CHAPTER 7
DATA GATHERING

Objectives

The main aims of the chapter are to:

- Discuss how to plan and run a successful data gathering program.
- Enable you to plan and run an interview.
- Enable you to design a simple questionnaire.
- Enable you to plan and carry out an observation.

7.1 Introduction

This chapter presents some techniques for data gathering which are commonly used in interaction design activities. In particular, data gathering is a central part of establishing requirements, and of evaluation. Within the requirements activity, the purpose of data gathering is to collect sufficient, accurate, and relevant data so that a set of stable requirements can be produced; within evaluation, data gathering is needed in order to capture users' reactions and performance with a system or prototype.

In this chapter we introduce three main techniques for gathering data: interviews, questionnaires, and observation. In the next chapter we discuss how to analyze and interpret the data collected. Interviews involve an interviewer asking one or more interviewees a set of questions, which may be highly structured or unstructured; interviews are usually synchronous and are often face-to-face, but they don't have to be. Questionnaires are a series of questions designed to be answered asynchronously, i.e. without the presence of the investigator; these may be on paper, or online. Observation may be direct or indirect. Direct observation involves spending time with individuals observing activity as it happens. Indirect observation involves making a record of the user's activity as it happens to be studied at a later date. All three techniques may be used to collect qualitative or quantitative data.

Although this is a small set of basic techniques, they are flexible and can be combined and extended in many ways. Indeed it is important not to focus on just one data gathering technique but to use them flexibly and in combination so as to avoid biases which are inherent in any one approach. The way in which each technique is used varies, depending on the interaction design activity being undertaken. More detailed descriptions of how they are used and additional techniques relevant only to specific activities of the lifecycle are given in later chapters (Chapter 10 for requirements, and Chapters 13–15 for evaluation).

7.2 Five Key Issues

Data gathering sessions need to be planned and carried out carefully. Specific issues relating to the three data gathering techniques are discussed in the following sections, but first we consider five key issues that require attention for any data gathering session to be successful: goal setting, identifying participants, the relationship between the data collector and the data provider, triangulation, and pilot studies.

7.2.1 Setting Goals

The main reason for gathering data at all is to glean information about something. For example, you might want to understand how technology fits into normal family life, or you might want to identify which of two icons representing 'send message' is easier to use, or you might want to find out whether the redesign you are planning for a hand-held meter reader is along the right lines. There are many different reasons for gathering data, and before beginning it is important to identify specific goals for the study. The goals that are set will influence the nature of the data gathering sessions, the data gathering techniques to be used, and also the analysis to be performed. Once the goals have been set, you can concentrate on what data to look for and what to do with it once it is gathered.

The goals may be expressed more or less formally, e.g. using some structured or even mathematical format, or using a simple description such as the ones in the previous paragraph, but whatever the format they should be clear and concise. In interaction design it is more usual to express goals for data gathering informally.

7.2.2 Identifying Participants

The goals you develop for your data gathering session will indicate the kind of people you want to gather data from. Those people who fit this profile are called the population. In some cases, the people you need to gather data from may be clearly identifiable – maybe because there is a small group of users and you have access to each one. However, it is more likely that you will need to choose the participants to include in your data gathering, and this is called sampling. The situation where you have access to all members of your target population is called saturation sampling, but this is quite rare. Assuming that you will be choosing to involve a proportion of your population in data gathering, then you have two options: probability sampling or non-probability sampling. In the former case, the most commonly used approaches are simple random sampling or stratified sampling; in the latter the most common are convenience sampling or volunteer panels.

Random sampling can be achieved by using a random number generator or by choosing every *n*th person in a list. Stratified sampling relies on being able to divide the population into groups (e.g. classes in a secondary school), and then applying random sampling. Both convenience sampling and volunteer panels rely less on you choosing the participants and more on participants being prepared to take part. The term convenience sampling is used to describe a situation where the sample includes those who were available rather than those specifically selected.

The crucial difference between probability and non-probability methods is that in the former you can apply statistical tests and generalize to the whole population, while in the latter such generalizations are not robust. Using statistics also requires having a sufficient number of participants. What exactly 'sufficient' means will depend on the type of data being collected and the kind of statistical tests that need to be applied. This can be a complex issue so if not confident with statistics, it is best to consult with a someone who knows about them. See Sue and Ritter (2012) for a more detailed treatment of sampling.

7.2.3 Relationship with Participants

One significant aspect of any data gathering is the relationship between the person (people) doing the gathering and the person (people) providing the data. Making sure that this relationship is clear and professional will help to clarify the nature of the study. One way in which this can be achieved is to ask participants to sign an informed consent form. The details of this form will vary, but it usually asks the participants to confirm that the purpose of the data gathering and how the data will be used have been explained to them and that they are happy to continue. It also often includes a statement that participants may withdraw at any time, and that in this case none of their data will be used in the study.

