothesis, independent variable, dependent variable, power of test, choice of statistical test, number of participants.

The same is not true for SSQSs, where there is no hypothesis – though usually there is a question, or research problem – where the themes that emerge from the data may be very different from what the researcher expected, and where the individual personalities of participants and their situations can have a huge influence over the progress of the study and the findings.

Because of the shortage of names for qualitative research methods, there is a temptation to call a study an 'ethnography' or a 'Grounded Theory' (both described below: sections 1.4 and 1.5) whether or not they have the hallmarks of those methods as presented in the literature. Data gathering for SSQSs typically involves the use of a semi-structured interview script or a partial plan for what to focus attention on in an observational study.

There is also some structure to the process of analysis, including systematic coding of the data, but usually not a rigid structure that constrains interpretation, as discussed in section 7. SSQSs are less structured than, for example, a survey, which would typically allow people to select from a range of pre-determined possible answers or to enter free-form text into a size-limited text box. Conversely, they are more structured than ethnography – at least when that term is used in its classical sense; see section 1.4.

1.2 A starting point: problems or opportunities

Most methods texts (e.g. Cairns and Cox, 2008; Lazar et al., 2010; Smith, 2008; Willig, 2008) start with methods and what they are good for, rather than starting with problems and how to select and adapt research methods to address those problems. Willig (2008, p.12) even structures her text around questions about each of the approaches she presents:

"What kind of knowledge does the methodology aim to produce? ... What kinds of assumptions does the methodology make about the world? ... How does the methodology conceptualise the role of the researcher in the research process?"

If applying a particular named method, it is important to understand it in these terms to be able to make an informed choice between methods. However, by starting at the other end – the purpose of the study and what resources are available – it should be possible to put together a suitable plan for conducting a SSQS that addresses the purpose, makes relevant assumptions about the world, and defines a suitable role for the researcher.

Some researchers become experts in particular methods and then seek out problems that are amenable to that method; for example, drawing from the social sciences rather than HCI, Giorgi and Giorgi (2008) report seeking out research problems that are amenable to their phenomenology approach. On the one hand, this enables researchers to gain expertise and authority in relation to particular methods; on the other, this risks seeing all problems one way: "To the man who only has a hammer, everything he encounters begins to look like a nail", to quote Abraham Maslow.

HCI is generally problem-focused, delivering technological solutions to identified user needs. Within this, there are two obvious roles for SSQSs: understanding current needs and practices, and evaluating the effects of new technologies in practice. The typical interest is in how to understand the 'real world' in terms that are useful for interaction design. This can often demand a '*bricolage*' approach to research, adopting and adapting methods to fit the constraints of a particular problem situation. On the one hand this makes it possible to address the most pressing problems or questions; on the other, the researcher is continually having to learn new skills, and can always feel like an amateur.

In the next section, I present a brief overview of relevant background work to set the context, focusing on qualitative methods and their application in HCI. Subsequent sections cover an approach to planning SSQSs based on the PRET A Rapporter framework (Blandford et al., 2008a) and discuss specific issues including the role of theory in SSQSs, assessing and ensuring quality in studies, and various roles the researcher can play in studies. This chapter closes with a checklist of issues to consider in planning, conducting and reporting on SSQSs.

1.3   A brief overview of qualitative methods

There has been a growing interest in the application of qualitative methods in HCI. Suchman's (1987) study of situated action was an early landmark in recognising the importance of studying interactions in their natural context, and how such studies could complement the findings of laboratory studies, whether controlled or employing richer but less structured techniques such as think aloud.

Sanderson and Fisher (1994) brought together a collection of papers presenting complementary approaches to the analysis of sequential data (e.g., sequences of events), based on a workshop at CHI 1992. Their focus was on data where sequential integrity had been preserved, and where sense was made of the data through

relevant techniques such as task analysis, video analysis, or conversation analysis. The interest in this collection of papers is not in the detail, but in the recognition that semi-structured qualitative studies had an established place in HCI at a time when cognitive and experimental methods held sway.

Since then, a range of methods have been developed for studying people's situated use and experiences of technology, based around ethnography, diaries, interviews, and similar forms of verbal and observable qualitative data (e.g. Lindtner et al. 2011; Mackay 1999; Odom et al. 2010; Skeels & Grudin 2009).

Some researchers have taken a strong position on the appropriateness or otherwise of particular methods. A couple of widely documented disagreements are briefly discussed below. This chapter avoids engaging in such 'methods wars'. Instead, the position, like that of Willig (2008) and Woolrych et al. (2011), is that there is no single correct 'method', or right way to apply a method: the textbook methods lay out a space of possible ways to conduct a study, and the details of any particular study need to be designed in a way that maximises the value, given the constraints and resources available. Before expanding on that theme, we briefly review ethnography – as applied in HCI – and Grounded Theory, as a descriptor that is widely used to describe exploratory qualitative studies.

1.4   Ethnography: the all-encompassing field method?

Miles and Huberman (1994, p.1) suggest, "The terms ethnography, field methods, qualitative inquiry, participant observation, … have become practically synonymous". Some researchers in HCI seem to treat these terms as synonymous too, whereas others have a particular view of what constitutes 'ethnography'. For the purposes of this chapter, an ethnography involves observation of technology-based work leading to rich descriptions of that work, without either the observation or the subsequent description being constrained by any particular structuring constructs. This is consistent with the view of Anderson (1994), and Randall and Rouncefield (2013).

Crabtree et al. (2009) present an overview – albeit couched in somewhat confrontational terms – of different approaches to ethnography in HCI. Button and Sharrock (2009) argue, on the basis of their own experience, that the study of work should involve "ethnomethodologically informed ethnography", although they do not define this succinctly. Crabtree et al. (2000, p.666) define it as study in which "members' reasoning and methods for accomplishing situations becomes the topic of enquiry".

Button and Sharrock (2009) present five maxims for conducting ethnomethodological studies of work: keep close to the work; examine the correspondence between work and the scheme of work; look for troubles great and small; take the lead from those who know the work; and identify where the work is done. They emphasise the importance of paying attention, not jumping to conclusions, valuing observation over verbal report, and keeping comprehensive notes. However, their guidance does not extend to any form of data analysis. In common with others (e.g. Heath & Luff, 1991; Von Lehn & Heath, 2005), the moves

that the researcher makes between observation in the situation of interest and the reporting of findings remain undocumented, and hence unavailable to the interested or critical reader.

According to Randall and Rouncefield (2013), ethnography is "a qualitative orientation to research that emphasises the detailed observation of people in naturally occurring settings". They assert that ethnography is not a method at all, but that data gathering "will be dictated not by strategic methodological considerations, but by the flow of activity within the social setting".

Anderson (1994) emphasises the role of the ethnographer as someone with an interpretive eye delivering an account of patterns observed, arguing that not all fieldwork is ethnography and that not everyone can be an ethnographer. In SSQSs, our focus is on methods where data gathering and analysis are more structured and open to scrutiny than these flavours of ethnography.

1.5 Grounded Theory: the SQSS method of choice?

I am introducing Grounded Theory (GT) early in this chapter because the term is widely used as a label for any method that involves systematic coding of data, regardless of the details of the study design, and because it is probably the most widely applied SSQS method in HCI.

GT is not a theory, but an approach to theory development – grounded in data – that has emerged from the social sciences. There are several accounts of GT and how to apply it, including Glaser and Strauss (2009), Corbin and Strauss (2008), Charmaz (2006), Adams et al. (2008), and Lazar et al. (2010).

Historically, there have been disputes on the details of how to conduct a GT: the disagreement between Glaser and Strauss, following their early joint work on Grounded Theory (Glaser and Strauss, 2009), has been well documented (e.g. Charmaz, 2008; Furniss et al., 2011a, Willig, 2008). Charmaz (2006) presents an overview of the evolution of different strains of GT prior to that date.

Grbich (2013) identifies three main versions of GT, which she refers to as Straussian, involving a detailed three-stage coding process; Glaserian, involving less coding but more shifting between levels of analysis to relate the details to the big picture; and Charmaz's, which has a stronger constructivist emphasis.

Charmaz (2008, p.83) summarises the distinguishing characteristics of GT methods as being:

- Simultaneous involvement in data collection and analysis;
- Developing analytic codes and categories "bottom up" from the data, rather than from preconceived hypotheses;
- Constructing mid-range theories of behaviour and processes;
- Creating analytic notes, or memos, to explain categories;

- Constantly comparing data with data, data with concept, and concept with concept;
- Theoretical sampling – that is, recruiting participants to help with theory construction by checking and refining conceptual categories, not for representativeness of a given population;
- Delaying literature review until after forming the analysis.

There is widespread agreement amongst those who describe how to apply GT that it should include interleaving between data gathering and analysis, that theoretical sampling should be employed, and that theory should be constructed from data through a process of constant comparative analysis. These characteristics define a region in the space of possible SSQSs, and highlight some of the dimensions on which qualitative studies can vary. I take the position that the term 'Grounded Theory' should be reserved for methods that have these characteristics, but even then it is not sufficient to describe the method simply as a Grounded Theory without also presenting details on what was actually done in data gathering and analysis.

As noted above, much qualitative research in HCI is presented as being Grounded Theory, or a variant on GT. For example, Wong and Blandford (2002) present Emergent Themes Analysis as being "based on Grounded Theory but tailored to take advantage of the exploratory and efficient data collection features of the CDM" – where CDM is the Critical Decision Method (Klein et al., 1989) as outlined in section 6.4.

McKechnie et al. (2012) describe their analysis of documents as a Grounded Theory, and also discuss the use of inter-rater reliability – both activities that are inconsistent with the distinguishing characteristics of GT methods if those are taken to include interleaving of data gathering and analysis and a constructivist stance. GT has been used as a 'bumper sticker' to describe a wide range of qualitative analysis approaches, many of which diverge significantly from GT as presented by the originators of that technique and their intellectual descendants.

Furniss et al. (2011a) present a reflective account of the experience of applying GT within a three-year project, focusing particularly on pragmatic 'lessons learnt'. These include practical issues such as managing time and the challenges of recruiting participants, and also theoretical issues such as reflecting on the role of existing theory – and the background of the analyst – in informing the analysis.

Being fully aware of relevant existing theory can pose a challenge to the researcher, particularly if the advice to delay literature review is heeded. If the researcher has limited awareness of relevant prior research in the particular domain, it can mean 'rediscovery' of theories or principles that are, in fact, already widely recognized, leading to the further question, "So what is new?" We return to the challenge of how to relate findings to pre-existing theory, or literature that emerges as being important through the analysis, in section 9.1.

2   Planning and conducting a study: PRET A Rapporter

Research generally has some kind of objective (or purpose) and some structure. A defining characteristic of SSQSs is that they have shape… but not too much: that there is some structure to guide the researcher in how to organise a study, what data to gather, how to analyse it, etc., but that that structure is not immutable, and can adapt to circumstances, evolving as needed to meet the overall goals of the study. The plan should be clear, but is likely to evolve over the course of a study, as understanding and circumstances change.

Thomas Green used to remind PhD students to "look after your GOST", where a GOST is a Grand Overall Scheme of Things – his point being that it is all too easy to let the aims of a research project and the fine details get out of synch, and that they need to be regularly reviewed and brought back into alignment.

We structure the core of this chapter in terms of the PRET A Rapporter (PRETAR) framework (Blandford et al., 2008a), a basic structure for designing, conducting and reporting studies. Before presenting this structure, though, it is important to emphasise the basic interconnectedness of all things: in the UK a few years ago there was a billboard advertisement, "You are not stuck in traffic. You are traffic" (Figure 1).

It is impossible to separate the components of a study and treat them completely independently – although they have some degree of independence. The style of data gathering influences what analysis can be performed; the relationship established with early participants may influence the recruitment of later participants; ethical considerations may influence what kinds of data can be gathered, etc. We return to this topic of interdependencies later; first, for simplicity of exposition, we present key considerations in planning a study using the PRETAR framework.

Figure 1: An example of interconnectedness

The PRETAR framework draws its inspiration from the DECIDE framework proposed by Rogers et al. (2011), but has a greater emphasis on the later – analysis and reporting – stages that are essential to any SSQS:

- Purpose: every study has a purpose, which may be more or less precisely defined; methods should be selected to address the purpose of the study. The purpose of a study may change as understanding develops, but few people are able to conduct an effective study without some idea of why they are doing it.
- Resources and constraints: every study has to be conducted with the available resources, also taking account of existing constraints that may limit what is possible.
- Ethical considerations often shape what is possible, particularly in terms of how data can be gathered and results reported.
- Techniques for data gathering need to be determined (working with the available resources to address the purpose of the study).
- Analysis techniques need to be appropriate to the data and the purpose of the study.
- Reporting needs to address the purpose of the study, and communicate it effectively to the intended audiences. In some cases, this will include an account of how and why the purpose has evolved, as well as the methods, results, etc.

Some authors have focused attention on one of these steps; for example, Kvale and Brinkmann (2009) focus primarily on data gathering while Miles and Huberman (1994), Grbich (2013) and Braun and Clarke (2006) focus on analysis, and Morse (1997) and Wolcott (2009) on reporting and other aspects of closing off a research project. However, these steps are not independent, and are typically interleaved in SSQSs. What matters is that they remain coherent – that there is a clear GOST. For example, the researcher's view of the purpose of a study may evolve as their understanding matures through data gathering and analysis.

As noted above (section 1.5), GT is based on tight coupling between data gathering and analysis; other analysis techniques assume no such coupling. These steps provide a useful structure to organise our discussion on planning and conducting a study, but should not be regarded as strictly sequential or independent.

3   Purpose

Every study has a purpose. That purpose might be to better understand:
- 'work', broadly conceived, and how interactive technologies support or fail to support that work (e.g. Hartswood et al., 2003; Hughes et al., 1994);
- people's experiences with a particular kind of technology (e.g. Palen, 1999; Kindberg et al., 2005; Mentis et al., 2013);
- how people exploit technologies to support cognition (e.g. Hutchins 1995);
- how people make sense of information with and without particular technological support (e.g. Attfield and Blandford, 2011); or
- many other aspects of the design and use of interactive technologies.

