Open Research Online

The Open University's repository of research publications
and other research outputs

Questionnaires, in-depth interviews and focus groups

Book Section

oro.open.ac.uk

2   Questionnaires, in-depth interviews and focus groups

ANNE ADAMS AND ANNA L. COX

2.1   Overview

With fast changing technologies and related human interaction issues, there is an increased need for timely evaluation of systems with distributed users in varying contexts (Pace, 2004). This has led to the increased use of questionnaires, in-depth interviews and focus groups in commercial usability and academic research contexts. Questionnaires are usually paper based or delivered online and consist of a set of questions which all participants are asked to complete. Once the questionnaire has been created, it can be delivered to a large number of participants with little effort. However, a large number of participants also means a large amount of data needing to be coded and analysed. Interviews, on the other hand, are usually conducted on a one-to-one basis. They require a large amount of the investigator's time during the interviews and also for transcribing and coding the data. Focus groups usually consist of one investigator and a number of participants in any one session. Although the views of any one participant cannot be probed to same degree as in an interview, the discussions that are facilitated within the groups often result in useful data in a shorter space of time than that required by one-to-one interviews.

All too often, however, researchers eager to identify usability problems quickly throw together a questionnaire, interview or focus group that, when analysed, produces very little of interest. What is often lacking is an understanding of how the research method design fits with the research questions (Creswell, 2003) and how to appropriately utilise these different approaches for specific HCI needs. The methods described in this chapter can be useful when used alone but are most useful when used together with other methods. Creswell (2003) provides a comprehensive analysis of the different quantitative and qualitative methods and how they can be mixed and matched for overall better quality research. Depending on what we are investigating, sometimes it is useful to start with a questionnaire and then, for example, follow up some specific points with an experiment, or a series of interviews, in order to fully explore some aspect of the phenomenon under study.

This chapter describes how to choose between and design questionnaires, interviews and focus group studies and using two examples illustrates the advantages of combining a number of approaches when conducting HCI research.

2.2   The methods

As with all research, before you start developing your materials, for example by composing questions for a questionnaire, it is important to consider:

- why you are asking the questions
- who the results are for
- what you expect to find from the answers (i.e. hypothesis – see Chapter 9)
- how you are going to analyse the data when you get them (see Chapter 7).

If you stop and reflect on these questions, it will be easier to compose more appropriate, accurate questions that provide useful findings. This reflection should also help you to understand which method will be the most appropriate for your needs. The methods discussed in this chapter enable us to investigate participants' subjective memories of an event or views on a topic whereas experiments and user testing give a more objective measure of what occurs (see Chapter 9). The choice of method therefore depends on the questions that you want answered. If you want to know what people do in a particular situation (e.g. which pieces of software they use and how frequently), then a questionnaire will probably be sufficient. If you want to identify why something has occurred, a questionnaire will provide less valid responses than in-depth interviews or focus groups because in-depth interviews and focus groups allow the respondent the freedom to express things in context that you may not have thought of before. Knowing what questions to ask is the hardest part of constructing a questionnaire. Triangulating between qualitative and quantitative approaches can help with this problem.

This section of the chapter will outline how to design questionnaires, interview schedules and focus groups before considering in more detail the issues of sampling, piloting your study, triangulation of data and trust, privacy and ethics.

2.2.1   Using questionnaires

Questionnaire design

It is important to understand that a questionnaire is a tool and as such it must be usable so that the reader can easily understand, interpret and complete it. This in turn will increase the accuracy of responses. As a research tool, we must consider two important concepts in questionnaire design: 'reliability' and 'validity'. Reliability refers to the consistency of a measure whilst validity refers to its ability to measure what it is supposed to be measuring. There are several ways in which these factors can be statistically analysed (e.g. a 'reliability coefficient' between questions) and several books that deal with these issues in depth (e.g. Aiken, 1996). These concepts and analysis methods are very valuable if you wish your tool to accurately assess attitudes or opinions. In the past researchers have concentrated on reliability as it is easier to measure. However, reliability without validity is useless. You may be reliably testing something completely different

from what you want to test. For example, you want to test interface layout but the respondents are actually giving responses for aesthetics. This is an issue that relates strongly to the questionnaire wording and scales (see below).