It is common practice in many countries to use an informed consent form when running evaluation sessions, particularly where the participants are members of the public, or are volunteers in a research project. The informed consent form is intended to protect the interests of both the data gatherer and the data provider (see Chapter 13). The gatherer wants to know that the data she collects can be used in her analysis, presented to interested parties, and published in reports (as appropriate). The data provider wants reassurance that the information he gives will not be used for other purposes, or in any context that would be detrimental to him. For example, he wants to be sure that personal contact information and other personal details are not made public. This is especially true when people with disabilities or children are being interviewed. In the case of children, using an informed consent form reassures parents that their children will not be asked threatening, inappropriate, or embarrassing questions, or be asked to look at disturbing or violent images. In these cases, parents are asked to sign the form. Figure 7.1 shows an example of a typical informed consent form.

Crowdsourcing Design for Citizen Science Organizations

SHORT VERSION OF CONSENT FORM for participants at the University of Maryland – 18 YEARS AND OLDER

You are invited to participate in a research project being conducted by the researchers listed on the bottom of the page. In order for us to be allowed to use any data you wish to provide, we must have your consent.

In simplest terms, we hope you will use the mobile phone, tabletop, and project website at the University of Maryland to

- Take pictures
- Share observations about the sights you see on campus
- Share ideas that you have to improve the design of the phone or tabletop application or website
- Comment on pictures, observations, and design ideas of others

The researchers and others using CampusNet will be able to look at your comments and pictures on the tabletop and/or website, and we may ask if you are willing to answer a few more questions (either on paper, by phone, or face-to-face) about your whole experience. You may stop participating at any time.

A long version of this consent form is available for your review and signature, or you may opt to sign this shorter one by *checking off all the boxes that reflect your wishes and signing and dating the form below.*

___I agree that any photos I take using the CampusNet application may be uploaded to the tabletop at the University of Maryland and/or a website now under development.

___I agree to allow any comments, observations, and profile information that I choose to share with others via the online application to be visible to others who use the application at the same time or after me.

___I agree to be videotaped/audiotaped during my participation in this study.

___I agree to complete a short questionnaire during or after my participation in this study.

NAME [Please print]	
SIGNATURE	
DATE	

[Contact information of Senior Researcher responsible for the project]

Figure 7.1 Example of an informed consent form

However, this kind of consent is not generally required when collecting data for the requirements activity where a contract usually exists in some form between the data collector and the data provider. For example, consider the situation where a consultant is hired to gather data from a company in order to establish a set of requirements for a new interactive system to support timesheet entry. The employees of this company would be the users of the system, and the consultant would therefore expect to have access to the employees to gather data about the timesheet activity. In addition, the company would expect its employees to cooperate in this exercise. In this case, there is already a contract in place which covers the data gathering activity, and therefore an informed consent form is less likely to be required. As with most ethical issues, the important thing is to consider the situation carefully and make a judgment based on the specific circumstances. Increasingly, projects that involve collecting data from humans are being reviewed to ensure that participants' personal information is protected.

Incentives for completing a questionnaire might be needed in some circumstances because there is no clear and direct advantage to the respondents, but in other circumstances, respondents may see it as part of their job to complete the questionnaire. For example, if the questionnaires form part of the requirements activity for a new mobile sales application to support sales executives, then it is likely that sales executives will complete a questionnaire about their job if they are told that the new device will impact their day-to-day lives. In this case, the motivation for providing the required information is clear. However, if you are collecting data to understand how appealing

a new interactive website is for school children, different incentives would be appropriate. Here, the advantage for the individuals to complete a questionnaire is not so obvious.

7.2.4 Triangulation

Triangulation is a term used to refer to the investigation of a phenomenon from (at least) two different perspectives (Denzin, 2006; Jupp, 2006). Four types of triangulation have been defined (Jupp, 2006):

1. Triangulation of data means that data is drawn from different sources at different times, in different places, or from different people (possibly by using a different sampling technique).
2. Investigator triangulation means that different researchers (observers, interviewers, etc.) have been used to collect and interpret the data.
3. Triangulation of theories means the use of different theoretical frameworks through which to view the data or findings.
4. Methodological triangulation means to employ different data gathering techniques.

The last of these is the most common form of triangulation. One application of triangulation (and again the most common) is to validate the results of some inquiry by pointing to similar results yielded through the use of different perspectives. However, validation through triangulation is difficult to achieve. Different data gathering methods result in different kinds of data, which may or may not be compatible. Using different theoretical frameworks may or may not result in complementary findings, but to achieve theoretical triangulation would require the theories to have similar philosophical underpinnings. Using more than one data gathering technique, and more than one data analysis approach, is good practice, but achieving true triangulation is rare.

7.2.5 Pilot Studies

Apilot study is a small trial run of the main study. The aim is to make sure that the proposed method is viable before embarking on the real study. Data gathering participants can be (and usually are) very unpredictable, even when a lot of time and effort has been spent carefully planning the data gathering session. Plans should be tested by doing a pilot study before launching into the main study. For example, the equipment and instructions that are to be used can be checked, the questions for an interview or in a questionnaire can be tested for clarity, and an experimental procedure can be confirmed as viable. Potential problems can be identified in advance so that they can be corrected. Distributing 500 questionnaires and then being told that two of the questions were very confusing wastes time, annoys participants, and is an expensive error that could have been avoided by doing a pilot study.