The recent 'turn to the wild' (Rogers, 2012), in which novel products are designed in situ, working directly with the intended users of those products, introduces yet more

possible purposes for qualitative studies: to understand how a new product changes attitudes and behaviours, and how the design of the product might be adapted to better support people's needs and aspirations.

Crabtree et al. (2009) argue that the purpose of an ethnographic study in HCI is to inform system design. They claim that ethnographic research to inform systems design has shifted from a study of work towards a study of culture, and that this shift is "harmful". In creating an either/or stand-off, the authors pit the contrasting ethnographic focuses against each other, apparently disregarding the possibility that each has a place in informing design, and that different ethnographic studies of the same context can serve different purposes. The same is true of SSQSs: they can address many different questions that inform design – though whether 'informing design' should mean that there are explicitly stated 'implications for design' is a further question.

Dourish (2006) suggests that there is a widespread expectation that studies will always include "implications for design" as an explicit theme, and that this is expected in reporting even if that was not the purpose of the study. He argues that designers need a rich understanding of the situation for which they are designing, and that one of the important roles for ethnography is to expose and describe that cultural context for design, without necessarily making the explicit link to implications for design. This might be regarded as an argument for helping designers to put themselves in the user's shoes when they are not designing for themselves, for example, designing specialist products to support work or other activities in which they are not experts themselves.

Obviously, not all studies should be qualitative, and certainly not all should be semi-structured. Yardley (2000, p.220) differentiates between the typical purposes of qualitative and quantitative studies:

> "Quantitative studies … ensure the 'horizontal generalization' of their findings across research settings … qualitative researchers aspire instead to … 'vertical generalization', i.e., an endeavour to link the particular to the abstract and to the work of others".

In HCI, qualitative studies – whether structured, semi-structured or ethnographic – most typically focus on understanding technology use, or future technology needs, in situated settings, recognizing that laboratory studies are limited when it comes to investigating issues around real-world use.

In summary, there are many purposes for which SSQSs are well suited. There are others that demand other techniques, such as controlled laboratory studies or ethnography (in the sense discussed in section 1.4); what matters is that the study design suits the study purpose.

There are, of course, purposes that cannot, in practice, be addressed reliably, for legal, safety, privacy, ethical or similar reasons. For example, in safety-critical situations, the presence of researchers could be a distraction when conditions become demanding, so it may not be possible to study the details of interactions at

times of greatest stress. It is not theoretically or practically possible to address every imaginable research question in HCI.

The centrality of purpose is emphasised by Wolcott (2009, p.34), who advocates "a candidate for the opening sentence for scholarly writing: 'The purpose of this study is…'". While the purpose should drive the study design, and might evolve in the light of early study findings, it may also have to be crafted to fit the available resources.

4 Resources and constraints

Every study has to be designed to work with the available resources. Where resources are limited by, for example, time or budgetary restraints, it is necessary to 'cut your coat according to your cloth' – i.e., to fit ambitions and hence purpose to what is possible with the available resources.

Resource considerations need to cover – at least – time, funding, equipment available for data collection and analysis, availability of places to conduct the study, availability of participants, and expertise. In many cases, it is also necessary to have advocacy and support from people with influence in the intended study settings. Of these resources, three that merit further discussion are advocacy, participants and expertise.

4.1 Advocacy

Sometimes, studies are devised and run in collaboration with 'problem owners' (e.g. Randell et al, 2013), but other studies are conceived by a research team outside a particular domain setting. In some cases, it is essential to get support from a domain specialist.

For example, in the work of the author and co-workers, with a shift in emphasis in healthcare from hospital to home, we are interested in how medical devices are taken up and used in the home, and how products that were originally developed for use by clinical staff in hospitals can be adapted to be suitable for home use. There are some products that are well established for home use, such as nebulisers, blood glucose monitors and dialysis machines, and others that are making the transition from hospital to home, such as patient-controlled analgesia and intravenous administration of chemotherapy.

We followed several lines of enquiry to identify clinicians who expressed an interest in patients' experiences of intravenous therapies at home, but all eventually drew a blank. In contrast, we identified several nephrologists who were sufficiently interested in patients' experience of home haemodialysis to introduce us to their patients. This has led to a productive study (e.g. Rajkomar et al., 2013).

In other cases, it may be important to obtain permission to conduct a study in a particular location; for example, Perera (2006) investigated under what circumstances people forgot their chip-and-pin cards in shops, and she needed to obtain permission from shop managers to conduct observational studies on their premises.

In another case, in work for a Masters project (O'Connor, 2011), support was needed from the nursing manager in the emergency department; as soon as she was contacted, it was clear that she was keen to support others in their education, and this was the main factor in her supporting the project, rather that its inherent interest. In yet other cases, there is no particular advocate or manager to involve; for example, studies of diabetes patients (e.g. O'Kane & Mentis, 2012) involved direct recruitment of participants without mediation from specialists.

There may be a hidden cost in negotiating support from advocates, but this often brings with it the benefits of close engagement with the study domain, introductions to potential participants, and longer-term impact through the engagement of stakeholders.

4.2   Participants: recruitment and sampling

When recruiting participants for a study, with or without the advocacy of an intermediary, it is important to consider their motivations for participation. This is partly coupled with ethical considerations (section 5), and partly with how to incentivise people to participate at all. People may agree or elect to participate in studies for many different reasons: maybe it is low-cost (in terms of time and effort), and people just want to be helpful.

This was probably the case in the studies of ambulance control carried out by Blandford & Wong (2004). The immediate benefits to participants, ambulance controllers, were relatively small, beyond the sense that someone else was interested in their work and valued their expertise. However, the costs of participation were also low – continue to do your job as normal, and talk about the work in slack periods when you would otherwise simply be waiting for the next call.

In other cases, participants may be inherently interested in the project – as in some of our work on serendipity (Makri & Blandford, 2012) – or perceive some personal benefit, for example, in reflecting on how you manage your time (Kamsin et al., 2012). Some people may participate for financial reward, or to return a favour. And of course, people's motivations for participating may be mixed.

Where the topic is one that participants might be sensitive about, for example, intimate health issues, it can sometimes help to have pre-existing common ground between the researcher doing data gathering and the participant, such as being of the same sex or a similar age. Where multiple researchers are available, this might mean matching them well to participants; where there is a single researcher, it might mean reviewing the purpose of the study to be sure that data gathering is likely to be productive. In the section on ethics (section 5), we discuss relationships with participants and how these relate to recruitment and motivations for participating.

The choice of approaches to recruitment depends on the purpose of the study and the kinds of participant needed. Possible approaches include:

- Direct contact – e.g. approaching individuals in the workplace, with authorisation from local managers if needed, or approaching people in public spaces, with due regard for safety, informed consent, etc..
- Mediated contact: an introduction by someone else, such as a line manager in the workplace, another 'gatekeeper' – for example, a teacher, or the organiser of a relevant 'special interest' group – or friends or other participants.
- Advertising: on noticeboards in physical space, through targeted email lists, via online lists and social networks.

As social media and other technologies evolve, new approaches to recruiting study participants are emerging. What matters is that the approach to recruitment is effective in terms of recruiting both a suitable number of participants and appropriate participants for the aims of the study.

Two questions that come up frequently are how many participants should be included and how they should be sampled. The answer to both is 'it depends' – on the aims of the study, and what is possible with the available resources.

Although not common in HCI, it is possible to conduct a study with a single participant, as a rich case study. For example, Attfield et al. (2008) gathered observations, interview data and examples of artefacts produced from a single journalist as that journalist prepared an article from inception to publication. The aim of the study was to understand the phases of work, how information was transformed through that work, and how technology supported the work.

Such a case study provides a rich understanding of the interaction, but care has to be taken over generalizing: ideally, such a case will be compared with known features of comparable cases, in terms of both similarities and contrasts. In poorly understood areas, even a single rich case study can add to our overall understanding of the design, deployment and use of interactive technologies. But most studies involve many more participants than one.

Smith (2008, p.14) draws the distinction between idiographic and nomothetic research as follows:

> "The nomothetic approach assumes that the behaviour of a particular person is the outcome of laws that apply to all, and the aim of science is to reveal these general laws. The idiographic approach would, in contrast, focus on the interplay of factors which may be quite specific to the individual."

In other words, nomothetic research relies more on large samples and statistical techniques to establish generalizations, for example, through controlled experiments. SSQSs typically involve much smaller numbers of participants, occasionally as few as one, but more commonly 10-20, but gathers rich data with each. In this sense, SSQSs are idiographic, and care must be taken with generalizing beyond the study setting. This is a topic to which we return in section 10.1.

GT researchers resist specifying numbers of participants required. Rather, they advocate continuing to gather data until the theoretical categories of the analysis are saturated. Charmaz (2006, p. 113) explains: "Categories are 'saturated' when

gathering fresh data no longer sparks new theoretical insights, nor reveals new properties of your core theoretical categories". In other words, you stop gathering data when it no longer advances the study. This presupposes an iterative approach to data gathering and analysis, which is the case for GT, but not for other styles of qualitative research where all data may be gathered before analysis commences.

In practice, there are often pragmatic factors that determine how many participants to involve in a study. One might be the time available: it can take a long time to recruit each participant, to arrange and conduct data gathering, and analyse the data. Another might be the availability of participants who satisfy the recruitment criteria, for example, performing a particular role in an organisation or having particular experience. A shorter study with fewer participants needs to be more focused in order to deliver insight, because otherwise it risks delivering shallow data from which it is almost impossible to derive valuable insight.

Thought must be given to how to sample participants. Sometimes, the criteria are quite broad , for example, people who enjoy playing video games, and it is possible to recruit through public advertising. Sometimes, they are focused, such as on people with a particular job role within an organisation.

For other studies, the aim might be to obtain a representative sample; for example, in a study of lawyers' use of information resources (Makri et al., 2008), our aim was to involve lawyers across the range of seniority, from undergraduate students to senior partners in a law firm and professors in a university law department.

However, in qualitative research it is rare to aim for probability sampling, as one would for quantitative studies. Marshall (1996) discusses three different approaches to sampling for qualitative research: convenience, judgement (also called purposeful), and theoretical.

1. Convenience sampling involves working with the most accessible participants, and is therefore the easiest approach.
2. Judgment sampling, in which the "researcher actively selects the most productive sample to answer the research question" (p. 523), is the most commonly used in HCI.
3. Theoretical sampling is advocated within GT, and involves recruiting participants who are most likely to help build the theory that is emerging through data gathering and analysis.

Miles and Huberman (1994, p.28) list no fewer than 16 different approaches to sampling, such as maximum variation, extreme or deviant case, typical case, and stratified purposeful, each with a particular value in terms of data gathering and analysis.

An approach to sampling that can be particularly useful for accessing hard-to-reach populations, for example, people using a particular specialist device, is snowball sampling, where each participant introduces the researcher to further participants who satisfy their inclusion criteria.

Atkinson and Flint (2001) highlight some of the limitations of this approach, in terms of participant diversity and consequent generalizability of findings. Slightly tongue-in-cheek, they also describe "scrounging sampling": the increasingly desperate acquisition of participants to make up numbers almost regardless of suitability. While few authors are likely to admit to applying scrounging sampling as a strategy for recruitment, it is important to explain clearly how and why participants have been recruited and the likely consequences of the recruitment strategy on findings.

The same issue arises when there might be barriers to recruitment. Buckley et al. (2007) highlight the dangers of 'consent bias', whereby those with more positive outcomes are more likely to agree to participate in a study. Although most of the literature on consent bias relates to healthcare studies, there are similar risks in HCI studies, particularly where the technology under investigation is related to a sensitive personal issue, such as behaviour change technologies. Atkinson and Flint (2001) also highlight the risks of 'gatekeeper bias', where those in authority – for example, clinicians or teachers – filter out potential participants whom they consider less suitable.

In summary, when planning a study, it is important to consider questions of recruitment and relationship management:

- Who the appropriate participants are and how they should be recruited;
- Where and when to work with them in data gathering; and
- How (or whether) to engage with them more broadly from the start to the end of a study.

Throughout recruitment, study design and data analysis, it is important to remain aware of participants' motivations for participating, and their expectations of the outcome, whether this is, for example, the expectation of novel interaction designs, or simply to gain the experience of participating. These factors are addressed further in section 5.

When dealing with sensitive topics where people may have reasons for sharing or withholding certain information, or for behaving in particular ways, it is also important to be aware of motivations and their possible effects on the data that is gathered. Such considerations imply the need (a) to review data gathering techniques to maximise the likelihood of gathering valid data (see section 6) and (b) to reflect on the data quality and implications for the findings (see section 10.1).

4.3   Expertise of the research team

There are at least two aspects to expertise: that in qualitative research and that in the study domain.

There is no shortcut to acquiring expertise in qualitative research. Courses, textbooks and research papers provide essential foundations, and different resources resonate with – and are therefore most useful to – different people. Corbin and Strauss (2008, p.27) emphasise the importance of planning and practice:

"Persons sometimes think that they can go out into the field and conduct interviews or observations with no training or preparation. Often these persons are disappointed when their participants are less than informative and the data are sparse, at best."

Kidder and Fine (1987, p.60) describe the evolving focus of qualitative research: "A daily chore of a participant observer is deciding which question to ask next of whom." There is no substitute for planning, practice and reflecting on what can be learnt from each interview or observation session.

Yardley (2000, p.218), comments on the trend towards precisely defined methods:

"This trend is fuelled by the tendency of those who are new to qualitative research, and dismayed by the scope and complexity of the field, to adhere gratefully to any set of clear-cut procedures provided by proponents of a particular form of analysis."

As noted elsewhere, there is an interdependence between methods, research questions and resources; fixed methods have their place, but can rarely be applied cleanly to address a real research problem (Furniss et al., 2011a), and may sometimes be used as labels to describe an approach that could not, in practice, conform exactly to the specified procedure.