Questionnaire length and structure

It is commonly accepted that a questionnaire should not be over long. People's short attention spans mean that long questionnaires are completed less accurately as people rush to finish them. This is also true for obvious question repetition with respondents biased towards simply repeating what they said before whether it is accurate or not. One major problem with very long questionnaires is the likelihood of participants skim reading them, which increases the likelihood of participants misinterpreting complex questions. This is also a problem with background information or instructions given at the beginning of the questionnaire. Many, although not all, of these problems can be counteracted with careful design of the questionnaire (see below). It is important to understand that there is a not a simple golden number of questions that can be given as a limit for questionnaires. The amount of motivation felt by participants to complete the questionnaire can affect how much they are prepared to concentrate on completing it. For example, participants who feel the outcomes of the research will directly benefit them (e.g. the application will be redesigned and used by the participants themselves) may feel more motivated to complete a questionnaire. Participants who feel the research is irrelevant to them (e.g. they rarely use these applications) may feel less motivated to complete it. Since a spread of different types of users is often required, it is important to understand these variations when designing and piloting the questionnaire.

To increase the usability and effectiveness of the questionnaire tool it is important to consider how you structure it. This means reviewing the sequence of questions very carefully. Grouping questions together under a common themed heading will help the respondent contextualize the subsequent questions. This approach will also help you identify how the sequence is likely to affect the respondent. The order in which questions are presented may bias your respondent to give more or less favorable responses. This can also happen with the response scales whereby a respondent gives the same answer throughout a section without reading the questions. To counteract this you can either counterbalance the questions (i.e. a negative and positive statement) or you can counterbalance the response (i.e. going from positive to negative then negative to positive). However, you must be careful how you design this in, as it might cause errors with people not realising that the ordering has changed. Piloting the questionnaire should help you identify these problems and correct them in the final version.

Question wording

When designing questions it is important to consider if each question will have the same meaning for everyone; e.g. networks are a common technology term but the expression 'networking' can be commonly used to mean establishing social

relationships. It is, therefore important to make sure your frame of reference is clear. Providing definitions or examples is a useful way to overcome these problems. Some researchers provide scenarios at the beginning of sections to contextualize the questions. Be careful, however, that these don't bias responses.

Keeping questions as short and simple as possible will increase the likelihood that they will be understood as well as the accuracy of responses. Questions that are complicated by double negatives or loaded words are particularly hard for a respondent to answer (e.g. 'which of these applications did you not consider unreliable?'). It is also important not to ask more than one question at once as you will find the respondent only answering one (e.g. 'How do you rate the system response times to urgent and non-urgent request?'). Similarly it is essential not to ask a question that requires them to interpret your meaning of a term (e.g. 'Do you attend online tutorials on a regular basis?' What is meant by the concept 'regular basis', is it once a day or once a week?). Providing a range within which to choose a response will help to clarify these choices. Ranges can also help respondents feel happier about answering some questions. For example, being asked to write down your age can feel far more invasive than choosing from a selection of age ranges.

Ultimately, it is always important to consider what biases you may be relaying through the wording of a question. Leading questions are frequently the major problem with most questionnaires (e.g. 'Why do you think online assessment is wrong?'). Similarly questions that invite a socially desirable response will produced a biased set of responses (e.g. 'In online conferences which description best describes yourself: an "enthusiast", "occasional participant" or "lurker"?' Who would realistically describe themselves as a lurker?). Finally, assuming a respondent will be able to remember accurate details of events several months ago is unrealistic and will produce inaccurate responses.

Different types of questions and scales

It is important to understand the different types of questions that can be included within a questionnaire as they require different levels of interpretation by the reader. This in turn impacts on the level of concentration required by the participant and thus the ease in completing the questionnaire. There are four main types of questions:

- simple factual questions – requiring a yes/no responses, e.g. 'Do you have a computer at home?'
- complex factual questions – requiring some interpretation or analysis, e.g. 'How many times have you used this application today?'
- opinion and attitudinal questions – requiring more alternatives and deeper concentration
- open ended questions – requiring participants' full concentration.

Each of these types of questions suits different scales of responses, e.g. factual questions can be recorded as yes/no answers or with a space to enter an appropriate

The advice on the help pages was:
Very Good   Good   Acceptable   Poor   Very poor

I found the help page advice useful:
Strongly agree   1 2 3 4 5   Strongly disagree

Fig 2.1 *Two types of Likert scale*

quantity. Open-ended questions will require a box large enough to record the amount of detail you want and that participants are likely to give. There is little dispute about these scales and forms of data entry. However, opinion and attitude questions have a wide variety of rating scales and are rife with complex arguments. A full appraisal of these scales would be too lengthy for this chapter (see instead Aiken, 1996). The following paragraph, will therefore concentrate on the most popular scales and relevant issues.