If it is difficult to find people to participate or if access to participants is limited, colleagues or peers can be asked to comment. Getting comments from peers is quick and inexpensive and can be a substitute for a pilot study. It is important to note that anyone involved in a pilot study cannot be involved in the main study. Why? Because they will know more about the study and this can distort the results.

BOX 7.1

Data, Information, and Conclusions

There is an important difference between raw data, information, and conclusions. Data is what you collect; this is then analyzed and interpreted and conclusions drawn. Information is gained from analyzing and interpreting the data and conclusions represent the actions to be taken based on the information. For example, you might want to know whether a particular screen layout has improved the user's understanding of the application. In this case, the raw data collected might include the time it takes for a set of users to perform a particular task, the users' comments regarding their use of the application, biometric readings about the users' heart rates while using the application, and so on. At this stage, the data is raw. Information will emerge once this raw data has been analyzed and the results interpreted. For example, you may find after analyzing the data that people with more than 5 years' experience find the new design frustrating and take longer to achieve their goals, while those with less than 2 years' experience find the design helpful and complete tasks more quickly. Your interpretation may be that the new layout has improved novices' understanding but has irritated more experienced users, and you may conclude that the layout needs to be redesigned. ■

7.3 Data Recording

Capturing data is necessary so that the results of a data gathering session may be taken away and analyzed. Some forms of data gathering such as questionnaires, diaries, interaction logging, and collecting work artifacts are self-documenting and no further recording is necessary, but for other techniques there is a choice of recording approaches. The most common of these are taking notes, audio recording, taking photographs, and video recording. These may be used individually or in combination. For example, an interview may be audio recorded and then to help the interviewer in later analysis, a photograph of the interviewee may be taken.

Which data recording approaches are used will depend on the context, time and resources available, and the sensitivity of the situation; the choice of data recording approach will affect the level of detail collected, and how intrusive the data gathering will be. In most settings, audio recording, photographs, and notes will be sufficient. In others it is essential to collect video data so as to record in detail the intricacies of the activity and its context. Three common data recording approaches are discussed below.

7.3.1 Notes Plus Photographs

Taking notes (by hand or by typing) is the least technical and most flexible way of recording data. Handwritten notes may be transcribed, in whole or in part. While this may seem tedious, it is usually the first step in the analysis, and this activity gives the analyst a good overview of the quality and contents of the data collected. Even though tools exist for supporting data collection and analysis, the advantages of handwritten notes include that pen and paper are much less intrusive than a keyboard, and they are extremely flexible. Disadvantages with notes include that it can be difficult and tiring to write (or type) and listen or observe at the same time, it is easy to lose concentration, biases creep in, handwriting can be difficult to decipher, and the speed of writing (or typing) is limited. However, working with another person solves some of these problems and provides another perspective.

If appropriate, photograph(s) and short videos, captured via smartphones or other handheld devices, of artifacts, events, and the environment can be used to supplement notes and hand-drawn sketches, provided permission has been given.

7.3.2 Audio Plus Photographs

Audio recording can be a useful alternative to note taking and is less intrusive than video. In observation, it allows observers to focus on the activity rather than trying to capture every spoken word. In an interview, it allows the interviewer to pay more attention to the interviewee rather than try to take notes as well as listen, but transcribing a lot of audio data is time-consuming. However, it isn't always necessary to transcribe all of it – often only sections are needed, depending on why the data was collected. Many studies do not need a great level of detail, and instead, recordings are used as a reminder and as a source of anecdotes for reports. It is also surprising how evocative it can be to hear audio recordings of people or places from when you collected the data. If you are using audio recording as the main or only data collection technique then it is important that the quality is good and it is advisable to check this before starting your data collection.

Audio recording can be supplemented with photographs, as mentioned above.

7.3.3 Video

Video has the advantage of capturing both visual and audio data but video recording has some additional planning issues that need to be addressed, and it can be intrusive (no matter how well you plan it) (Denzin and Lincoln, 2011). Heath *et al* (2010) identify several of these issues including:

- Deciding whether to fix the camera's position or use a roving recorder. This decision depends on the activity being recorded and the purpose to which the video data will be put – e.g. for illustrative purposes only or for detailed data analysis. In some cases, such as pervasive games, a roving camera is the only way to capture the required action.
- Deciding where to point the camera in order to capture what is required. Heath and his colleagues suggest carrying out fieldwork for a short time before starting to video record in order to become familiar with the environment and be able to identify suitable recording locations. Involving the participants themselves in deciding what and where to record also helps to capture relevant action.
- Understanding the impact of the recording on participants. It is often assumed that video recording will have an impact on participants and their behavior but Heath *et al* (2010) suggest taking an empirical approach to the question and examining the data itself to see whether there is any evidence of behavior orienting to the camera.

Activity 7.1

Imagine you are a consultant who is employed to help develop a new computerized garden planning tool to be used by amateur and professional garden designers. Your goal is to find out how garden designers use an early prototype as they walk around their clients' gardens sketching design ideas, taking notes, and asking the clients about what they like and how they and their families use the garden. What are the advantages and disadvantages of the three approaches to data recording discussed above, in this environment?