As well as expertise in qualitative methods, the level of expertise in the study context can have a huge influence over the quality and kind of study conducted. When the study focuses on a widely used technology, or an activity that most people engage in, such as time management (e.g. Kamsin et al., 2012) or in-car navigation (e.g. Curzon et al., 2002), any disparity in expertise between researcher and participants is unlikely to be critical, although the researcher should reflect on how their expertise might influence their data gathering or analysis.

Where the study is of a highly specialised device, or in a specialist context, the expertise of the researcher(s) can have a significant effect on both the conduct and the outcomes of a study. At times, naivety can be an asset, allowing one to ask important questions that would be overlooked by someone with more domain expertise. At other times, naivety can result in the researcher failing to note or interpret important features of the study context.

Pennathur et al. (2013, p.216) discuss this in the context of a study of technology use in an operating theatre:

"There was a possibility for bias and/or inconsistencies during identification of hazards in the [operating theatre] due to the involvement of observers with different expertise, and consequently the aspects that they may prioritise during observations."

Domain expertise may also cause the researcher to become drawn into the on-going activity, potentially limiting their ability to record observations systematically – effectively becoming a practitioner rather than a researcher, insofar as these roles may conflict.

In preparing to conduct a study, it is important to consider the effects of expertise and to determine whether or not specific training in the technology or work being studied is required.

4.4 Other resources

There will be other resources and constraints that create and limit possibilities for the research design. These include the availability of equipment, funding – for example, for travel and to pay participants –, time, and suitable places to conduct research. Here, we briefly discuss some of these issues, while avoiding stating the obvious – variants on the theme of "don't plan to use resources that you don't have or can't acquire!".

Where a study takes place can shape that study significantly. Studies that take place within the context of work, home or other natural setting are sometimes referred to as 'situated' or 'in the wild' (e.g. Rogers, 2012). Studies that take place outside the context of work include laboratory studies – involving, for example, think-aloud protocol – and some interview studies – those that take place in 'neutral' spaces.

There are also intermediate points, such as the use of simulation labs, or the use of spaces that are 'like' the work setting, where participants have access to some, but not all, features of the natural work setting.

Observational studies most commonly take place 'in the wild', where the 'wild' may be a workplace, the home, or some other location where the activity of interest takes place, that is, the technology of interest is used. Interview studies may take place in the 'wild' or in another place that is comfortable for participants, quiet enough to record and ensure appropriate privacy, and safe for both participant and interviewer. Of course, there are also study types where researcher and participant are at-a-distance from each other, such as diary studies and remote interviews.

Rogers et al. (2011, p.227) discuss the uses of data recording tools including notes, audio recording, still camera, and video camera. All of these can be useful tools for data recording, depending on the situations in which data is being gathered. For instance, still photographs of equipment that has been appropriated by users, or a record of the locations in which technology was being used or how it was configured, provide a permanent record to support analysis, and to illustrate use in reports. As an example, Figure 2 shows how a home haemodialysis machine was marked up to remind the user to change a setting every time the machine was used.

Screen capture software can give a valuable record of user interactions with desktop systems. Particular qualitative methods such as the use of cultural probes (Gaver and Dunne, 1999) or engaging participants in keeping video diaries, or testing ubiquitous computing technologies, may require particular specialist equipment for data gathering. When it comes to data analysis, coloured pencils, highlighter pens and paper are often the best tools for small studies. For larger studies, computer-based tools to support qualitative data analysis (e.g. NVivo or AtlasTI) can help with managing and keeping track of data, but require an investment of time to learn to use them effectively.
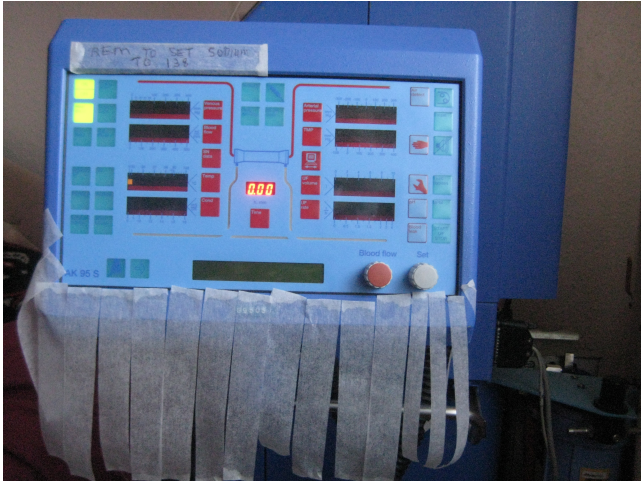
Figure 2: HHD machine with reminder "rem to set sodium to 136". The strips hanging at the bottom also show how the machine is being used as a temporary place to store cut strips for future use.

In addition to the costs of equipment, the other main costs for studies are typically the costs of travel and participant fees. Within HCI, there has been little discussion around the ethics and practicality of paying participant fees for studies. In disciplines where this has been studied, most notably medicine, there is little agreement on policy for paying participants (e.g. Grady et al., 2005; Fry et al., 2005). The ethical concerns in medicine are typically much greater than those in HCI, where the likelihood of harm is much lower. In HCI, it is common practice to recompense participants for their time and any costs they incur without making the payment, whether cash or gift certificates, so large that people are likely to participate just for the money.

5   Ethics and informed consent

Traditionally, ethics has been concerned with the avoidance of harm, and most established ethical clearance processes focus on this. 'VIP' is a useful mnemonic for the main concerns, **V**ulnerability, **I**nformed consent, and **P**rivacy:

- **V**ulnerability: particular care needs to be taken when recruiting participants from groups that might be regarded as vulnerable, such as children, the elderly, or people with a particular condition (illness, addiction, etc.).

- **I**nformed consent: where possible, participants should be informed of the purpose of the study, and of their right to withdraw at any time. It is common practice to provide a written information sheet outlining the purpose of the study, what is expected of participants, how their data will be stored and used, and how findings will be reported. Depending on the circumstances, it may be appropriate to gather either written or verbal informed consent; if written, the record should be kept securely, and separately from data. With the growing use of social media, and of research methods making use of such data, from, for example, Twitter or online forums, there are situations where gathering informed consent is impractical or maybe even impossible; in such

situations, it is important to weigh up the value of the research and means of ensuring that respect for confidentiality is maintained, bearing in mind that although such data has been made publicly available, the authors may not have considered all possible uses of the data and may feel a strong sense of ownership. If in doubt, discuss possible ethical concerns with local experts.

- **P**rivacy and confidentiality should be respected, in data gathering, management and reporting.

Willig (2008, p.16) lists informed consent, the avoidance of deception, the right to withdraw, debriefing, and confidentiality as primary considerations in the ethical conduct of research.

However, the work of the author and co-workers with clinicians and patients (Furniss et al., 2011b; Rajkomar & Blandford, 2012, Rajkomar et al., 2013) has highlighted the fact that ethics goes beyond these principles. It should be about doing good, not just avoiding doing harm. This might require a long-term perspective: understanding current design and user experiences to guide the design of future technologies. That long-term view may not directly address the desire of research participants to see immediate benefit.

What motivates an individual technology user to engage with research on the design and use of that technology? Corbin and Strauss (2008) suggest that one reason for participating in a study may be in order to make one's voice heard.

This is not, however, universally the case. For example, in one of our studies of medical technologies (Rajkomar et al., 2013), participants were concerned to be seen as experts – because they might have had rights withdrawn if they were not – so it was *not* a benefit to them to have a chance to critique the design and usability of the system. For such participants, it may be about the 'common good': about being prepared to invest time and expertise for long-term benefits. For others, there is an indirect pay-back in terms of having their expertise and experience recognised and valued, or of being listened to, or having a chance to reflect on their condition or their use of technology.

If the study involves using a novel technology, there may well be elements of curiosity, opportunities to learn, and experiencing pleasure in people's motivations for taking part. Some people will be attracted by financial and similar incentives. There are probably many other complex motivations for participating in research. As researchers, we need to understand those motivations better, respect them, and work with them. Where possible, researchers need to 'repay' participants and others who facilitate research, and manage expectations where those expectations may be unrealistic – such as having a fully functioning new system within a few months.

Finally, Rogers et al. (2011, p.224) point out that the relationship between researcher and participants must remain "clear and professional". They suggest that requiring participants to sign an informed consent form helps in achieving this: true in some situations, but not in others, where verbal consent may be less costly and distracting for participants.

6   Techniques for data gathering

The most common techniques for data gathering in SSQSs are outlined below: observation; contextual inquiry; semi-structured interviews; think-aloud; focus groups; and diary studies. The increasing focus on the use of technologies while mobile, in the home, and in other locations are leading to yet more ways of gathering qualitative data. As Rode (2011, p.123) notes:

> "As new technologies develop, they allow new possibilities for fieldwork – remote interviews, participant-observation through games, or blogs, or virtual worlds, and following the lives of one's informants via twitter."

The possibilities are seemingly endless, and growing. The limit may be the imagination of the research team.

Whatever methods of data gathering are employed, it is wise to pilot test them before launching into extensive data gathering – both to check that the data gathering is as effective as possible and to ensure that the resulting data can be analysed as planned to address the purpose of the study. If the study design is highly iterative (e.g. using Grounded Theory as outlined in section 1.5), then it is important to review the approach to data gathering before every data-gathering episode. If the data gathering and analysis are more independent, as in some other research designs, it is more important to include an explicit piloting stage to check that the approach to data gathering is working well: for example, to ensure that interview questions are effective or that participant instructions are clear).

6.1   Observation

There are many possible forms of observation, direct and indirect. Flick (2009, p.222) proposes five dimensions on which observational studies may vary:

- Covert vs. overt: to what extent are participants aware of being observed?
- Non-participant vs. participant: to what extent does the observer become part of the situation being observed?
- Systematic vs. unsystematic: how structured are the observation notes that are kept?
- Natural vs. controlled context: how realistic is the environment in which observation takes place?
- Self-observation vs. observation of others: how much attention is paid to the researcher's reflexive self-observation in data gathering?

In other words: there is no single right way to conduct an observational study. Indeed, the way it is conducted will often evolve over time, as the researcher's understanding of the context and ability to participate constructively and helpfully in it develop.

Flick (2009, p.223) identifies seven phases for planning an observational study:

- Selection of setting(s) for observation;
- Determining what is to be documented in each observation;

- Training of observers (see discussion on expertise, section 4.3)
- Descriptive observations to gain an overview of the context;
- Focused observations on the aspects of the context that are of interest;
- Selective observations of central aspects of the context;
- Finish when theoretical saturation has been reached – i.e., when nothing further is being learned about the context.

These phases – particularly selective observations and theoretical saturation – convey a particular view of observation as developing a focused theory, much in the style of Grounded Theory (section 1.5). Nevertheless, the broader idea of careful preparation for a study and recognition that the nature of observations will evolve over time are important for nearly all observational studies.

Willig (2008, p.28) discusses the nature of data gathering, including the importance of keeping detailed notes – such as near-verbatim quotations from participants and "concrete descriptions of the setting, people and events involved". She refers to these as "substantive notes", which may be supplemented by "methodological notes" – reflecting on the method applied in the research in practice – and "analytical notes", which constitute the beginning of data analysis (section 7). She also notes that data collection and analysis may be more or less tightly integrated – a theme to which we return in section 9.3.

6.2   Contextual Inquiry

Contextual inquiry (Beyer and Holtzblatt, 1998) is a widely reported method for conducting and recording observational studies in HCI, as a stage in a broader process of contextual design. According to Holtzblatt and Beyer (2013), "Contextual Design prescribes interviews that are not pure ethnographic observations, but involve the user in discussion and reflection on their own actions, intents, and values". In other words, contextual inquiry involves interleaving observation with focused, situated interview questions concerning the work at hand and the roles of technology in that work.

More importantly, Holtzblatt and Beyer (2013) present clear principles underpinning contextual design, and a process model for conducting design, including the contextual inquiry approach to data gathering. This includes a basic principle of the relationship between researcher and participants: that although the researcher may be more expert in human factors or system design, it is the participants who are experts in their work and in the use of systems to support that work.

Holtzblatt and Beyer (2013) present five models, flow, cultural, sequence, physical, and artefact, that are intermediate representations to describe work and the work context, and for which contextual inquiry is intended to provide data. Although contextual inquiry is often regarded as a component of contextual design, it has been applied independently as an approach to data gathering in research (e.g. Blandford and Wong, 2004).

6.3   Think Aloud

In contextual inquiry, the researcher is clearly present, shaping the data gathering through the questions he or she asks; in contrast, in a think-aloud study the researcher retreats into the background. Think aloud involves the users of a system articulating their thoughts as they work with a system. It typically focuses on the interaction with a particular interface, and so is well suited to identifying strengths and limitations of that interface, as well as the ways that people structure their tasks using the interface. Think aloud is most commonly used in laboratory studies, but also has a valuable role in some situated studies, as people demonstrate their use of particular systems in supporting their work (e.g. Makri et al., 2007).

Variants on the think-aloud approach are used in many disciplines, including cognitive psychology (e.g. Ericsson and Simon, 1980), education research (e.g. Charters, 2003) and HCI. Boren and Ramey (2000) conducted a review analysing the ways in which think aloud had been used in a variety of HCI studies, and conclude that, although most researchers cited Ericsson and Simon (1980) as their source for the method, the details of think alouds varied substantively from study to study.

Boren and Ramey (2000, p.263) highlight four key principles from Ericsson and Simon (1980) to which a think aloud study should conform:

1. Only 'hard' verbal data should be collected and analysed: "The only data considered must be what the participant attends to and in what order."
2. Detailed instructions for how to think aloud should be given: "Encourage the participant to speak constantly 'as if alone in the room' without regard for coherency." They also recommend that participants should have a chance to practise thinking aloud prior to the study.
3. If participants fall silent, they should be reminded succinctly to verbalise their thoughts.
4. Other interventions should be avoided, and attention should not be drawn to the researcher's presence. This is in stark contrast to the approach of contextual inquiry (section 6.2).