The most commonly used attitude/opinion scale is the Likert Scale. There are different versions of this scale used in questionnaires (see Figure 2.1). Ideally a full Likert scaling should be undertaken by developing groups of items that are statistically tested to identify the degree to which they measure the same things (Aiken, 1996). However, many designers have neither the time nor the skills to take this approach and so utilise the basic look of the Likert scale (Aldridge and Levine, 2001) in a checklist type format. This means that the importance of each question/item may vary, so adding up across those items would be useless. What is often done is that percentages across users are given, e.g. '85 per cent of users didn't find the help page advice useful'. It is worth noting that this approach will lower both the reliability and validity of the tool. A useful way to counteract this is to triangulate the findings with other approaches.

One continuing debate amongst questionnaire researchers is the question of a middle category. Some researchers feel that it gives the respondent a true representation of the scale. Others argue that respondents opt for the middle category as an opt-out from making a decision either way. Ultimately, a well-designed questionnaire will counteract all of these problems.

2.2.2 Conducting interviews

Conducting interviews is a time-consuming process that requires careful planning. In this section we outline some of these planning issues, notably: interview structure; style; the setting; and recording the data.

One might think that conducting an interview is much like helping someone complete a questionnaire by reading out the questions and writing down their answers. However, as discussed in the overview of this chapter, researchers make use of interviews when they wish to obtain more detailed and thorough information on a topic than might be gleaned from a questionnaire. During an interview the investigator often follows a schedule of pre-prepared questions, but is also able

to deviate where necessary in order to maximise the information obtained. These are known as structured or semi-structured interviews. It is important to consider that the more structured an interview is the less likely it is for a participant to feel at ease and reveal important and relevant issues. However, the less structured it is the harder it is to analyse afterwards. If a structured approach is taken, it is still important to be flexible enough to allow jumping between questions that the participants have started to answer in a previous response. Sticking rigidly to a structure can simply annoy your respondents when you ask a question they've already answered. For semi-structured and unstructured interviews it is important that the interview structure is kept flexible so that key issues not identified before the interview are allowed to emerge through the discussion. If there are key points that need to be discussed, these should flow from the discussion rather than being forced on the interviewee. This takes practice to perform while keeping an informal feel to the structure. A useful aid is to have a list of issues and to tick them off if the interviewee mentions them along the way. It is also good to link things already being discussed to new issues on the list, so the conversation feels natural and not forced. Don't forget that conducting an interview is a skill and the more you complete the better you will be at doing them.

The format for an in-depth interview or focus group can take on many forms (e.g. scenario-led, reflective accounts, task-led). Yet there are some basic guidelines that can be followed for all of these approaches. Initially, it is important to put participants at their ease with general introductions and permissions (consent) for interviewing and recording the interview/focus group, as for any study involving people (see Chapter 1). Details on confidentiality should also be given about the anonymity of the information and sensitivity concerning its later usage. At this point it is useful to start the talking with a brief background to the study. This helps contextualise the focus of the questions and their responses. However, you must be very careful not to give away too much detail that would bias participants' responses. After this initial introduction it is useful to consider the interview or focus group in four main stages.

1 **Background:** At this point obtain the subjects' background details, such as their experience with equipment and experience in general. Do not attempt to ask any detailed personal background questions here as this will be received much better at the end of an interview, when a trust relationship has been established.
2 **Letting off steam:** It is now useful to ask general questions which allow the participant to 'let off steam'. Often participants have in their mind a set of key points that they want to tell you about (e.g. their pet hates or loves). If you do not allow them to unburden themselves of these issues at the beginning, you will find all your questions resulting in these points repeatedly resurfacing.
3 **Addressing issues:** Any issues that have not naturally been dealt with through (2) should now be introduced. This stage could require further scenario prompting – although this should be used sparingly and with great caution so as not to bias the respondent towards specific issues or responses. If there

are potentially sensitive issues that you want to ask the participant, leave them till the end of this section. This will help to increase a sense of trust between the participant and researcher, making it easier for the participant to talk with more comfort about these issues.

4 **Tying up/debriefing:** There should always be a rounding off of the session with a summing up of all the issues, so that the respondent believes they have presented all the information they want to. There should also be a reaffirmation that the information will be dealt with in the strictest confidence. Finally a short debriefing session may be required to present in more detail what the research will be used for and what the aim of the research is.

An appropriate interviewing style can aid natural interactions without fostering any intimidation from those being interviewed. Although there are many different styles and approaches to choose from, the student–tutor approach is the easiest and quickest to pick up and implement. In this approach the interviewer takes on the role of a student who is asking questions from an expert. This approach should accentuate to the end user how much their opinions will be valued. It is also a useful approach to support novice end users in revealing problems they have with applications.