Comment

Handwritten notes do not require specialist equipment. They are unobtrusive and very flexible but difficult to do while walking around a garden. If it starts to rain, there is no equipment to get wet, but notes may get soggy and difficult to read (and write!). Video captures more information, e.g. the landscape, where the designers are looking, sketches, comments, etc., but it is more intrusive and you must hold the camera. Video may also be tricky to capture if it starts to rain. Short video sequences captured on a smartphone are easier to collect and tend to be less obtrusive. Audio may be a good compromise, but integrating sketches and other artifacts later can be more difficult.

Garden planning is a highly visual, aesthetic activity, so it would be important to supplement note taking and audio recording with photographs captured using either a digital camera or a smartphone. ■

7.4 Interviews

Interviews can be thought of as a "conversation with a purpose" (Kahn and Cannell, 1957). How like an ordinary conversation the interview can be depends on the type of interview method used. There are four main types of interviews: open-ended or unstructured, structured, semi-structured, and group interviews (Fontana and Frey, 2005). The first three types are named according to how much control the interviewer imposes on the conversation by following a predetermined set of questions. The fourth involves a small group guided by a facilitator.

The most appropriate approach to interviewing depends on the purpose of the interview, the questions to be addressed, and the stage in the lifecycle. For example, if the goal is to gain first impressions about how users react to a new design idea, such as an interactive sign, then an informal, open-ended interview is often the best approach. But if the goal is to get feedback about a particular design feature, such as the layout of a new web browser, then a structured interview or questionnaire is often better. This is because the goals and questions are more specific in the latter case.

7.4.1 Unstructured Interviews

Open-ended or unstructured interviews are at one end of a spectrum of how much control the interviewer has over the interview process. They are exploratory and are more like conversations around a particular topic; they often go into considerable depth. Questions posed by the interviewer are open, meaning that there is no particular expectation about the format or content of answers. For example, the first question asked of all participants might be: 'What are the advantages of using a touch screen?' Here, the interviewee is free to answer as fully or as briefly as she wishes and both interviewer and interviewee can steer the interview. For example, often the interviewer will say: "Can you tell me a bit more about . . ." This is referred to as probing.

Despite being unstructured and open, it is always advisable for the interviewer to have a plan of the main topics to be covered, so that she can make sure that all the topics of interest are included. Going into an interview without an agenda should not be confused with being open to hearing new ideas (see Section 7.4.5 on planning an interview). One of the skills necessary for conducting an unstructured interview is getting the balance right between making sure that answers to relevant questions are obtained, while at the same time being prepared to follow new lines of enquiry that were not anticipated.

Abenefit of unstructured interviews is that they generate rich data that is often interrelated and complex, i.e. data that gives a deep understanding of the topic. In addition, interviewees may mention issues that the interviewer has not considered. Alot of unstructured data is generated and the interviews will not be consistent across participants since each interview takes on its own format. Unstructured interviews can therefore be time-consuming to analyze, although themes can often be identified across interviews using techniques from grounded theory and other approaches discussed in Chapter 8. These characteristics need to be taken into account when deciding which type of interview to choose.

7.4.2 Structured Interviews

In structured interviews, the interviewer asks predetermined questions similar to those in a questionnaire (see Section 7.5), and the same questions are used with each participant so the study is standardized. The questions need to be short and clearly worded, and they are typically closed questions, which means that they require an answer from a predetermined set of alternatives (this may include an 'other' option, but ideally this would not be chosen very often). Closed questions work well if the range of possible answers is known, and when participants are in a rush. Structured interviews are only really useful when the goals are clearly understood and specific questions can be identified. Example questions for a structured interview might be:

- Which of the following websites do you visit most frequently: amazon.com, google.com, msn.com?
- How often do you visit this website: every day, once a week, once a month, less often than once a month?
- Do you ever purchase anything online: yes/no? If your answer is yes, how often do you purchase things online: every day, once a week, once a month, less frequently than once a month?

Questions in a structured interview should be worded exactly the same for each participant, and they should be asked in the same order.

7.4.3 Semi-structured Interviews

Semi-structured interviews combine features of structured and unstructured interviews and use both closed and open questions. The interviewer has a basic script for guidance, so that the same topics are covered with each interviewee. The interviewer starts with preplanned questions and then probes the interviewee to say more until no new relevant information is forthcoming. For example:

> ***Which music websites do you visit most frequently? <Answer: mentions several but stresses that she prefers hottestmusic.com>***
>
> ***Why? <Answer: says that she likes the site layout>***
>
> ***Tell me more about the site layout <Silence, followed by an answer describing the site's layout>***
>
> ***Anything else that you like about the site? <Answer: describes the animations>***
>
> ***Thanks. Are there any other reasons for visiting this site so often that you haven't mentioned?***

It is important not to pre-empt an answer by phrasing a question to suggest that a particular answer is expected. For example, 'You seemed to like this use of color . . .' assumes that this is the case and will probably encourage the interviewee to answer that this is true so as not to offend the interviewer. Children are particularly prone to behave in this way (see Box 7.2 for more on data gathering

with children). The body language of the interviewer, for example whether she is smiling, scowling, looking disapproving, etc., can have a strong influence on whether the interviewee will agree with a question, and the interviewee needs to have time to speak and not be moved on too quickly.