Norgaard and Hornbaek (2006) conducted a study of how think aloud methods are used in practice, observing studies in seven different companies. They noted (p.271) that many of the studies did not conform to the guidelines above. For example, they included questions about people's perceptions, expectations, and interpretations during the TA study; exhibited a "tendency that evaluators end up focusing too much on already known problems"; and "evaluators seem to prioritize problems regarding usability over problems regarding utility".

In other words, as with most other data-gathering techniques, there are in practice many different ways to go about gathering data, and these are shaped by the interests of the researcher, the purpose of the study, and the practicalities of the situation.

One aspect of think aloud that has received little attention is how participants are instructed – not just in how to think aloud, but also in the tasks to be performed.

Sometimes (e.g. Makri et al., 2007) these are naturalistic tasks chosen by the participants themselves, so that the researcher is essentially observing the participant completing a task that is part of, or aligned to, their on-going work. In other cases, tasks need to be defined for participants, and care needs to be taken to ensure that these tasks are appropriate, realistic and suitably engaging. While researcher-defined tasks are widely used in usability studies, they are less common in SSQSs, which are generally concerned with understanding technology use 'in the wild'.

6.4   Semi-structured Interviews

Think alouds are one way to gather verbal data from participants about the perceptions and use of technology; interviews are another widespread way of gathering verbal data. Interviews may be more or less structured: a completely structured interview is akin to a questionnaire, in that all questions are pre-determined, although a variety of answers may be expected; a completely unstructured interview is more like a conversation, albeit one with a particular focus and purpose. Semi-structured interviews fall between these poles, in that many questions – or at least themes – will be planned ahead of time, but lines of enquiry will be pursued within the interview, to follow up on interesting and unexpected avenues that emerge.

Interviews are best suited for understanding people's perceptions and experiences. As Flick (1998, p.222) puts it: "Practices are only accessible through observation; interviews and narratives merely make the accounts of practices accessible."

People's ability to self-report facts accurately is limited; for example, Blandford and Rugg (2002) asked participants to describe how they completed a routine task, and then to show us how they completed it. The practical demonstration revealed many steps and nuances that were absent from the verbal account: these details were taken for granted, so 'obvious' that participants did not even think to mention them.

Arthur and Nazroo (2003) emphasise the importance of careful preparation for interviews, and particularly the preparation of a "topic guide" (otherwise known as an interview schedule or interview guide). Their focus is on identifying topics to cover rather than particular questions to ask in the interview. It can be useful to have prepared important questions 'verbatim' – not because the question should then be asked rigidly as prepared, but because it identifies one way of asking it, which is particularly valuable if the interviewer has a 'blank' during the interview. Arthur and Nazroo advocate planning the topic guide within a frame comprising:

- Introduction;
- Opening questions;
- Core in-depth questions; and
- Closure.

This planning corresponds to the stages of an interview process as described by Legard et al. (2003), who present two views on in-depth interviewing. One starts from the premise that knowledge is 'given' and that the researcher's task is to dig it

out; although they do not use the term, this is in a positivist tradition. The other view is a constructivist one: that knowledge is created and negotiated through the conversation between interviewer and interviewee. Legard et al. (p.143) emphasize the importance of building a relationship, noting that the interviewer is a "research instrument", but also that researchers need "a degree of humility, the ability to be recipients of the participant's wisdom without needing to compete by demonstrating their own."

They present the interview process as having six stages, all of which need to be planned for:

1. Arrival: the first meeting between interviewee and interviewer has a crucial effect on the success of the interview; it is important to put participants at their ease.
2. Introducing the research: this involves ensuring that the participant is aware of the purpose of the research, and has given informed consent, that they are happy to have the interview recorded, and understand their right to withdraw.
3. Beginning the interview: the early stages are usually about giving the participant confidence and gathering background facts to contextualize the rest of the interview.
4. During the interview: the body of the interview will be shaped by the themes of interest for the research. Participants are likely to be thinking in a focused way about topics that they do not normally consider in such depth in their everyday lives.
5. Ending the interview: Legard et al. emphasize the need to signal the end so that the participant can prepare for it and ensure there are no loose ends.
6. After the interview: participants should be thanked and told what will happen next with their data. Many participants think of additional things to say once the recorder is off, and these may be noted. Legard et al. emphasise the importance of participants being "left feeling 'well'" (p.146), as discussed in section 5.

Legard et al. present various strategies for questioning, including the use of broad and narrow questions, avoiding leading questions, and making sure all questions are clear and succinct.

Within the core phase of interviewing, one technique to help with recall is the use of examples, asking people to focus on the details of specific incidents rather than generalizations. For example, the critical incident technique (Flanagan, 1954) can be used to elicit details of unusual and memorable past events, which in the context of HCI might include times when a technology failed or when particular demands were placed on a system.

A variant of this approach, the critical decision method (CDM), is presented in detail by Klein at al. (1989): in brief, their approach involves working with participants to reconstruct their thought processes while dealing with a problematic situation that involved working with partial knowledge and making difficult decisions. CDM helps

to elicit aspects of expertise that are particularly well suited to studying technology use in high-pressure environments where the situation is changing rapidly and decisions need to be made, as in control rooms, operating theatres, and flight decks.

Charmaz (2006) describes an "intensive interview" as a "directed conversation". Her focus is on interviewing within grounded theory (section 1.5), and on eliciting participants' experiences. She emphasizes the importance of listening, of being sensitive, of encouraging participants to talk, of asking open-ended questions, and not being judgemental. Although the participant should do most of the talking, the interviewer will shape the dialogue, steering the discussion towards areas of research interest while attending less to areas that are out of scope.

Charmaz emphasizes the "contextual and negotiated" (p.27) qualities of an interview: the interviewer is a participant in the shaping of the conversation, and therefore, the interviewer's role needs to be reflected in the outcome of a study. This is a theme to which we return in section 9.2.

6.5   Focus groups

Focus groups may be an alternative to interviews, but have important differences. The researcher typically takes a role as facilitator and shaper, but the main interactions are between participants, whose responses build on and react to each other's. The composition of a focus group can have a great effect on the dynamic and outcome in terms of data gathered. Sometimes a decision will be made to gather data through focus groups to exploit the positive aspects of group dynamics; at other times, the decision will be more pragmatic.

For example, Adams et al. (2005) gathered data from individual practising doctors through interviews, partly because doctors typically had their own offices (a location for an interview), but also because they had very busy diaries. Each interview, therefore, had to be scheduled for a time when the participant was available (and many had to be delayed or rescheduled due to the demands of work). However, they gathered data from trainee nurses through focus groups because the nurses formed a cohort who knew each other reasonably well and often had breaks at the same time, so it was both easier and more productive to conduct focus groups than interviews.

6.6   Diary studies

Diary studies enable participants to record data in their own time – such as at particular times of day, or when a particular trigger occurs. Diary entries may be more or less structured; for example, the Experience Sampling Method (Consolvo and Walker, 2003) typically requires participants to report their current status in a short, structured form, often on a PDA / smart phone, whereas video diaries may allow participants to audio-record their thoughts, with accompanying video, with minimal structure.

Kamsin et al. (2012) investigated people's time management strategies and tools using both interviews and video diaries. While interviews gave good insights into

people's overall strategies and priorities, the immediacy of video diaries delivered a greater sense of the challenges that people faced in juggling the demands on their time and of the central role that email plays in many academics' time management.

6.7 Summary

Analysis can only work with the data that is collected. Therefore, it is important to gather the best possible data, working within the resources of the project (as discussed in section 4). In some situations, data gathering and analysis are treated as being semi-independent from each other, with all analysis following the end of data gathering. In other situations, the two are interleaved – whether in the rich way advocated in GT, or by interleaving stages of data gathering and analysis as the study proceeds (e.g. as the theoretical focus develops, or as different data gathering methods are applied to address the problem from different angles – see discussion of triangulation in section 10.2).

7 Analysis

Most data for SSQSs exists in the form of field notes, audio files, photographs and videos. The first step of analysis is generally to transform these into a form that is easier to work with – such as transcribing audio, annotating or coding video. This may be done at different levels of detail; for example, selectively transcribing text that is directly relevant to the theme of the study through to a full transcription of all words, phatic utterances, pauses and intonations. The decision about which details to include should be guided by the purpose of the study, and hence the style of analysis to be completed. Some researchers choose to transcribe data themselves, as the very act of transcribing, and maybe making notes at the same time, is a useful step in becoming familiar with the data and getting immersed in it. Others prefer to pay a good typist to transcribe data, because, for example, they consider this a poor use of their time.
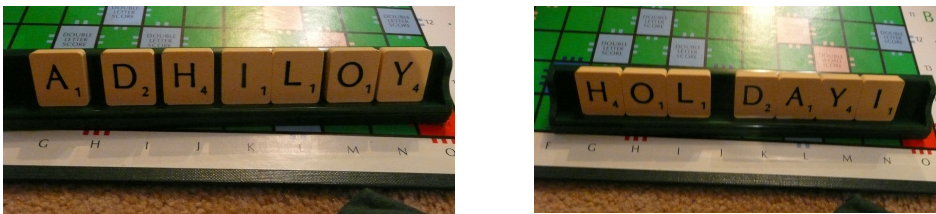
Figure 3: Reorganising Scrabble letters

Similarly, people make different decisions about which tools to use for analysis. Decisions may be based on prior experience – such as having used Qualitative Data Analysis (QDA) tools such as ATLASti or NVivo, on the size and manageability of the dataset, and on personal preference. Any tool creates mediating representations between the analyst and the data, allowing the analyst to 'see' the data in new ways, just as reorganising Scrabble letters can help the player to 'see' words they may not have noticed previously (Figure 3). One researcher may choose to use a set of tables in a word processor to organise and make sense of data; another might create an

affinity diagram (see Figure 4 for an example). Corbin and Strauss (2008, p.xi) emphasise that tools should "support and not 'take over' or 'direct' the research process", but note the value of tools in making analysis more systematic, contributing to reliability and an audit trail through the analysis.

Figure 4: Example of affinity diagramming with post-it notes, reproduced with permission from Stawarz (2012).

7.1   Different approaches to coding and iteratively analysing data

As noted above, an identifying feature of SSQSs is that they involve some form of coding of the data – i.e. of creating useful descriptors of units of data, such as single words, phrases, extended utterances, objects featuring in photographs, actions noted in videos, etc., and then of comparing and contrasting coded units to construct an analytical narrative based on the data. Grounded theory and thematic analysis (as outlined in sections 1.5 and 7.2) exemplify ways of coding data for analysis. There is a 'space' of approaches to coding qualitative data.

At one extreme, codes are simply 'buckets' in which to organise concepts identified in the data. Taking a utilitarian HCI approach, as a form of requirements gathering and system evaluation, CASSM (Blandford et al., 2008b; Blandford, in press) is a systematic approach to identifying the concepts that users are invoking when working with a system; where possible, these should be implemented in the system design (Johnson and Henderson, 2011).

A CASSM analysis involves gathering verbal data and classifying it in terms of user concepts. For example, in a study of ambulance control (Blandford et al., 2002), controllers were found to be working with two concepts both of which they referred to as 'calls': emergency calls being received; and the incidents to which those calls referred. The call management system they were working with at the time of the study allowed them to process emergency calls reasonably easily, but did not support incident management, which is important, particularly when a major incident occurs and many people call to report the same incident. A CASSM analysis

is an SSQS, but the data analysis is simple, being mainly concerned with coding concepts by assigning them to 'buckets'.

Based in a positivist social sciences tradition, Miles and Huberman (1994, p.64) advocate creating a preliminary list of codes prior to conducting fieldwork, and then refining it through analysis. They also advocate having multiple coders who can check that there is a shared understanding of codes to achieve "an unequivocal, common vision of what the codes mean and which blocks of data best fit which code."

At another point on the spectrum of approaches, Corbin and Strauss (2008, pp. 10 & 32) emphasise the centrality of the individual researcher in creating, interpreting and reporting the study:

> "Concepts and theories are *constructed* by researchers out of stories that are constructed by research participants who are trying to explain and make sense out of their experiences and/or lives. ………… Sensitivity stands in contrast to objectivity. It requires that a researcher put him- or herself into the research. Sensitivity means having insight, being tuned in to, being able to pick up on relevant issues, events, and happenings in data. It means being able to present the view of participants and taking the role of the other through immersion in data."

There are many available resources, such as those from the social sciences, that describe approaches to qualitative data analysis in detail. For example, Grbich (2013) presents over a dozen different approaches including: what she terms "classical ethnography", which is much more structured than the approach described in section 1.4; three variants of grounded theory; cyber ethnography, focusing on internet use; and various approaches for analysing existing data. The challenge for the HCI researcher is to navigate their way through the space of possibilities, understanding the theoretical perspectives from which different authors are writing and constructing their own approach that is appropriate to the research question at hand, their own biases and competencies, and the resources available.

As well as the various approaches to GT, thematic analysis can be a valuable approach to analysing qualitative data, and exemplifies more of the space of possible approaches to analysis.

7.2   Thematic Analysis

In contrast to GT, where data collection and analysis are interleaved, thematic analysis assumes that a dataset already exists, and focuses attention on how that data might be analysed. Braun and Clarke (2006) argue that "thematising meanings" is a generic skill across qualitative methods and that thematic analysis builds directly on this skill. They contrast thematic analysis with qualitative techniques such as conversation analysis or interpretative phenomenological analysis, which are founded on a particular theoretical position and are typically applied in relatively tightly defined ways, but are rarely used in HCI. Rather, Braun and Clarke place thematic analysis in a 'camp' of techniques that can be applied across a range of theoretical positions, and that tries to steer a path between 'anything goes'

unstructured analysis and an approach that is overly constrained. They make the obvious but important point (p.80), "What is important is that the theoretical framework and methods match what the researcher wants to know, and that they acknowledge these decisions, and recognise them *as* decisions."