The setting for the interview can influence its success as the more natural the setting, for the respondents, the more likely they are to give naturalistic responses. If you are asking them about technologies, they would use at home it easier for them to respond in their home environment. Keeping eye contact during an interview is also imperative for keeping the natural, conversational nature of the interview. This becomes very hard if you are trying to write down most of what is said and makes audio or video recording of the session extremely valuable. It is, however, useful to consider the impact that a recording device will have on the interview structure and responses. A large, imposing recording device will be likely to intimidate the respondent and thus their responses to the questions. However, it is worth noting that after a while the respondent often relaxes and forgets the presence of the recording device. It is important to ask permission to use any recording device and highlight to the participant why you need the device (e.g. 'I can't write down what people say quickly enough'). Again, reiterate here how the information will be kept in the strictest confidence and anonymised before being disseminated in an aggregated form. Although some people refuse permission to be recorded, it is very rare and notes can always be used in these circumstances.

To ensure that the data is complete enough for accurate analysis there are some key points that need to be considered throughout the interview.

• Don't do all the talking, you should talk at most for 5–15 per cent of the time.
• Don't ask leading questions or questions that are too restrictive. If you do your answers are likely to be reduced to a series of simple yes/no responses.
• Don't let participants drift off the focus, which is more likely in a group than with an individual. But at the same time be flexible about their responses to

questions and don't restrict them to answering your questions in the order you've got them in.

- Do remember that there are no right or wrong answers, only the respondents' opinions.
- Don't ever give the respondents your opinions on the topic as this may bias them to give responses that will please you.
- Do remember to ask questions that probe but do not prompt a response. If a respondent gives a vague response ask them to qualify it (e.g. 'What do you mean by that?'), but don't give them a response to agree to (e.g. 'Do you mean this . . . ?').
- Do ask a respondent to qualify what they mean when they uses jargon (e.g. 'What does eTMA mean?').
- Do remember, in a focus group, not to let dominant personalities steal the limelight by giving you all the responses.

2.2.3 Focus groups

Focus groups are very similar to interviews and therefore many of the guidelines for conducting interviews also apply to conducting focus groups. A focus group should not exceed six or seven participants (eight at a maximum) and it should also be no smaller than three people. Too large and people are more likely to break off to talk in sub-groups and leave people out of the discussion. Too small and it is hard to keep the conversation going in enough depth for the participants not to feel intimidated by the situation. It is also easier to have a homogeneous group of people for a focus group as they find it easier to talk to one another. Watch out, though, they may be more eager to impress one another and be biased in their responses. To counterbalance this likelihood it is useful to have comparative focus groups to compare responses with (Lunt and Livingstone, 1996).

When choosing whether to complete an interview or focus group often the decision is made on purely logistic reasons (e.g. it is very difficult to get seven software designers from different companies to agree to attend a focus group). However, there are some issues that for ethical or privacy reasons are better dealt with in an in-depth interview. Conversely a focus group will allow for easier reflection on collaborative experiences (Lunt and Livingstone, 1996; Bruseberg and McDonagh-Philp, 2002).

The role of the investigator during the focus group is to facilitate the discussion of a number of topics and to ensure that the data is recorded. It is often useful to allow the members of the group a few moments to consider each question before you invite them to start the discussion. The investigator needs to ensure that each member of the group gets an opportunity to put forward their views and that no single person dominates the meeting. One might consider ensuring that each person takes a turn at explaining their view on the topic before moving on to the next person in the group. The investigator should also reflect back to the group what the group view or views are on a topic.

2.2.4   Important issues to consider

Sampling

In all research the issue of sampling is a crucial one that is often overlooked. In experiments this is often discussed in reference to participants (see Chapter 1). Within questionnaire research there is a wide variety of sampling methods used (e.g. random, systematic, stratified, multi-stage clustered, quotas) that are reviewed in detail by many other books (e.g. Czaja and Blair, 1996; Creswell, 2003; Aldridge and Levine, 2001). However, there are some key terms and concepts that need to be understood. The 'theoretical population' is the population you may want to generalise your research to (e.g. internet users across the world). The study population is the population you can acquire access to (e.g. computer users within the UK). The 'sampling frame' is the reference point that will allow you to select appropriate people for your study (e.g. a telephone book). But you must realise that whatever sampling frame you use will bias how representative your sample is of the wider population. For example, selecting people from a telephone book will exclude those who are not in the book (e.g. without a phone, ex-directory). With your own research limitations taken into consideration (e.g. time, cost, opportunity) it is useful to remember that you will never get an ideal sample.

Piloting

It is essential to identify potential problems before the expensive, time-consuming, full-scale research is undertaken. Initial, small-scale studies (pilots) help to identify how the questions are likely to be interpreted. It is important to seriously consider any research issues that occur at this point and use them to improve your questionnaire design or interview techniques. These pilots, although small scale, should attempt to obtain a spread of different types of user so that variations in bias, expectations and abilities can be understood and accounted for. Giving your questionnaire to a selection of colleagues who understand questionnaire design can help to correct some of the problems that are so easy to miss. Be careful, though, not to take criticism too personally – there is no questionnaire that could not be criticised to some extent.