Probes are a useful device for getting more information, especially neutral probes such as 'Do you want to tell me anything else?', and prompts which remind interviewees if they forget terms or names help to move the interview along. Semi-structured interviews are intended to be broadly replicable, so probing and prompting should aim to help the interview along without introducing bias.

BOX 7.2
Working with Children

Children think and react to situations differently from adults. Therefore, if children are to be included in data gathering sessions, then child-friendly methods are needed to make them feel at ease so that they will communicate with you. For example, for very young children of pre-reading or early reading age, data gathering sessions need to rely on images and chat rather than written instructions or questionnaires. Many researchers who work with children have developed sets of 'smileys', such as those shown in Figure 7.2, so that children can select the one that most closely represents their feelings (e.g. Read *et al*, 2002).

Figure 7.2 A smileyometer gauge for early readers

Source: **Figure 2, Janet Read, Stuart MacFarlane and Chris Casey "Endurability, Engagement and Expectations: Measuring Children's Fun" Department of Computing, University of Central Lancashire. Reproduced with permission.**

Several other techniques for data gathering with children have been developed. For example, in KidReporter (Bekker *et al*, 2003) children are asked to produce newspaper articles on the topic being investigated, while the Mission from Mars approach (Dindler *et al*, 2005) involves children explaining everyday experiences over an audio connection to a researcher pretending to be a Martian.

Druin (2002) identifies four roles for children in the design of technology (particularly for learning): user, tester, informant, and design partner. In the role of user children use the technology while adults observe, as tester children test prototypes of technology, as informant children take part in the design process at various stages, and as design partner children are equal stakeholders throughout the design process.

Guha *et al* (2013) work with children as technology design partners. They have found that unexpected innovations result when working as an intergenerational team, i.e. adults and children working together. The method they use is called cooperative inquiry (Druin, 2002; Guha *et al*, 2013), based on Scandinavian cooperative design practices, participatory design, and contextual inquiry. Many techniques can be used in cooperative inquiry, such as sketching ideas and brainstorming, and observational research.

Researchers also use a variety of participatory design methods in design-based research (DBR), a methodology that is common in the fields of learning sciences and interaction design for children. In DBR, researchers design theory-driven learning environments, test these designs in authentic educational contexts, and then use the resulting research findings to inform further iterative cycles of design and testing. Yip *et al* (2013) employ these methodologies in designing educational environments and technologies for children's science learning. They find that children play very different design roles based on their prior knowledge. Children who had experience in the learning environment often were able to improve the practical and pragmatic aspects of technology designed for those environments. On the other hand, children who had explicit design experience were more able to generate open and unconstrained ideas regarding aesthetics, features, and novel ideas related to technology.

Ahn *et al* (2014) and Clegg *et al* (2014) also used participatory design and DBR methods to create a social media application called ScienceKit (see Figure 7.3), where children can share aspects of their daily lives via mechanisms commonly seen in popular apps such as Instagram, but in the process engage in scientific inquiry in everyday life. Their studies illuminate how combining design activities with children with focused studies of their technology use helps researchers to understand: (i) how children learn with social media, as their design ideas and use of technologies directly inform what kind of learning behavior is possible with new tools, (ii) how iterative implementation of the designed technologies with children yield further insights that can be fed back into additional design iterations, and (iii) result in technologies that are usable and engaging, but also theoretically informed to positively benefit children cognitively and socially. By enacting cycles of participatory design, studies of

learning, and implementation, research studies can yield deeper insights about both child–computer interaction and issues of children's social and cognitive development.

Figure 7.3 Children using the ScienceKit app which was developed as part of a design-based research project.

Source: **Ahn *et al*, Seeing the Unseen Learner: Designing and Using Social Media to Recognize Children's Science Dispositions in Action. 2014. Reproduced with permission of Taylor and Francis Group LLC.**

Duveskog *et al* (2009) designed a story-based interactive digital platform to educate children about HIV and AIDS in southern Tanzania. They included secondary school children, university counseling students, HIV counseling experts, and experts in ICT in their team; groups were involved at different times through the design process. For example, before the implementation, interviews were conducted with secondary school children to elicit stories of their HIV and AIDS experiences. Other students produced drawings to illustrate their stories. Later in development, students in a local drama group recorded voices for the characters in the stories, and once a pilot system was developed, counseling students tested the platform. Using these different forms of communication helped the students to think about and communicate their ideas and feelings. ■

What the examples in Box 7.2 demonstrate is that technology developers have to be prepared to adapt their data collection techniques to suit the participants with whom they work – in those cases, children. Similarly, different approaches are needed when working with users from different cultures. Winschiers-Theophilus *et al* (2012) comment that: “Many attempts have been made to adapt participatory and user-centered design methods to specific regions by localizing usability measures or incorporating cultural models of people's interpersonal interactions and communicative habits into analytic tools. However, our failure to successfully apply user-centered methods, evaluations, or benchmarks in developing regions, or to assess the efficacy of cross-cultural projects according to ‘universally valid’ a priori measures calls for the reframing of relationships between cultural contexts and meaning in design” p. 90. In their work with local communities in Namibia they had to find ways of involving the local participants, which included developing a variety of visual and other techniques to communicate ideas and capture the collective understanding and feelings inherent in the local cultures of the people with whom they worked. (See also Winschiers-Theophilus and Bidwell (2013) and Case Study 11.3.)