Braun and Clarke identify six phases of thematic analysis (pp. 87-88):

1. Familiarising with the data: simply reading and re-reading the data, making notes of ideas that spring to mind.
2. Generating initial codes: coding the entire dataset systematically and collating data that is relevant to each code. They define codes as labels that "identify a feature of the data (semantic content or latent) that appears interesting to the analyst".
3. Searching for themes: gathering codes (and related data) into candidate themes for further analysis.
4. Reviewing themes: checking whether the themes work with the data and creating a thematic "map" of the analysis.
5. Defining and naming themes: refining the themes and the overall narrative iteratively.
6. Producing the report: which will, in turn require a further level of reflection on the themes, the narrative and the examples used to illustrate themes.

These phases represent an approach to iteratively deepening engagement with the data through layers of analysis.

Consistent with their overall flexible approach, Braun and Clarke (2006, pp.88-89) are not prescriptive about whether an analysis should be informed (or driven) by a particular theory, or whether it should be driven by the analyst's interpretation of the data:

> "Coding will, to some extent, depend on whether the themes are more 'data-driven' or 'theory-driven' – in the former, the themes will depend on the data, but in the latter, you might approach the data with specific questions in mind that you wish to code around."

This is a theme to which we return in section 9.1.

8   Reporting

As with any writing, the reporting of an SSQS has to be appropriate to the audience. If the study has been commissioned to deliver findings rapidly as part of a commercial development process, the reporting should be appropriately succinct and focused, whereas if it is part of a PhD thesis or other large academic project the reporting is more likely to focus on novel contribution and relationship to theory and previous literature.

There are many texts dedicated to the topic of how to write – whether in terms of the practicalities of getting and staying motivated, or of structuring text, or of structuring argument and addressing the intended audience. Thimbleby (2008) encourages the reader to write early and often, to draft and redraft, not to expect the first version to

be the best. He observes that we are implicitly taught to write just once by tight deadlines and that the practice of writing and rewriting that is essential to refining ideas and communicating them effectively is a difficult one to develop.

The same principle is advocated, equally strongly, by Wolcott (2009), who goes so far as to discuss "shitty first drafts" (p.51), to be followed by better second drafts and excellent third drafts. Writing is much improved by getting feedback from others, so it is helpful to get into the practice of getting feedback even for early drafts. In some cases, it can be particularly helpful to get feedback – of a draft that is not too 'shitty' – from study participants, as a form of validation (see section 10.2).

Within some research traditions, there are recognised structures that are widely conformed to; for example, in the sciences, a standard format is: aims, background, method, results, discussion, conclusion – and this format is sometimes advocated in traditions such as design where the material does not fit so naturally into this shape.

For qualitative studies, Wolcott (2009) argues strongly that this is not an effective structure, because presentation of background material delays the presentation of the key substance of the study. He argues that only essential background material should be included as part of the introduction, and that other related work should be introduced as needed through the narrative.

There is no one correct approach to structuring, and it can certainly be very challenging to fit the reporting of SSQSs into the standard 'scientific' structure. Unlike most quantitative research, where the researcher's understanding of the problem is unlikely to change much during a study, unless the hypothesis is poorly founded or the method inadequately planned or executed, during an SSQS the researcher is likely to learn much about the problem, and to 'see' it in different ways as understanding matures (Furniss et al., 2011a).

To take a simple example: a researcher doing a situated study in an unfamiliar environment is learning about the study context – beyond what can be read in published material about it – while doing the study, and yet the details of the context are part of the background to the research, and not, usually, research findings. The boundaries between method of analysis and results, between results and discussion, and between discussion and conclusions can seem just as blurred, particularly as understanding deepens through iterations of analysis.

If the final understanding and all the literature that relates to that understanding is presented up-front, the actual findings can seem underwhelming, even though they were not anticipated at the beginning. In such cases, it is often valuable to take the reader through highlights of the journey that the researcher has travelled so that the reader is exposed to some of the delight of discovery that the research team experienced – assuming that the researchers started from a sensible place.

For example, one study of the author's started with the purpose of understanding how underground train controllers use technology and work together, with the intention of conducting Distributed Cognition analyses of different control rooms to understand variability in design and practices. As data gathering proceeded, it became clear that commonalities were much greater than contrasts, and that a more

interesting question was how the culture and use of technology has evolved to maintain safety. We, the researchers, decided to focus the background section of the report (Smith et al., 2009) on principles of train control, based on both literature and our early data gathering, and then to contextualise our findings in terms of the literature on resilience (such as Rochlin, 1999).

Understanding can develop, both as further data is gathered (e.g. Charmaz, 2006) and as new theoretical perspectives are encountered as ways of making sense of the data (e.g. Furniss et al., 2011a). Braun and Clarke (2006, p.80) make an important point in noting that an "account of themes 'emerging' or being 'discovered' is a passive account of the process of analysis, and it denies the *active* role the researcher always plays in identifying patterns/themes".

Again, this highlights the fact that there are alternative ways of reporting, depending on the role(s) that the researchers perceive themselves as having played in the research process. Bringing the researcher into the narrative makes explicit their role, which may make the research findings seem less objective or authoritative than a more 'distanced' account.

Within HCI, the highly personalised account is rare, as it can seem to be at odds with the expectation that one is delivering an account that is appropriately objective to inform design. And yet there may be times, as, for example, when delivering rich accounts of user experience to help designers 'put themselves in the users' shoes', when such a personalised account is both more honest and more effective than a depersonalised one.

It is possible for accounts to be 'too honest', if that results in participants being disadvantaged through their participation in the study in any way. Lipson (1997) highlights pitfalls of reporting that apply generally to qualitative studies. In the context of HCI studies, the issues may be more focused around whether any non-participant readers would be able to identify any participants by reading the account, and how participants might feel about the way their contribution has been reported. For example, if a study includes a focus on errors that people make with technology then it needs to be reported in a way that does not make participants feel either stupid or vulnerable.

It is also possible to be too honest in reporting the journey at such a fine-grained level of detail that the reader is bored and cannot discern important information from trivial details. It is important to be accountable while presenting the study at an appropriate level of abstraction.

In summary, there is no one correct way of reporting a qualitative study, in terms of 'voice' (e.g. to what extent the researcher is present in the narrative) or structure. The researcher should understand what is possible and what disciplinary rules they are violating if they choose to write in an unconventional way. As Wolcott (2009, p.66) puts it, "Before you begin to rock the boat, make sure you are in it". What matters is that:

- the purpose of the account is clear, and that the account focuses on the purpose;

- essential information is presented, such as what was actually done (rather than delivering textbook accounts of methods), while respecting participants and their confidentiality;
- it addresses the intended audience (whether this be practitioners, other HCI researchers or specialists in the domain of the study);
- it is related well to relevant prior work, so that it is clear what is novel about this study;
- the findings are presented at a level of abstraction such that the novel contribution and the extent to which the findings generalize to other settings are clear; and
- it is coherent as a narrative.

As noted above, it is almost impossible to get writing right first time, and an iterative process of drafting, getting feedback from others, re-reading the draft critically (preferably after a break, to gain some distance from it), and re-drafting is essential. It is also important to know when to stop, though, because perfection is unachievable!

It is also worth considering whether there are multiple audiences or angles from the same study, which may be written up separately (while avoiding self-plagiarism by making sure that multiple reports address different questions within the overall study purpose). An informal test of self-plagiarism is whether each paper can cite the other and be clearly different.

Reporting multiple angles separately can be particularly advantageous when each paper needs to be fitted within a tight word or page limit. Tight constraints can, in practice, be very helpful for communicating effectively as it forces the author to think about what really matters in the narrative, to omit spurious information, and to write succinctly. However, writing well takes time: Pascal is widely credited with the apology, "I would have written a shorter letter, but I did not have the time."

9 Factors that shape a study

As well as resources, constraints and ethical considerations, there are various less tangible factors that also shape any study. These include the way that pre-existing theory can be used to inform data gathering, analysis, and/or reporting of a study, and also the biases, understanding and experience of the researcher(s) involved in the project. These and other factors together create a web of interdependencies.

9.1 The role of pre-existing theory in data gathering, analysis and reporting

No researcher is a *tabula rasa*: each comes to a study with pre-existing understanding, experience, interests, etc. Hertzum et al. (2001) consider this to be "chilling": that there is no objective, shared understanding, even with an activity as superficially simple as identifying usability difficulties from think-aloud data. If this is true for analysing pre-determined data with a pre-defined question, it clearly has an even greater effect on the research that is conducted if the researcher is shaping

the entire study. For the individual, it may be difficult to identify or articulate many of the individual factors that shape the research they conduct, but one obvious factor is the role of theory in an SSQS. Theory may be most prominent towards the end of research project, may come into play during the analysis, or may shape the research from the outset.

Morse (1997) argues that the role of qualitative research is to deliver theory, whereas the role of quantitative research is to test theory. This is consistent with the focus of GT, in which theory is 'grounded' in data, and of thematic analysis, in which theory 'emerges', or is discovered / created, from data through analysis. Furniss et al. (2011a) present an account of a study of human factors (HF) practitioners' practices that includes an example of this relationship with theory. Based on the findings from a series of semi-structured interviews, a high-level theory was developed around the idea of 'downstream utility', which was seeded by the work of Wixon (2003) and developed into the use of a flowing river metaphor for describing how context-shaping factors influence the flow of a HF project.

Other researchers may seek a theoretical framework to help them make sense of data that seems very interesting but specific to the context of study. For example, Furniss et al. (2011a) were already familiar with the theory of distributed cognition (DC: Hollan et al., 2000; Furniss & Blandford, 2006). Although DC was not used for structuring data gathering, we thought it would be a useful framework for thinking about information flows in a HF project, providing a 'theoretical lens' on the analysis.

In contrast, Adams et al (2005) had to explicitly search for a theory to account for their findings. We had studied several different digital library (DL) deployment projects and found that making DLs more accessible to healthcare practitioners, by making them available through shared computers in the workplace, reduced their use when it was expected to increase it. Conversely, a project that had placed clinical librarians as members of multi-disciplinary care teams had increased use of DLs.

We explored theories such as DC and activity theory (Kaptelinin, 2013), but these did not help in accounting for our data. After some searching, we came across the theory of communities of practice (Wenger, 1998), which resonated with our data. This theory helped us to make sense of the data in a way that moved us from some interesting but idiosyncratic findings that were only relevant to our particular study contexts to findings that had some generalizability, and hence could be applied in other settings where new technology was being deployed.

Finally, others may intentionally structure data gathering and analysis around a particular theoretical framework, or are so steeped in a particular theoretical perspective that their approach to both data gathering and analysis is shaped by that perspective. For example, when studying people's strategies for information seeking, Makri et al. (2008) shaped their approach to data gathering and analysis around the work of Ellis et al. (1993, 1997). Where this is done, it is important not to trust an existing theoretical framework unquestioningly, but to test and extend that framework: are there counter-examples that challenge the accuracy of the existing

framework? Are there examples that go beyond the framework and introduce important extensions to it?

In summary: theories, of different kinds, can serve useful roles in SSQSs: in structuring the gathering and/or analysis of data and reporting of findings. A theory can be a 'lens', providing 'sensitizing concepts' that impose a partial structure on data that is gathered, helping to shape and focus data gathering. Similarly, a theory can help in shaping analysis, suggesting initial codes for analysing the data. In both of these cases, it is important not simply to accept a theory, but to test it, looking for evidence that might extend or contradict the established theory, while being mindful that there has to be a balance between power and generality in any theory. Make a theory too general and it typically loses analytical power, so not every extension to a theory is valuable.

9.2   The role of the researcher

A consideration that crops up repeatedly but is rarely discussed explicitly in reporting SSQSs is the role of the researcher, and the degree to which the researcher shapes the data gathering as it happens.

As noted in section 4.3, the training and expertise of the researcher can have a significant influence on their role, particularly in terms of what data they are most sensitised to, and hence alert and responsive to. This will include their training in the research methods, in the principles and practices of using the technology of interest, in particular theories, and in the practices of participants, such as details of particular job roles).

Some kinds of studies, such as diary studies and think-aloud studies, typically involve relatively little engagement between researcher and participant, so that once the study is initiated the researcher is reliant on having designed it well and on participants recording data as anticipated, which is why good pilot testing is advisable!

Such studies are relatively easy to describe in terms of how participants were instructed. Interventions by the researcher may be planned, and may 'nudge' data gathering, but on the whole the approach does not evolve significantly during the study, and the role of the researcher is limited – it is, for example, likely that substituting one researcher for another would have little effect on what data is gathered. Similarly, an analysis of documents or online forum data, for instance, allow the researcher to be objective, at least in regard to the data gathering since all they are doing is selecting the data of focus.

The same is true of some interview and observational studies, particularly those that are relatively structured. It is important to be aware of, and reflect on, the effect that observation may have had on participants' behaviour. Heisenberg showed that even at a subatomic level the act of observing changed the state of the thing observed; it is not possible to plan the perfect study of a situation without both influencing and being influenced by that situation. The very act of observing, therefore, may influence participants' behaviour.

In interviewing, the role of the researcher is likely to be smallest when the interview is structured. Some observational studies involve the researcher acting as a 'fly on the wall', trying to minimise the effect of their presence on the activity being observed. It may be a reasonable approximation to assume that the presence of the researcher has little influence on the data that is gathered, although it is important to reflect on the likelihood that observational factors such as the Hawthorne effect, in which participants were found to perform better when being observed (Roethlisberger & Dickson, 1939), might have an impact on findings.

For such studies, it is simplest and clearest to take an objective view. If, for example the purpose of the study is to understand how a particular group of professionals use technology to support their working practices, and the implications therefore for design, then the data gathered might emphasise some features of the work and downplay others. It is unlikely, however, that the presence or behaviour of the researcher has a major influence on how participants behave or what they report.