Triangulation

Triangulation is the term given to the use of other qualitative methods, literature and experimentation to evaluate research findings. It enables researchers to overcome the weakness or intrinsic biases that come from qualitative methods where the data is derived from a small number of observations and, rather than viewing quantitative methods as rivals, takes the view that we can gain most by employing a mixture of methods and thereby avoiding the limitations of any single method.

In order to validate the results from a qualitative study we can obtain multiple data sources and multiple observers, employ current theory and make use of multiple methods. Obtaining data from multiple sources might involve

interviewing people from different groups (different companies, professions, etc.) and identifying issues that are common across groups. Obtaining data from multiple observers would require having more than one person conducting the interviews and ensuring that issues are identified by all observers/investigators thus avoiding any experimenter bias. Theoretical triangulation requires looking in the established literature for references that support our findings (and ones that don't!) in order to add weight to our own conclusions. And using methodological triangulation requires the use of multiple methods to investigate the same phenomenon – perhaps by using this book to identify which other methods would be appropriate.

Trust, privacy and ethics

A poorly designed series of questions (e.g. unclear, lengthy, confusing questions with lots of typographic errors) can decrease a respondent's trust in the researcher. Leading questions can also make people believe there is a hidden agenda behind the questions thus decreasing their trust in the research. Providing additional information where needed can, in contrast, increase a sense of trust. Careful emphasis of the purpose of the research can encourage a sense of purpose in the respondents. However, as with experiments, the respondents' consent (verbal or written) to record and use the data must be obtained. This is often completed at the beginning of an interview or focus groups and in a tick-box format on a questionnaire. Another important ethical requirement is the need to establish a clear understanding with the respondents that they can withdraw from the research at any point.

Trust is closely connected to a respondent's perception of privacy (Adams and Sasse, 2001). If a respondent feels that their answers are likely to risk their personal privacy they will either not complete the research or distort their responses to put themselves in a more favourable light. Participants also need assurances of confidentiality – details of how the information is going to be stored and who will access it. Of key concern is how identifiable the information is. Anonymous information will produce less concern as the information will not be personally identifiable. It can be helpful to emphasize, either verbally or within the questionnaire, that any data will be anonymised. For example, participants may need to know that this means that their data will be merged together with other responses, e.g. 'Twenty per cent of participants gave this response', or 'Eight our of ten agreed with this statement.' Giving participants a number (e.g. participant 246) also helps to increase perceptions of anonymity and thus the likelihood of accurate responses. It is also important to emphasise that information that identifies groups of people (e.g. specific departments or organisations) will also be anonymised in later dissemination of findings. These issues can be summarised as follows:

- Trust and ethics: informed consent, right to withdraw, enhancing trust through question and questionnaire design
- Privacy and ethics: confidentiality, anonymity, sensitive use of information.

2.3   Applying the methods

In this section two examples (Adams and Sasse, 1999; Adams, Blandford and Lunt, 2005) are given for the different methods to help clarify a series of points that were found through each study. The same studies are being used in the chapter on qualitative analysis (see Chapter 7) and so both chapters can be used to cross-reference for the whole research process.

2.3.1 Questionnaire and interview study

Most organisations try to protect their systems from unauthorised access, usually through passwords. Considerable resources are spent designing secure authentication mechanisms, but the number of security breaches and problems is still increasing (Hitchings, 1995). Unauthorised access to systems, and resulting theft of information or misuse of the system, is usually due to hackers 'cracking' user passwords, or obtaining them through social engineering. At the same time many organisations have major problems with users forgetting their passwords. The cost in re-instating passwords is an ever-growing problem within modern organisations. A review of previous research (Adams and Sasse, 1999) identified that there is a need for balance in password construction between secure and memorable passwords. Secure passwords are often considered less memorable and vice versa. The development of usable yet effective system security is a field growing in importance. This study sought to identify relationships between memorability and security to support users in designing memorable yet secure passwords.

Method

A total of 139 questionnaire responses were received, half of which were from one particular telecommunications company. The other half of the responses came from organisations throughout the world. Participants were recruited via email and web interest groups. This approach could be argued to have biased subject sampling to technologically experienced respondents. It should be noted, however, that respondents were varied in both computer experience (less than 1 year to 10 years) and password experience (1 to over 10 passwords, some using over 30 passwords).