7.4.4 Focus Groups

Interviews are often conducted with one interviewer and one interviewee, but it is also common to interview people in groups. One form of group interview that is frequently used in marketing, political campaigning, and social sciences research is the focus group.

Normally three to ten people are involved, and the discussion is led by a trained facilitator. Participants are selected to provide a representative sample of the target population. For example, in an evaluation of a university website, a group of administrators, faculty, and students may form three separate focus groups because they use the web for different purposes. In requirements activities it is quite common to hold a focus group in order to identify conflicts in terminology or expectations from different sections within one department or organization.

The benefit of a focus group is that it allows diverse or sensitive issues to be raised that might otherwise be missed. The method assumes that individuals develop opinions within a social context by talking with others, which means that this approach is more appropriate for investigating community issues rather than individual experiences. Focus groups aim to enable people to put forward their own opinions in a supportive environment. Apreset agenda is developed to guide the discussion, but there is sufficient flexibility for the facilitator to follow unanticipated issues as they are raised. The facilitator guides and prompts discussion and skillfully encourages quiet people to participate and stops verbose ones from dominating the discussion. The discussion is usually recorded for later analysis and participants may be invited to explain their comments more fully at a later date.

7.4.5 Planning and Conducting an Interview

Planning an interview involves developing the set of questions or topics to be covered, collating any documentation to give to the interviewee (such as consent form or project description), checking that recording equipment works in advance and you know how to use it, working out the structure of the interview, and organizing a suitable time and place.

Developing Interview Questions

Questions for an interview may be open or closed. Open questions are best suited where the goal of the session is exploratory; closed

questions can only be used where the possible answers are known in advance. An unstructured interview will usually consist entirely of open questions, while a structured interview will usually consist of closed questions. A semi-structured interview may use a combination of both types.

Dilemma

What They Say and What They do

What users say isn't always what they do. When asked a question, people sometimes give the answers that they think show them in the best light, or they may just forget what happened, or they may want to please the interviewer by answering in the way they anticipate will satisfy the interviewer. For example, in a study looking at the maintenance of telecommunications software, the developers stated that most of their job involved reading documentation, but when observed, it was found that searching and looking at source code was much more common than looking at documentation (Singer *et al*, 1997).

So, can interviewers believe all the responses they get? Are the respondents telling the truth or are they simply giving the answers that they think the interviewer wants to hear?

It isn't possible to avoid this behavior, but it is important to be aware of it and to reduce such biases by choosing questions carefully, getting a large number of participants, or by using a combination of data gathering techniques. ■

The following guidelines for developing interview questions are derived from Robson (2011):

- Compound sentences can be confusing, so split them into two separate questions. For example, instead of, 'How do you like this smartphone app compared with previous ones that you have owned?' say, 'How do you like this smartphone app?' 'Have you owned other smartphone apps?' If so, 'How did you like them?' This is easier for the interviewee to respond to and easier for the interviewer to record.
- Interviewees may not understand jargon or complex language and might be too embarrassed to admit it, so explain things to them in layman's terms.
- Try to keep questions neutral. For example, if you ask 'Why do you like this style of interaction?' this question assumes that the person does like it and will discourage some interviewees from stating their real feelings.

Activity 7.2

Several e-readers for reading ebooks, watching movies, and browsing photographs are available on the market (see Figure 7.4). These devices are thin and lightweight and are ideally designed for reading books, newspapers, and magazines. The exact design differs between makes and models, but they all support book reading that is intended to be as comfortable as reading a paper book.

Figure 7.4 (a) Sony's e-reader, (b) Amazon's Kindle, and (c) Apple's iPad

Source: **(a) Courtesy of Sony Europe Limited, (b) and (c) ©PA Images.**

The developers of a new e-reader want to find out how appealing it will be to young people under 18 years of age. To this end, they have asked you to conduct some interviews for them.

1. What is the goal of your data gathering session?
2. Suggest ways of recording the interview data.
3. Suggest a set of questions that are suitable for use in an unstructured interview that seek opinions about e-readers and their appeal to the under-18s.
4. Based on the results of the unstructured interviews, the developers of the new e-reader have found that two important acceptance factors are whether the device can be handled easily and whether the typeface and appearance can be altered. Write a set of semi-structured interview questions to evaluate these two aspects. If you have an e-reader available, show it to two of your peers and ask them to comment on your questions. Refine the questions based on their comments.