For example, in our study (sections 4.2 & 7.1) of how ambulance controllers use technology to maintain awareness of the situation, both within the control room and in the outside world of ambulances and incidents (Blandford and Wong, 2004), it seemed reasonable to assume that the way we related to study participants had little influence on their performance as professionals. The way we collected, analysed and reported data therefore downplayed our role as researchers in that process. The focus was on being systematic and thorough in data gathering and analysis, and transparent in reporting, so that the reader could trace how conclusions related to data.

Such a view places the researcher outside the study setting as an objective observer of phenomena, assuming that those phenomena have a truth that is independent of the role of the researcher in gathering data.

More active participation brings the researcher into the frame, and increases their influence on the data being gathered. This is most obvious in studies involving action research, in which the researcher is intentionally intervening to assess the effects of interventions on perceptions, processes and outcomes (Kock, 2013). It is also likely to be the case where the researcher acts as a participant observer, playing an active role within the study context, and to a lesser extent in approaches such as Contextual Inquiry (Holtzblatt & Beyer, 2013), which bring the researcher into the observation / interview space, though data gathering is still shaped mainly by the activities being performed.

In other studies, the researchers and their relationship with participants is central to the research process: the relationship has a strong influence over what information is shared and how it is shared by participants, how it is interpreted by the researcher, and how it is reported and may be interpreted by the reader. For example, Rode et al. (2004) discuss their approach of exploring families' use of programmable technologies in the home by using fuzzy felt props as being "provocative", aiming to establish "rich dialog" with participants.

Semi-structured interviews inevitably bring in the interests of the researcher as well as the participant. To pretend that they are objective is to downplay the individuality of each researcher and of the relationship between researcher and participant. Willig (2008) emphasises the role of the interview as a dialogue between people. Where the interview strays into potentially sensitive areas, such as negative feelings around technology use, it is surely unethical to remain artificially detached from the setting. In such situations, it is impossible to substitute one researcher for another: the researcher is effectively a research instrument.

Where the research is truly exploratory, it is impossible to plan all the details of the study ahead of time and get them all right: the details have to evolve as understanding of the context and subject matter matures. This evolution is made explicit in the processes and ethos of Grounded Theory, but applies equally to other SSQSs that do not follow all the principles of GT. Such constructivist research demands reflexivity of the researcher in data gathering, analysis and reporting.

9.3 A web of considerations

As should already be apparent, there are many connections and interdependencies when designing, conducting and reporting SSQSs. And these phases of work are not generally distinct. Through engaging with the study setting, the researcher learns more about what is possible in terms of data gathering, and more about the nuances of the research question, so the purpose of the study may change, at least in subtle ways, as understanding evolves.

Unlike most quantitative studies, which can conveniently be treated as starting with a hypothesis and finishing with a conclusion – even if the truth is not quite that simple – many SSQSs are effectively journeys, in which the researcher travels alongside the participants, making discoveries that are shared through the reporting of the study. The focus for data gathering and analysis, therefore, may change, shaped by current understanding as the study proceeds. Furthermore, as discussed in earlier sections, the study is shaped by the individuals – researchers and participants – engaged in it, by any extant theory that is exploited at any stage in the study, by resources and constraints, and by ethical considerations.

Figure 5 shows an unrealistically simplified research process (which is the one often promoted by traditional reporting structures), highlighting the key stages of planning based on the purpose of the study: data gathering; analysis and reporting. All three are shaped by the expertise and understanding of the research team, any extant theory, ethical considerations, and resources and constraints, which typically have the greatest impact on data gathering.
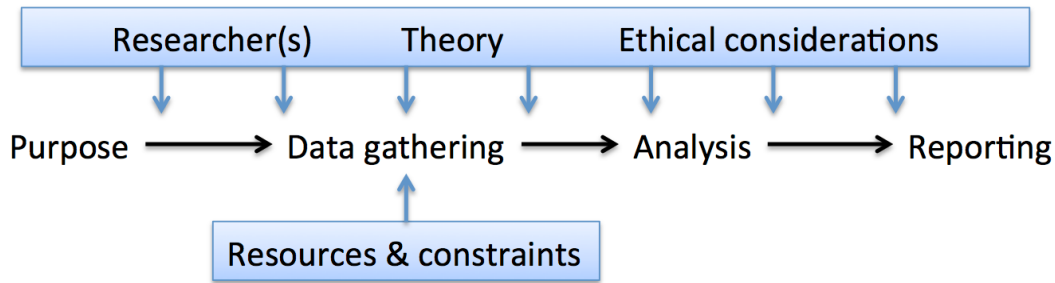
Figure 5: an idealised process based on aims, methods, results, and discussion

Figure 6 shows a process that is slightly closer to reality, with feedback and evolution in all stages as a study progresses, but still shaped by the same external factors. Data gathering and analysis may be more or less closely coupled. Early analysis may lead to revisions in the purpose of the study. The process of reporting often leads to new understanding of the problem. The overall purpose may be broken down into sub-questions that are best addressed through complementary studies involving different data gathering and analysis methods. These studies may be reported singly or together. Described in this way, the process can appear complicated and daunting, but in any particular instance the space of possibilities at any moment is not very great so, while every study is unique, it does not need to be unmanageable.
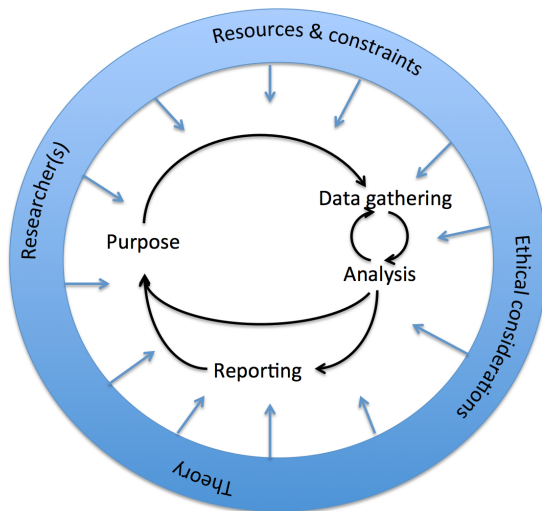
Figure 6: closer to reality: a journey shaped by many factors

One further factor that can have a great influence on findings is the participants, their motivations for taking part, and hence the data that is gathered in a study. The purpose of the study will determine who are ideal or possible participants, which may relate more-or-less directly to people's likely motivations for participating. This in turn should shape, and be shaped by, the recruitment strategy. Participants will shape what data gathering and validation is possible and hence the quality of data analysis, which will determine the actual outcomes of the study. These outcomes should address the purpose of the study (Figure 7). As discussed above, all of these stages will also be constrained by ethical considerations.
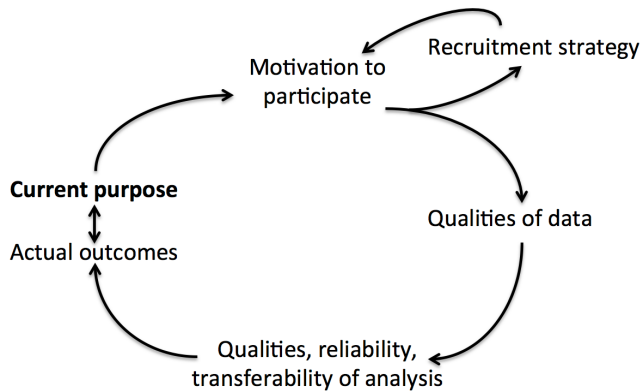
Figure 7: interdependencies between the purpose of a study, recruitment of participants and outcomes – which should match the purpose.

10 Assessing and ensuring quality in qualitative research

One of the challenges for qualitative researchers in HCI is that there is little consensus on what constitutes quality in qualitative research. Many reviewers adopt a particular stance, such as positivist or constructivist/interpretive, and immediately criticise research that does not conform to the expected paradigm. Arguably, on the one hand it is incumbent on the authors of a qualitative paper to present their approach and the rationale for it clearly, while on the other hand the reviewer has a responsibility to have appropriate expertise or an open mind, or to decline to review.

10.1 Quality criteria for constructivist research

In quantitative research, there are widely agreed criteria for quality such as internal validity, concerned with whether the experiment was properly conducted without confounding variables, and external validity, concerned with the generalizability of results. Kidder and Fine (1987, p.58) highlight some of the challenges in agreeing criteria in qualitative research by drawing an analogy between the work of a biographer and a qualitative researcher, quoting a psychohistorian: "When two quantitative researchers arrive at the same conclusions, we call it 'reliability,' but when two biographers write the same story we call it 'plagiarism'."

Implicitly, novelty and interest are assessment criteria in the latter case, though presumably other criteria, such as being justified on the basis of the available evidence, also apply. Within the space of SSQSs, there are different possible evaluative criteria.

Yardley (2000, p.219) proposes four essential characteristics of good qualitative research, which are also echoed by other authors:

- *Sensitivity to context*: e.g., taking account of previous relevant research, as well as 'listening' deeply to participants' perspectives and being sensitive to ethical considerations. Klein and Myers (1999) emphasise the importance of enabling the reader to comprehend fully the context of the research.

- *Commitment and rigour*: e.g., engaging well with the topic and with participants, completing a thorough data collection, and conducting a thorough analysis.

- *Transparency and coherence*: e.g., making it clear how data was analysed and conclusions drawn. Similarly, Henwood and Pidgeon (1992) advocate keeping close to the data so that the link between data and conclusions is clear, and maintaining a 'paper trail' that is open to external audit to expose the layers of analysis.

- *Impact and importance*: e.g., articulating clearly both the theoretical and practical significance of findings. In HCI studies, this may, but not necessarily, include "implications for design" (Dourish, 2006). It may also include insight – that the study helps to understand work, interaction or experience with technology in a new way. Klein and Myers (1999) argue that importance is achieved through *abstraction and generalization* – i.e. relating the particulars of the study to general principles. Henwood and Pidgeon (1992) focus on *transferability*, arguing that researchers should report on the contextual aspects of the study that allow the reader to assess the sphere of relevance of findings. Transferability is similar, but not identical, to the idea of generalizability, being more focused on the question of how readily the findings from one study can be applied to a different context.

Focusing on constructivist research, Henwood and Pidgeon (1992, p.105-108) list additional quality criteria, including:

- *Reflexivity*: the role of the researcher in the research should be recorded, and made apparent as appropriate. Klein and Myers (1999) advocate critical reflection on how data is "socially constructed" between researchers and participants.
- *Theoretical sampling and negative case analysis*: selecting cases that do not fit an emerging conceptual system helps to challenge assumptions. Henwood and Pidgeon consider this to be closely related to the constant comparative analysis advocated by Glaser, Strauss and others within the GT tradition. As well as being sensitive to contradictory evidence, Klein and Myers (1999) note the importance of being open to multiple interpretations (e.g. contradictory views of the same situation from different participants), yielding multiple narratives.

Coming from Information Systems research, Klein and Myers (1999) present principles based on a philosophical rationale. As well as echoing many of the criteria above, they also highlight the importance of:

- *The Hermeneutic Circle:* by this, they mean recognising that understanding is achieved by iterating between a focus on details and an understanding of the whole (similar to looking after the GOST of the study).
- *Suspicion:* i.e. being sensitive to possible systematic distortions in participants' narrative, for instance deriving from the way participants were recruited, or shaped by people's motivations for participating in the study.

These criteria for quality all depend on the researcher conducting the data gathering, analysis and reporting rigorously and honestly, and presenting the process with clarity and transparency.

10.2 External validation: inter-rater reliability, triangulation and respondent validation

There are also various approaches that give external validation of an analysis, which may be appropriate and feasible under some circumstances. These include employing multiple coders, triangulation of data sources, and respondent validation. These methods are typically built into the study design where they are used.

For some studies, the use of multiple coders, and maybe also measuring inter-rater reliability, is relevant. Miles and Huberman (1994, p.11) emphasise the importance of conclusions being verified, whether by reference back to field notes, achieving "intersubjective consensus" through discussion with colleagues, or replicating findings in another dataset. Miles and Huberman (1994) focus on the agreement of codes between multiple analysts – an approach that can be validated through measures of inter-rater reliability if coding is done independently.

Pennathur et al. (2013, p.207) work in a similar tradition, developing an approach to analysis that involves achieving group consensus for reconciling discrepancies between coders rather than computing inter-rater reliability. This requires that a set of codes has been previously agreed; in their case, these were based on the SEIPS model (Carayon et al., 2006) – i.e., on a particular theoretical perspective.

Having multiple independent coders of data and checking inter-rater reliability is appropriate for studies where codes and their meanings have been agreed and where the analysis and reporting relies heavily on those codes. It is not an appropriate way to validate a rich interpretive analysis.

In other situations, including many constructivist studies, it is possible to employ triangulation, which involves comparing multiple data sources or different methods of gathering data to corroborate findings. Mackay and Fayard (1997) argue that triangulation across scientific and design disciplines (introducing methods and theories from both) is particularly valuable in HCI. Rogers et al. (2011, p.225) list four different approaches to triangulation:

- *Triangulation of data*: data from different sources is compared; this helps with assessing the generalizability of findings.
- *Investigator triangulation*: different researchers collect and interpret the data; this is like the use of multiple coders as advocated by Miles and Huberman (1994).
- *Triangulation of theories*: using different theoretical frameworks as lenses on the data or findings.
- *Methodological triangulation*: employing different data gathering techniques can help to ensure that the outcome is not a simple function of the way that data was gathered.

Mays and Pope (2000) propose that, rather than supporting validation directly, triangulation encourages a more reflexive analysis of the available data. It depends on what form of triangulation is adopted as to how it can support data validation and give greater confidence in the findings, as outlined by Rogers et al. (2011).

Another widely discussed and used approach is respondent validation, or 'member checking,' in which study participants are invited to review the study findings to validate the researchers' interpretation of the data. A variant on this is to have other representatives of the same group – i.e., people like the participants – review the findings. While some (e.g. Lincoln and Guba, 1985) regard this as a strong check, others (e.g. Mays and Pope, 2000) highlight weaknesses in the approach, including dealing with discrepancies in the responses of participants, which effectively represent new data to be analysed, and managing the different priorities and focuses of participants and researchers.