A pilot questionnaire was used to obtain initial quantitative and qualitative data to re-design the questionnaire. Although the questionnaire was placed on the web, the pilot feedback ascertained that anonymity was of key importance. Yet ensuring that participants did not complete and submit several copies was also important. This problem led to the development of a small program that would automatically anonymise participants' personal details and return the data entered via email for analysis (uniquely identified according to IP address). There was a personal details section in the questionnaire, but this did not have to be completed, giving the participants the freedom to privacy that they required. This was not

Table 2.1 *Mean correlation coefficients between automaticity, memorability and frequency of password usage*

	Responses	Correlation	Significance
Password usage by Memory problems	137	−0.220	$p < 0.05$
Automaticity by Memory problems	136	−0.634	$p < 0.05$
Requiredsecurity changes by Memory problems	122	−0.2079	$p < 0.05$

an ideal solution to the problems noted above, but, it did provide an adequate solutions to most of the problems identified.

The questions ranged from factual to open-ended questions with the appropriate scales as noted above. There was one main open-ended question at the end of the questionnaire that asked participants to add any other issues they felt relevant. The responses identified a very large response rate for this question (some participants wrote over a page) and many of these highlighted the limitations of the questionnaire design, for example identifying questions not previously noted before.

Findings and re-design

Although this questionnaire took a broad approach to the subject area, it focused on password related user behaviours, in particular memorability problems. Several correlations were identified using a Spearman's Correlation Coefficient (see Chapter 6). Several of the significant relationships found (summarised in Table 2.1), although never identified before, were relatively obvious, for example a significant ($p < 0.05$) correlation between 'infrequently used passwords' and 'frequent memory problems' and between 'frequently used passwords' and 'infrequent memory problems'. There were also significant correlations ($p < 0.005$) between 'have to think first' (before password recall) and 'frequent memory problems' (with the same password) and at the opposite end of this relationship between 'automatic' (password recall) and 'infrequent memory problems'. As noted earlier in the chapter, these findings highlight what was happening with the users, their passwords and their behaviours, but not why. The results could suggest that password memorability is partially reliant on frequency of use, which produces automaticity. This would tie in with observations on encoding specificity and the implicit vs. explicit memory models (Graf and Mandler, 1984; Graf and Schacter, 1985). An experiment could have been devised to investigate that this hypothesis was correct (see Chapter 1).

However, there was a surprising significant correlation ($p < 0.05$) between the 'desire to decrease security' and 'frequent memory problems' with the same password. Another interesting finding was that 50 per cent of respondents wrote their passwords down in one form or another. Both of these findings again highlighted a need to know why these issues were occurring. Results from the

open-ended sections of the questionnaire also suggested that this narrow approach was not addressing key problems with password usage as an authentication device. It was then decided that the questionnaire required a follow-up study to incorporate a more in-depth exploratory analysis of these issues. In-depth interviews with 30 participants from two comparable organisations (the original telecommunications company and a building research organisation) were then used to gather more detailed data.

The in-depth interviews uncovered a whole raft of relevant usability issues that meant that a re-conception of the research was needed. Ultimately, the focus of the original questionnaire on memorability issues was not the major problem with passwords as an authentication mechanism. The follow-on interviews took a more flexible approach to data gathering with a focus on themes identified from the open-ended sections of the questionnaire. The interviews were able to identify a whole range of security issues centreing on users complying with or circumventing password mechanisms. These password behaviours were found to relate to perceptions of 'security levels', 'information sensitivity' and 'compatibility with work practices'. The qualitative analysis chapter (see Chapter 7) will highlight how different stages of the analysis revealed more detailed unpacking of these issues.

The interviews uncovered an array of interesting issues that, because of their sensitive nature, would never have been identified through a questionnaire. The respondents needed assurance that exact details of their passwords would not be required. They also needed continual declarations that the data would be confidential and only used in an aggregated form. Respondents' motivation to complete the study was increased greatly by the knowledge that the findings would be relayed back to the organisation's security department to improve systems. Through these assertions, participants felt at ease to reveal negative as well as positive password behaviours (e.g. putting passwords on post-it notes, sharing passwords, using other people's passwords to circumvent organisational procedures). The interviews were conducted at the participants' place of work with a semi-structured flow to them and a student–tutor style of interaction. All these approaches worked to put the participants at their ease and uncover sensitive yet very valuable information.

2.3.2 Interview and focus group study

The following series of studies were closely interlinked with an iterative development of questions and style closely interwoven with the analysis procedure. Computer technology and the 'knowledge society' have inspired speculation about their effects on society and its organisations (Gallivan, 2000). Social contexts and work practices can have a significant impact on a community's engagement with new technology systems (Duncker, 2002). Lave and Wenger (1991) argue that social practices shape how we learn and, in turn, who we become. However, those social practices grow from and interlink with organisational structures.