Comment

1. The goal is to seek opinions about whether e-readers would be appealing to people under 18.
2. Taking notes might be cumbersome and distracting to the interviewee, and it would be easy to miss important points. An alternative is to audio record the session. Video recording is not needed as it isn't necessary to see the interviewee. However, it would be useful to have a camera at hand to take photographs of any aspects of the device referred to by the interviewee.
3. Possible questions include: Do you find reading a book on the e-reader comfortable? Do you know how to turn the page in an ebook? In what way(s) does the e-reader affect your ability to become engrossed in the story you are reading?
4. Semi-structured interview questions may be open or closed. Some closed questions that you might ask include:
   - Have you used an e-reader before?
   - Would you like to read a book using an e-reader?
   - In your opinion, is the e-reader easy to handle?

   Some open questions, with follow-on probes, include:

- What do you like most about the e-reader? Why?
- What do you like least about the e-reader? Why?
- Please give me an example where the e-reader was uncomfortable or difficult to use.

It is helpful when collecting answers to list the possible responses together with boxes that can just be checked (i.e. ticked). Here's how we could convert some of the questions from Activity 7.2.

1. Have you used an e-reader before? (Explore previous knowledge)

   Interviewer checks box ☐ *Yes* ☐ *No* ☐ *Don't remember/know*

2. Would you like to read a book using an e-reader? (Explore initial reaction, then explore the response)

   Interviewer checks box ☐ *Yes* ☐ *No* ☐ *Don't know*

3. Why?

   If response is 'Yes' or 'No,' interviewer says, 'Which of the following statements represents your feelings best?'

   For 'Yes,' interviewer checks the box

   ☐ *I don't like carrying heavy books*

   ☐ *This is fun/cool*

   ☐ *My friend told me they are great*

   ☐ *It's the way of the future*

   ☐ *Another reason (interviewer notes the reason)*

   For 'No,' interviewer checks the box

   ☐ *I don't like using gadgets if I can avoid it*

   ☐ *I can't read the screen clearly*

   ☐ *I prefer the feel of paper*

   ☐ *Another reason (interviewer notes the reason)*

4. In your opinion, is an e-reader easy to handle or cumbersome?

   Interviewer checks box

   ☐ *Easy to handle*

   ☐ *Cumbersome*

   ☐ *Neither*

Running the Interview

Before starting, make sure that the aims of the interview have been communicated to and understood by the interviewees, and they feel comfortable. Some simple techniques can help here, such as finding out about their world before the interview so that you can dress, act, and speak in a manner that will be familiar. This is particularly important when working with children, seniors, people from different ethnic and cultural groups, people who have disabilities, and seriously ill patients.

During the interview, it is better to listen more than to talk, to respond with sympathy but without bias, and to appear to enjoy the interview (Robson, 2011). Robson suggests the following steps for conducting an interview:

1. An introduction in which the interviewer introduces herself and explains why she is doing the interview, reassures interviewees regarding any ethical issues, and asks if they mind being recorded, if appropriate. This should be exactly the same for each interviewee.
2. A warm-up session where easy, non-threatening questions come first. These may include questions about demographic information, such as 'What area of the country do you live in?'
3. A main session in which the questions are presented in a logical sequence, with the more probing ones at the end. In a semi-structured interview the order of questions may vary between participants, depending on the course of the conversation, how much probing is done, and what seems more natural.
4. A cool-off period consisting of a few easy questions (to defuse any tension that may have arisen).
5. A closing session in which the interviewer thanks the interviewee and switches off the recorder or puts her notebook away, signaling that the interview has ended.

7.4.6 Other Forms of Interview

Conducting face-to-face interviews and focus groups can sometimes be impractical, especially when the participants live in different geographical areas. Skype, email, and phone-based interactions, sometimes with screen-sharing software, are therefore increasing in popularity. These are carried out similarly to face-to-face sessions, although such issues as dropped Skype connections and insufficient Internet bandwidth for reliable video can be a challenge to conducting them. However, there are some advantages to remote focus groups and interviews, especially when done through audio-only channels. For example, the participants are in their own environment and are more relaxed, participants don't have to worry about what they wear, who other people are, or interact in an unnatural environment surrounded by strangers; for interviews that involve sensitive issues, interviewees may prefer to be anonymous. In addition, participants can leave the conversation whenever they want to by just putting down the phone, which adds to their sense of security. While it is questionable whether data collected face-to-face can be compared directly with data collected remotely, it seems that remote phone-based group or individual interviews are preferable at least in some circumstances.

> Link to more information on telephone focus groups, at http://mnav.com/shocking-truth/ and for some interesting thoughts on remote usability testing, see http://www.uxbooth.com/articles/hidden-benefits-remote-research/

Retrospective interviews, i.e. interviews which reflect on an activity or a data gathering session in the recent past, may be conducted with participants to check that the interviewer has correctly understood what was happening.

7.4.7 Enriching the Interview Experience

Face-to-face interviews often take place in a neutral environment, e.g. a meeting room away from the interviewee's normal place of work or their home. In such situations the interview location provides an artificial context that is different from the interviewee's normal tasks. In these circumstances it can be difficult for interviewees to give full answers to the questions posed. To help combat this, interviews can be enriched by using props such as prototypes or work artifacts that the interviewee or interviewer brings along, or descriptions of common tasks (examples of these kinds of props are scenarios and prototypes, which are covered in Chapters 10 and 11). These props can be used to provide context for the interviewees and help to ground the data in a real setting. Figure 7.5 illustrates the use of prototypes in a focus group setting.