Rather than conducting standard respondent validation, Henwood and Pidgeon (1992) suggest that negotiating interpretations with participants may sometimes be an effective approach to validating interpretations. However, they also recognise that neither of these approaches is universally applicable – as when, for example, participants have reason to object to a particular interpretation of the data.

A further, informal, check is face validity: do the findings of the study make sense? Are they credible? On its own, face validity is a very weak test, and should always be viewed with a critical eye, but the converse can be helpful: findings that lack face validity are rightly viewed with suspicion.

Barbour (2001) suggests that in healthcare research there is a tendency towards what I would term a 'checklist mentality': what she calls "technical fixes" are being required by funders and/or reviewers to ensure the rigour of qualitative research. She highlights five such fixes: purposive sampling; grounded theory; multiple coding; triangulation; and respondent validation. For each, she discusses the potential benefits: reducing bias; supporting original theorising; enhancing inter-rater reliability; checking internal validity; and checking researchers' interpretations, respectively. However, she also highlights pragmatic limitations of each approach in practice, and argues (p.1117), "They can strengthen the rigour of qualitative research only if they are embedded in a broad understanding of qualitative research design and data analysis."

10.3 Building quality and value indicators into reporting

In HCI, the question of transferability, generalizability or scope of findings is important: if design decisions for future systems are to be based on findings from qualitative studies, there has to be confidence in their broader applicability, or an understanding of how broad their applicability is. Some confidence in generalizability can come from relating findings to established theory or by triangulating findings across different data sources (section 10.2): if the findings from the current study are consistent with those from other studies – whether

represented directly in their findings or through theory that abstracts from findings – then that is one source of confidence.

If findings differ in interesting ways from theory or previous studies then that merits further discussion: is this because of some important difference in study conditions such as taking place in a different kind of setting, or working with a different user population? Otherwise, how else might the discrepancy be accounted for? Alternatively, particularly where there is no relevant prior theory or data, the findings from a qualitative study might indicate the need for further research to test those findings.

The quality of studies varies for many reasons, often linked to what is possible with the available resources, including the experience and expertise of the researcher(s), the time available, or the ease of recruiting an appropriate group of participants. While it is obviously important to conduct the best possible study with the resources available, it is also important to report findings in a way that makes it possible for the reader to assess the quality of the research. The reader should be able to answer questions such as:

- What confidence do I have in the results and conclusions of this study? What is the evidence to support my judgement?
- What can I learn from this study (relative to what was known before)? What is novel?
- How can I build on this study? – whether on the methods, the findings, or gaps in knowledge that it has exposed.

For reviewers of papers reporting SSQSs, the question is perhaps more basic: is this paper worth publishing? It is probably impossible to conduct a 'perfect' qualitative study: with more resources, it is almost always possible to do a better job. So the question is: what's good enough?

In the preceding sections, I have outlined some of the dimensions on which qualitative studies vary. In this section I have reviewed some perspectives on quality in qualitative studies, emphasising that there is not a 'one size fits all' approach, but that methods and approaches to quality control and validation should be coherent, and appropriate to the purpose, resources and methods of the study.

11 A checklist for designing and reporting SSQSs

Drawing all these themes together, we can identify a 'space' of considerations for designing, conducting and reporting SSQSs. The actual details of a study design involve nuanced interrelationships between design aspects, and only some possible combinations of design decisions are coherent. Established named methods, such as Grounded Theory or Thematic Analysis, occupy regions of this space, but there are other possible designs of SSQSs that address important research questions and work with the available resources and constraints to deliver important HCI findings, and that do not have a particular name.

Any study demands sensitivity and adaptation to the situation. It also needs to be coherent and clear: you need to 'look after your GOST'. Clearly presented methods abound in textbooks and papers. They are there to be adopted and adapted to be fit for purpose.

Figure 8 presents a checklist of questions that should be considered in the design, conduct and reporting of SSQSs. As discussed above (section 8), there will be decisions that need to be made but that are not an essential part of reporting: that should focus on key decisions, key changes in plan and their likely impact on findings.

	Planning and conducting study	*Additional considerations for reporting*
Purpose (§3)	What is the purpose of the study? Why is it an important study to conduct? What gap in knowledge is it filling?	Did the purpose of the study change? If so, why and how? What are the novel and important findings? Why do they matter to the reader?
Resources and constraints (§4)	What resources do you have to work with? What constraints limit possibilities?	Were there any novel features of the way resources were used (e.g. new technology probes or innovative use of social media)? Did the availability of resources (e.g. time) limit what was possible in important ways?
Researcher attributes and role (§4.3 & §9.2)	How many people are in the research team, and what are their roles? What knowledge and expertise does each researcher bring to the study? What training will each receive? To what extent will the researcher participate in the situation being observed (for observational studies)? What is the intended relationship between researcher(s) and participants?	Are there attributes of the research team that will have influenced the study in important ways? What role(s) did the researcher(s) play in the study setting? How did the relationship that was established with each participant influence the data that was gathered (if it's possible to tell)?
Advocacy (§4.1)	Do you need advocate(s) within the study setting? How will you identify and work with them?	Who did you work with, and what was their influence (e.g. in terms of helping to refine research questions or recruit participants)?
Participant recruitment (§4.2)	What is the approach to sampling participants? How (practically) will participants be recruited? What are inclusion and exclusion criteria, and how might they evolve over the course of the study? What is the anticipated relationship between researcher(s) and participant?	How were participants recruited in practice? Were there compromises that needed to be made, and what is the likely impact of this on the quality, reliability or generalizability of findings? What roles did researcher(s) and participant(s) take in the study?
Location and intervention (§4.4)	Where will the study take place? What forms of intervention are planned (e.g. introduction of novel prototype designs)? How naturalistic is the study?	Did the location(s) in which the study took place, or any interventions, influence outcomes in any important ways?
Role of theory (§9.1)	To what extent, and how, will theory play a role in data gathering, analysis and/or reporting?	How, if at all, did established theory shape the study? How do the findings relate to established theory?
Ethical considerations (§5)	Are there important ethical considerations that need to be addressed? How will you ensure that participants benefit as far as possible from participation? What will participants be told about the study? How will data be stored and anonymised?	Did ethical considerations shape the study in important ways? If so, how? How were participants informed about the study and what would be done with the data?
Techniques for data gathering (§6)	How will data be gathered (interviews, observation, etc.)? How will it be recorded? If multiple methods are to be used, how will they be sequenced and coordinated? What data will be gathered? How structured will the data gathering be? Will it be informed by theory? Define a protocol for observation, or a semi-structured interview script, or participant instructions (for think-aloud). How will data gathering be timed (e.g. to sample particular kinds of activity)?	How was data gathered in practice? How did data gathering change over the course of the study (if at all)? How were participants instructed (e.g. for a think-aloud study)?
Interleaving recruitment, data gathering and analysis	How interdependent will participant recruitment, data gathering and analysis be?	How was data gathering and analysis interleaved (if at all)? How did early analysis shape later data gathering?
Analysis of data (§7)	How will data be analysed? At what level of detail will transcription take place? What tools will be used to support analysis? Are codes pre-determined or identified through analysis? Are they agreed by a team? If there are multiple coders, is their coding independent or negotiated? Will participants be involved in analysis and / or validation? If the analysis is individual and reflexive, what steps will the researcher take to ensure the validity of	How was data analysed in practice? How iterative and reflexive was the analysis process? How was data validated in practice?

	findings?	
Reporting (§8)	Who is the audience? How will findings be reported?	What is novel? What is important? What is the evidence to support the claims being made?

Figure 8: A checklist for planning and reporting SSQSs

This checklist has doubtless overlooked some important decisions in planning and in reporting. Please do add any omissions as comments on this chapter.

12 Conclusion

Wolcott (2009, p.36) quotes a biologist, Paul Weiss, as claiming, "Nobody who followed the scientific method ever discovered anything interesting." Whether or not that is strictly true, one of the delights of exploratory qualitative studies is that they frequently deliver interesting, even surprising, findings.

In healthcare, there is a persistent view that randomised controlled trials are the 'gold standard' that defines criteria for quality in research (e.g. Concato et al., 2000). Although HCI has been less blighted by such a hierarchical view of research designs, there has nevertheless been a tendency to dismiss some forms of qualitative research as lacking rigour. While this is true of some studies, at other times it seems to be due to limited understanding of the culture, principles and processes of qualitative research.

All research demands trust: that the researcher did what they claim to have done, with integrity, and that the presentation is as accurate as possible. Because SSQSs are suitable for addressing a range of research questions, and because every study setting is different, there is not a 'one size fits all' method: methods need to be adapted to work with the resources and constraints of the project. Named methods should not be used as 'bumper stickers'; it is important to describe what was actually done at an appropriate level of detail to enable others to judge the quality of a study, and the implications for future research and practice.

My aim in this chapter has been to lay out a space of possibilities and considerations for Semi-Structured Qualitative Studies in HCI, and to provide pointers to literature where further details can be found. Not every qualitative research project is an ethnography or a GT. Not every project results in implications for design. There are many possible research questions, study designs and study outcomes. The challenge is to ensure that studies are of high quality, and outcomes of interest and value.

Acknowledgments

References

Adams, A., Blandford, A. & Lunt, P. (2005) Social empowerment and exclusion: A case study on digital libraries. In ACM Transactions on CHI. 12.2. 174-200.

Adams, A., Lunt, P., & Cairns, P. (2008). A qualitative approach to HCI research. In: Cairns, Paul and Cox, Anna eds. Research Methods for Human-Computer Interaction. Cambridge, UK: Cambridge University Press, pp. 138–157.

Anderson, R. J. (1994). Representations and requirements: The value of ethnography in system design. *Human-computer interaction*, *9*(3), 151-182.

Arthur, S., & Nazroo, J. (2003). Designing fieldwork strategies and materials. *Qualitative research practice: a guide for social science students and researchers*, *1*, 109-137

Atkinson, R., & Flint, J. (2001). Accessing hidden and hard-to-reach populations: Snowball research strategies. *Social research update*, *33*(1), 1-4.

Attfield, S. & Blandford, A. (2011) Making sense of digital footprints in team-based legal investigations: the acquisition of focus. Human–Computer Interaction Journal. 26(1&2). 38-71.

Attfield S., Fegan, S. & Blandford, A. (2008) Idea Generation and Material Consolidation: Tool Use and Intermediate Artefacts in a Case Study of Journalistic Writing. Cognition, Technology & Work.

Barbour, R. S. (2001). Checklists for improving rigour in qualitative research: a case of the tail wagging the dog?. *BMJ: British Medical Journal*, *322*(7294), 1115-1117.

Beyer, H., & Holtzblatt, K. (1998) *Contextual design: defining customer-centered systems*. Elsevier.

Blandford, A., Adams, A., Attfield, S., Buchanan, G., Gow, J., Makri, S., Rimmer, J. & Warwick, C. (2008a) The PRET A Rapporter framework: Evaluating digital libraries from the perspective of information work. *Information Processing & Management*. 44. 4-21.

Blandford, A. & Rugg, G. (2002) A case study on integrating contextual information with usability evaluation. *International Journal of Human-Computer Studies*. 57.1, 75-99.

Blandford, A., Green, T. R. G., Furniss, D. & Makri, S. (2008b) Evaluating system utility and conceptual fit using CASSM. International Journal of Human–Computer Studies. 66. 393-409.

Blandford (in press) Eliciting people's conceptual models of activities and systems. To appear in International Journal of Conceptual Structures and Smart Applications (IJCSSA)

Blandford, A. E., Wong, B. L. W., Connell, I. & Green, T. R. G. (2002) Multiple Viewpoints On Computer Supported Team Work: A Case Study On Ambulance Dispatch. In X. Faulkner, J. Finlay & F. Détienne (Eds.) *People and Computers XVI: Proc. HCI'02*. 139-156. Springer.

Blandford, A. & Wong, W. (2004) Situation Awareness in Emergency Medical Dispatch. *International Journal of Human–Computer Studies*. 61(4). 421-452.

Boren, T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *Professional Communication, IEEE Transactions on*, *43*(3), 261-278

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101

Buckley, B., Murphy, A. W., Byrne, M., & Glynn, L. (2007). Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischaemic heart disease. *Heart*, *93*(9), 1116-1120.

Button, G., & Sharrock, W. (2009). Studies of Work and the Workplace in HCI: Concepts and Techniques. *Synthesis Lectures on Human-Centered Informatics*, *2*(1), 1-96.

Cairns, P., & Cox, A. L. (2008) (Eds.) *Research methods for human-computer interaction*. Cambridge University Press.

Carayon, P., Hundt, A. S., Karsh, B. T., Gurses, A. P., Alvarado, C. J., Smith, M., & Brennan, P. F. (2006). Work system design for patient safety: the SEIPS model. *Quality and Safety in Health Care*, *15*(suppl 1), i50-i58

Charmaz, K. (2006). Constructing grounded theory: A practical guide through qualitative analysis. Pine Forge Press.

Charmaz, K. (2008). Grounded Theory. In J. A. Smith, (Ed.). *Qualitative psychology: a practical guide to research methods*. Sage.

Charters, E. (2013). The Use of Think-aloud Methods in Qualitative Research An Introduction to Think-aloud Methods

Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine*, *342*(25), 1887-1892

Consolvo, S., & Walker, M. (2003). Using the experience sampling method to evaluate ubicomp applications. *Pervasive Computing, IEEE*, *2*(2), 24-31

Corbin, J., & Strauss, A. (2008). Basics of qualitative research: Techniques and procedures for developing grounded theory. Sage.

Crabtree, A., Rodden, T., Tolmie, P., & Button, G. (2009). Ethnography considered harmful. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 879-888). ACM.

Curzon, P., Blandford, A., Butterworth, R. & Bhogal, R. (2002) Interaction Design Issues for Car Navigation Systems. In Sharp, Chalk, LePeuple & Rosbottom (Eds.) Proc. HCI 2002 (Vol. 2). 38-41. BCS

Dourish, P. (2006). Implications for design. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 541-550). ACM.