Technology as 'boundary objects' often traverses these structures and knowledge domains, supporting communication and collaboration (Star and Griesemer, 1989). However, the empowering or excluding nature of these technologies is rarely related back to specific HCI issues. In this set of studies (Adams *et al.*, 2005) we sought to review these issues in relation to interactions between social context, system design and implementation.

Method

A series of studies was conducted, over a four-year period, based around the introduction and use of digital libraries in a clinical and academic domain. Across four different settings (within these two domains) in-depth interview and focus group data were collected from 144 users: end-users, librarians, designers and management. The data were analysed with reference to 'communities of practice' to identify socially relevant issues, both specific to each domain and generic (see Chapter 7 for more details). The studies each had different influences but two themes focused the research throughout: work practices (both current and those changing) and the impact of digital resources on those practices. It was decided that, throughout all the studies, a pre-defined concept of a 'digital library' would not be employed. This approach was taken so as to allow the users to explore what they perceived as comprising a digital library.

As this research was qualitative in nature a detailed analysis of the context for each of the studies was essential. This analysis of the context of study provided a background to previous change initiatives, implementation strategies and usage practices within each organisational structure. This in turn provided a backdrop to user technology expectations which proved invaluable. After each study was completed an analysis of how it related to the other contexts (similarities and differences) was undertaken. This analysis helped to contextualise each study at a generic level (see Chapter 7).

The academic domain data were gathered from a London-based university that is split over several geographically distributed campuses. Computers with web-accessible digital libraries were placed within the library and in computer suites. Focus groups and in-depth interviews were used to gather data from 25 academics and librarians (with a 50/50 split) from 4 different campuses within the university (10 from humanities, 10 from computer science, 4 from business and 1 with a library managerial role). The first clinical setting studied was a London teaching hospital. In this hospital, computers had been placed on the wards, with web-accessible digital libraries. Focus groups and in-depth interviews were used to gather data from 73 hospital clinicians. Fifty per cent of the respondents were nurses while the other fifty per cent were junior doctors, consultants, surgeons, Allied Health Professionals (AHPs; e.g. occupational therapists), managers and IT department members. A further comparative clinical domain study was conducted in a provincial teaching hospital. In this hospital, although all computers allowed access to web-accessible digital libraries, they were not placed on the wards, but within specified offices and the library. Twenty in-depth interviews

were used to gather data from management, library, IT, consultant and nursing employees. Finally, an evaluation of an information intermediary's role within clinical communities of practice was undertaken. Twenty in-depth interviews were conducted across eight different clinical teams over a six-month period, as well as an observational study of one team and an information intermediary collaborating during a drop-in session.

Between the two domains a wide variety of electronic resources were used and a good spread of computer literacy and digital library usage was obtained. On average the level of literacy within the academic setting was higher than within the health service. However, within both domains usage patterns were similarly biased towards usage for research purposes. Finally, four broad issues guided the focus of interview and focus group questions within all the studies:

- perceptions of participants' role within the organisation and their information requirements
- perceptions of current information practices, social structures and organisational norms
- the impact of current practices, structures and norms on information resource awareness, acceptance and use
- technology perceptions (specifically of digital libraries) and how these affect other issues already identified.

Findings and re-design

The choice between in-depth interviews and focus groups was made primarily for logistical reasons and was adapted for the needs of each study. High-level clinicians, for example, were difficult to arrange interviews with, so focus groups with these participants would have been impossible to organise. In contrast, some focus groups were organised with the nursing staff and allied health professionals at the end of regular meetings. This meant that a homogeneous group of appropriate respondents was already gathered and focused on practice-based behaviours. The advantage of these sessions was that often one participant would bring up an account of the technology that would spark off a memory or a contradictory opinion from another participant. The richness of this data for socially relevant practices was evident.

To ensure interviews and focus groups were retained and not cancelled it became imperative to be extremely flexible. Some interviews were held in the lunch halls, others on the wards, while one interview with a surgeon was held in the surgery common room in between operations. A good recording device was essential for these varied and often noisy environments. Often the interviews had to stop because an urgent issue had come up or support for a tricky procedure was needed. It was essential to keep the momentum going so that time wasn't wasted when the interview resumed. It was useful, both during the interviews and afterwards, to make a set of notes summarising key points as a reminder for the participant of the previous discussion (e.g. 'You were telling me about when the

****** system was implemented ...'). It is important, though, not to let these kinds of notes take the interviewer's focus away from the interview (i.e. they should only be supplemental to the recording).

Although the interviews contained often sensitive discussions everyone agreed to the sessions being recorded. There was one participant, however, who asked for the tape to be turned off at one particular point in the interview. When the tape was turned off this participant went on to relay some of his opinions that he knew were rather contentious. The respondent, however, noted vehemently that he wanted these opinions noted and 'on-record' but that he simply did not want to have his voice recorded saying them. This highlights again the importance of being flexible in an interview and having a pen and pad available to make notes when required.

In contrast to the previous study, a grounded theory approach was taken from the outset of the research. As this approach to analysis was taken, a change in questioning evolved in conjunction with the ongoing analysis. It is interesting to note, then, that as the data collection proceeded there was a shift in the interviewing focus towards perceptions of organisational change and different approaches to the implementation of technologies.

Through the analysis and interviewing, it was identified that the four settings represented three different approaches to digital library implementation: making digital library resources available from existing computer systems in people's offices and the library (a traditional approach); making computer systems, and hence digital libraries, available in shared spaces (in this case, hospital wards); and employing information intermediaries to work with staff and library resources. These different approaches engendered different perceptions of the technology. Further analysis was undertaken to draw out these differences through a community of practice perspective (see Chapter 7). The traditional approach produced perceptions of technology as being irrelevant for current needs and community practices. Making technology available within shared physical space, but with poor design, support and implementation procedures, was widely perceived as a threat to current organisational structures. In contrast, technology implemented within the community, which could adapt to and change practices according to individual and group needs, supported by an information intermediary, was seen as empowering to both the community and the individual. Finally these findings were reviewed in light of the contextual details gathered during the data collection. See Chapter 7 for the details of how the anslysis developed as an integral part of the data collecting process.

2.4   Critique

Questionnaires can be an invaluable tool when usability data is needed from large numbers of disparate users. They can be both cost-effective and easier to analyse than other methods. However, it must be understood that they suit some

research questions better than others. As mentioned above, they are reasonably effective at obtaining data about what issues are of increased importance (e.g. 95 per cent of respondents didn't use the help pages). It must be remembered, however, that these rely on respondents' subjective memories as opposed to objective measures that can be made in experiments and some user trials (see Chapter 1). Questionnaires that deal with opinions or attitudes will require a lot more time and effort in design, construction and piloting, thus increasing their cost.

It is also worth mentioning that all questionnaire designers must consider potential biases from subject sampling, fatigue and misinterpreting questions. Although, many of these problems can be overcome through careful questionnaire design and standardisation, this takes a great deal of skill. Finally, it is important to understand that questionnaires are limited to the questions asked. If you miss out important issues in your questionnaire design they will not appear in the analysis. However, as identified in the questionnaire example above, open-ended questions can sometimes allow these issues to surface. Ultimately, though, the importance of situational or contextual issues will be missed if this is the only research approach taken.

In-depth interview and focus groups can provide a flexible and participatory method that contextualises users' perceptions and experiences. As noted in the example above, participants are far more likely to release sensitive data (e.g. related to privacy and trust) when they have formed some rapport with the researcher in question. It is interesting to note that both these approaches allow the participant to sequence their responses themselves. As seen in the qualitative analysis in Chapter 7, these sequencings can be enlightening in themselves. However, it must be remembered that each methodology has its own limitations. Qualitative approaches are frequently criticised for being subjective and subject to the researcher's own biases. These issues are often dealt with through a series of researcher quality and inter-rater validity checks (see Chapter 7). What cannot be disputed is the time-consuming nature of these approaches reducing effective timely input into the design process. Findings will also be reduced to lower numbers than can be produced by questionnaires, although focus groups can increase numbers with similar time outlays.

Additional literature

There has been little research into the HCI application of questionnaires, in-depth interviews and focus groups. Further detailed explanations of these methods can be found within the social science literature (e.g. Strauss and Corbin, 1990; Aiken, 1996; Lunt and Livingstone, 1996; Czaja and Blair, 1996; Aldridge and Levine, 2001; and Creswell, 2003). There has, however, been some usage of these methods within the domain. Pace (2004) details the use of in-depth interviews to identify a grounded theory analysis of online flow experiences. Folstad, Bark and Gulliksen (2006) describe the use of questionnaires to assess HCI practitioners' evaluation practices. This paper presents a simple yet ideal approach to questionnaires as a means of ascertaining a broad simple picture of what is happening within the

field. Focus groups are less used within HCI but are beginning to be more readily used in the design domain as a whole. Bruseberg and McDonagh-Philp (2002) describe how the use of focus groups can effectively support the industrial/product designer in design decisions. Ultimately there is a growing appreciation for these methods and for how, through appropriate application, they can strengthen the quality of HCI research.