Ellis, D., Cox, D. & Hall, K. (1993). A Comparison of the Information-seeking Patterns of Researchers in the Physical and Social Sciences. Journal of Documentation 49(4), pp. 356-369.

Ellis, D. & Haugan, M. (1997). Modelling the Information-seeking Patterns of Engineers and Research Scientists in an Industrial Environment. Journal of Documentation 53(4), pp. 384-403.

Ericsson, K.A. & Simon, H.A. (1984) Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.

Flanagan, J. C. (1954). The critical incident technique. *Psychological bulletin*, *51*(4), 327.

Flick, U. (2009). An introduction to qualitative research. Sage.

Fry, C. L., Ritter, A., Baldwin, S., Bowen, K. J., Gardiner, P., Holt, T., Jenkinson, R. & Johnston, J. (2005). Paying research participants: a study of current practices in Australia. Journal of medical ethics, 31(9), 542-547.

Furniss, D., & Blandford, A. (2006). Understanding Emergency Medical Dispatch in terms of Distributed Cognition: a case study. *Ergonomics*, *49*(12-13), 1174-1203.

Furniss, D., Blandford, A., & Curzon, P. (2011a). Confessions from a grounded theory PhD: experiences and lessons learnt. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 113-122). ACM.

Furniss, D., Blandford, A. & Mayer, A. (2011b) Unremarkable errors: Low-level disturbances in infusion pump use. Proc. BCS HCI Conference. 197-204. http://dl.acm.org/citation.cfm?id=2305353

Gaver, W., & Dunne, A. (1999). Projected realities: conceptual design for cultural effect. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 600-607). ACM.

Giorgi, A. & Giorgi, B. (2008) Phenomenology. In J. A. Smith, (Ed.). *Qualitative psychology: a practical guide to research methods*. Sage.

Glaser, B. G., & Strauss, A. L. (2009). The discovery of grounded theory: Strategies for qualitative research. Transaction Books

Grady, C., Dickert, N., Jawetz, T., Gensler, G., & Emanuel, E. (2005). An analysis of US practices of paying research participants. Contemporary clinical trials, 26(3), 365-375.

Grbich, C. (2013). *Qualitative data analysis: An introduction*. 2nd Edition. Sage.

Hartswood, M., Procter, R., Rouncefield, M., & Slack, R. (2003). Making a case in medical work: implications for the electronic medical record. *Computer Supported Cooperative Work*, *12*(3), 241-266.

Heath, C., & Luff, P. (1991). Collaborative activity and technological design: Task coordination in London Underground control rooms. In *Proceedings of the second conference on European Conference on Computer-Supported Cooperative Work* (pp. 65-80). Kluwer.

Henwood, K.L. and Pidgeon, N.F. (1992). Qualitative research and psychological theorising. British Journal of Psychology, 83, 97-111.

Hertzum, M., & Jacobsen, N. E. (2001). The evaluator effect: A chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction*, *13*(4), 421-443.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, *7*(2), 174-196.

Holtzblatt, Karen and Beyer, Hugh R. (2013): Contextual Design. In: Soegaard, Mads and Dam, Rikke Friis (eds.). "The Encyclopedia of Human-Computer Interaction, 2nd Ed.". Aarhus, Denmark: The Interaction Design Foundation. Available online at http://www.interaction-design.org/encyclopedia/contextual_design.html

Hughes, J., King, V., Rodden, T., & Andersen, H. (1994). Moving out from the control room: Ethnography in system design. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (pp. 429-439). ACM.

Hutchins, E. (1995). *Cognition in the Wild* (Vol. 262082314). Cambridge, MA: MIT press.

Johnson, J. and Henderson, A. (2011) Conceptual Models: Core to Good Design Morgan Claypool Synthesis Lectures on Human-Centered Informatics

Kamsin, A., Blandford, A., & Cox, A. L. (2012, May). Personal task management: my tools fall apart when I'm very busy! In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (pp. 1369-1374). ACM

Kaptelinin, Victor (2013): Activity Theory. In: Soegaard, Mads and Dam, Rikke Friis (eds.). "The Encyclopedia of Human-Computer Interaction, 2nd Ed.". Aarhus, Denmark: The Interaction Design Foundation. Available online at http://www.interaction-design.org/encyclopedia/activity_theory.html

Kidder, L. H., & Fine, M. (1987). Qualitative and quantitative methods: When stories converge. *New directions for program evaluation*, *1987*(35), 57-75.

Kindberg, T., Spasojevic, M., Fleck, R., & Sellen, A. (2005). The ubiquitous camera: An in-depth study of camera phone use. *Pervasive Computing, IEEE*, *4*(2), 42-50.

Klein, G. A., Calderwood, R. & Macgregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man and Cybernetics, 19(3), 462-472.

Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS quarterly*, 67-93.

Kock, Ned (2013): Action Research: Its Nature and Relationship to Human-Computer Interaction. In: Soegaard, Mads and Dam, Rikke Friis (eds.). "The Encyclopedia of Human-Computer Interaction, 2nd Ed.". Aarhus, Denmark: The Interaction Design Foundation. Available online at http://www.interaction-design.org/encyclopedia/action_research.html

Kvale, S. and Brinkmann, S. (2009) InterViews: Learning the Craft of Qualitative Research Interviewing. Sage.

Lazar, J., Feng, J. H., & Hochheiser, H. (2010). Research methods in human-computer interaction. Wiley.

Legard, R., Keegan, J., & Ward, K. (2003). In-depth interviews. In J. Richie & J. Lewis (Eds) Qualitative research practice: A guide for social science students and researchers, 138-169

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Sage.

Lindtner, S., Chen, J., Hayes, G. R., & Dourish, P. (2011). Towards a framework of publics: Re-encountering media sharing and its user. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(2), 5.

Lipson, J. G. (1997). The politics of publishing: protecting participants' confidentiality. In J. Morse (Ed.) *Completing a qualitative project: details and dialogue. Sage Publications*.

Lützhöft, M., Nyce, J. M., & Petersen, E. S. (2010). Epistemology in ethnography: assessing the quality of knowledge in human factors research. *Theoretical Issues in Ergonomics Science*, *11*(6), 532-545.

Mackay, W. (1999) Is paper safer? The role of paper flight strips in air traffic control. ACM Trans. CHI. 6. 311-340.

Mackay, W. E., & Fayard, A. L. (1997, August). HCI, natural science and design: a framework for triangulation across disciplines. In *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 223-234). ACM.

Makri, S., Blandford, A., Gow, J., Rimmer, J., Warwick, C. & Buchanan, G. (2007). A Library or Just another Information Resource? A Case Study of Users' Mental Models of Traditional and Digital Libraries. Journal of the American Society of Information Science and Technology. 58.3, 433-445. DOI http://dx.doi.org/10.1002/asi.20510

Makri, S., Blandford, A. & Cox, A. L. (2008). Using Information Behaviours to evaluate the functionality and usability of electronic resources: From Ellis's model to evaluation. *JASIST*. 59. 2244-2267.

Makri, S. & Blandford, A. (2012) Coming Across Information Serendipitously: Part 1 – A Process Model. Journal of Documentation. Vol. 68, No. 5, 684-705. DOI http://dx.doi.org/10.1108/00220411211256030

Marshall, M. N. (1996). Sampling for qualitative research. *Family practice*, *13*(6), 522-526.

Mays, N., & Pope, C. (2000). Qualitative research in health care: Assessing quality in qualitative research. *BMJ: British Medical Journal*, *320*(7226), 50.

McKechnie, L. E., Serantes, L. C., & Hoffman, C. (2012). Dancing around the edges: the use of postmodern approaches in information behaviour research as evident in the published proceedings of the biennial ISIC conferences, 1996-2010. In Proc. ISIC 2012.

Mentis, H. M., Reddy, M., & Rosson, M. B. (2013). Concealment of Emotion in an Emergency Room: Expanding Design for Emotion Awareness. *Computer Supported Cooperative Work (CSCW)*, *22*(1), 33-63.

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook. Sage.

Millen, D. R. (2000, August). Rapid ethnography: time deepening strategies for HCI field research. In Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques (pp. 280-286). ACM.

Morse, J. M. (1997). *Completing a qualitative project*. Sage publications.

Nørgaard, M., & Hornbæk, K. (2006, June). What do usability evaluators do in practice?: an explorative study of think-aloud testing. In *Proceedings of the 6th conference on Designing Interactive systems* (pp. 209-218). ACM

O'Connor, L. (2011) Workarounds in accident and emergency & intensive therapy departments: resilience, creation and consequences. MSc dissertation. Available from http://www.ucl.ac.uk/uclic/studying/taught-courses/distinction-projects/2009_theses/OConnorL.pdf

Odom, W., Harper, R., Sellen, A., Kirk, D., & Banks, R. (2010, April). Passing on & putting to rest: understanding bereavement in the context of interactive technologies. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1831-1840). ACM

O'Kane, A. A., & Mentis, H. (2012). Sharing medical data vs. health knowledge in chronic illness care. In *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts* (pp. 2417-2422). ACM.

Palen, L. (1999). Social, individual and technological issues for groupware calendar systems. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 17-24). ACM.

Pennathur, P. R., Thompson, D., Abernathy III, J. H., Martinez, E. A., Pronovost, P. J., Kim, G. R., Lubomski, L. H., Marsteller, J. A. & Gurses, A. P. (2013). Technologies in the wild (TiW): human factors implications for patient safety in the cardiovascular operating room. *Ergonomics*, *56*(2), 205-219.

Perera, M. (2006) Human error in context: a study of post completion error in chip and pin transactions. MSc dissertation. UCL.

Rajkomar, A. & Blandford, A. (2012) Understanding Infusion Administration in the ICU through Distributed Cognition. Journal of Biomedical Informatics. Vol. 45, No 3. 580-590. DOI http://dx.doi.org/10.1016/j.jbi.2012.02.003

Rajkomar, A., Blandford, A. and Mayer, A. (2013) Coping with complexity in home hemodialysis: a fresh perspective on time as a medium of Distributed Cognition. Cognition, Technology and Work. (online first) DOI 10.1007/s10111-013-0263-x

Randall, Dave and Rouncefield, Mark (2013): Ethnography. In: Soegaard, Mads and Dam, Rikke Friis (eds.). "The Encyclopedia of Human-Computer Interaction, 2nd Ed.". Aarhus, Denmark: The Interaction Design Foundation. Available online at http://www.interaction-design.org/encyclopedia/ethnography.html

Randell, R., Ruddle, R. A., Mello-Thoms, C., Thomas, R. G., Quirke, P., & Treanor, D. (2013). Virtual reality microscope versus conventional microscope regarding time to diagnosis: an experimental study. *Histopathology*, *62*(2), 351-358.

Reddy, M. & Dourish, P. (2002) A finger on the Pulse: Temporal Rhythms and information seeking in medical work. In Proceedings of ACM CSCW'02. ACM Press. 344-353.

Rochlin, G. (1999) Safe operation as a social construct. Ergonomics. 42.11. 1549-1560.

Rode, J. A., Toye, E. F., & Blackwell, A. F. (2004). The fuzzy felt ethnography—understanding the programming patterns of domestic appliances. *Personal and Ubiquitous Computing*, *8*(3-4), 161-176.

Rode, J. A. (2011). Reflexivity in digital anthropology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 123-132). ACM

Roethlisberger FJ, Dickson WJ: *Management and the Worker*. Cambridge, Mass., Harvard University Press; 1939.

Rogers, Y. (2012). *HCI Theory: Classical, Modern, and Contemporary.* Morgan & Claypool.

Rogers, Y., Sharp, H. & Preece, J. (2011) Interaction Design (3rd edition). Wiley.

Sanderson, P. M., & Fisher, C. (1994). Introduction to this special issue on exploratory sequential data analysis. *Human–Computer Interaction*, *9*(3-4), 247-250.

Skeels, M. M., & Grudin, J. (2009, May). When social networks cross boundaries: a case study of workplace use of facebook and linkedin. In *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 95-104). ACM.

Smith, J. A. (Ed.). (2008). *Qualitative psychology: a practical guide to research methods.* Sage.

Smith, P., Blandford, A. & Back, J. (2009) Questioning, exploring, narrating and playing in the control room to maintain system safety. Cognition Technology and Work. 279-291. DOI 10.1007/s10111-008-0116-1

Stawarz, K. (2012) An Ergonomic Evaluation of the Potential Impact of Touch-Screen Tablets on Office Workers. MSc dissertation. Available from http://www.ucl.ac.uk/uclic/studying/taught-courses/distinction-projects/2011_theses/Stawarz_2011

Suchman, L. (1987) Plans and situated actions. The problem of human computer communication. Cambridge: Cambridge University Press.

Thimbleby, H. (2008). Write now. In P. Cairns & A. Cox (Eds.) *Research Methods in Human-Computer Interaction*, 196-211.

Vom Lehn, D., & Heath, C. (2005). Accounting for new technology in museum exhibitions. *International Journal of Arts Management*, 11-21.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge university press

Willig, C. (2008). *Introducing qualitative research in psychology*. 2nd Edition. Open University Press

Wixon, D. (2003), Evaluating Usability Methods: why the current literature fails the practitioner. Interactions, 29-34

Wolcott, H. F. (2009). *Writing Up Qualitative Research.* 3rd Edition. Sage

Wong, W. & Blandford, A. (2002) Analysing Ambulance Dispatcher Decision Making: Trialling Emergent Themes Analysis. In *Proc. HF2002* (electronic publication).

Woolrych, A., Hornbæk, K., Frøkjær, E., & Cockton, G. (2011). Ingredients and meals rather than recipes: A proposal for research that does not treat usability evaluation methods as indivisible wholes. *International Journal of Human-Computer Interaction*, *27*(10), 940-970.

Yardley, L. (2000). Dilemmas in qualitative health research. *Psychology and health*, *15*(2), 215-228.