Figure 7.5 Enriching a focus group with prototypes. Here storyboards are displayed on the wall for all participants to see

For example, Jones *et al* (2004) used diaries as a basis for interviews. They performed a study to probe the extent to which certain places are associated with particular activities and information needs. Each participant was asked to keep a diary in which they entered information about where they were and what they were doing at 30 minute intervals. The interview questions were then based around their diary entries.

7.5 Questionnaires

Questionnaires are a well-established technique for collecting demographic data and users' opinions. They are similar to interviews in that they can have closed or open questions but they can be distributed to a larger number of participants so more data can be collected than would normally be possible in an interview study. Furthermore, the issues of involving people who are located in remote locations or cannot attend an interview at a particular time can be dealt with more conveniently. Often a message is sent electronically to potential participants to direct them to an online questionnaire.

Effort and skill are needed to ensure that questions are clearly worded and the data collected can be analyzed efficiently. Well-designed questionnaires are good at getting answers to specific questions from a large group of people. Questionnaires can be used on their own or in conjunction with other methods to clarify or deepen understanding. For example, information obtained through interviews with a small selection of interviewees might be corroborated by sending a questionnaire to a wider group to confirm the conclusions.

Questionnaire questions and structured interview questions are similar, so how do you know when to use which technique? Essentially, the difference lies in the motivation of the respondent to answer the questions. If you think that this motivation is high enough to complete a questionnaire without anyone else present, then a questionnaire will be appropriate. On the other hand, if the respondents need some persuasion to answer the questions, it would be better to use an interview format and ask the questions face-to-face through a structured interview. For example, structured interviews are easier and quicker to conduct in situations in which people will not stop to complete a questionnaire, such as at a train station or while walking to their next meeting.

It can be harder to develop good questionnaire questions compared with structured interview questions because the interviewer is not available to explain them or to clarify any ambiguities. Because of this, it is important that questions are specific; when possible, closed questions should be asked and a range of answers offered, including a 'no opinion' or 'none of these' option. Finally, negative questions can be confusing and may lead to the respondents giving false information, although some questionnaire designers use a mixture of negative and positive questions deliberately because it helps to check the users' intentions.

7.5.1 Questionnaire Structure

Many questionnaires start by asking for basic demographic information (gender, age, place of birth) and details of relevant experience (the time or number of years spent using computers, or the level of expertise within the domain under study, etc.). This background information is useful for putting the questionnaire responses into context. For example, if two respondents conflict, these different perspectives may be due to their level of experience – a group of people who are using a social networking site for the first time are likely to express different opinions to another group with five years' experience of such sites. However, only contextual information that is relevant to the study goal needs to be collected. In the example above, it is unlikely that the person's shoe size will provide relevant context to their responses!

Specific questions that contribute to the data gathering goal usually follow these more general questions. If the questionnaire is long, the questions may be subdivided into related topics to make it easier and more logical to complete.

The following is a checklist of general advice for designing a questionnaire:

- Think about the ordering of questions. The impact of a question can be influenced by question order.
- Consider whether you need different versions of the questionnaire for different populations.
- Provide clear instructions on how to complete the questionnaire. For example, if only one of the boxes needs to be checked, then say so. Questionnaires can make their message clear with careful wording and good typography.
- Abalance must be struck between using white space and the need to keep the questionnaire as compact as possible.

7.5.2 Question and Response Format

Different formats of question and response can be chosen. For example, with a closed question, it may be appropriate to indicate only one response, or it may be appropriate to indicate several. Sometimes it is better to ask users to locate their answer within a range. Selecting the most appropriate question and response format makes it easier for respondents to answer clearly. Some commonly used formats are described below.

Check Boxes and Ranges

The range of answers to demographic questionnaires is predictable. Gender, for example, has two options, male or female, so providing the two options and asking respondents to circle a response makes sense for collecting this information. Asimilar approach can be adopted if details of age are needed. But since some people do not like to give their exact age, many questionnaires ask respondents to specify their age as a range. Acommon design error arises when the ranges overlap. For example, specifying two ranges as 15–20, 20–25 will cause confusion: which box do people who are 20 years old check? Making the ranges 14–19, 20–24 avoids this problem.

Afrequently asked question about ranges is whether the interval must be equal in all cases. The answer is no – it depends on what you want to know. For example, if you want to identify people who might use a website about life insurance, you will most likely be interested in people with jobs who are 21–65 years old. You could, therefore, have just three ranges: under 21, 21–65, and over 65. In contrast, if you wanted to see how the population's political views varied across the generations, you might be interested in looking at 10-year cohort groups for people over 21, in which case the following ranges would be appropriate: under 21, 22–31, 32–41, etc.

Rating Scales

There are a number of different types of rating scales that can be used, each with its own purpose (see Oppenheim, 1998). Here we describe two commonly used scales: the Likert and semantic differential scales. The purpose of these is to elicit a range of responses to a question that can be compared across respondents. They are good for getting people to make judgments about things, e.g. how easy, how usable, and the like.

The success of Likert scales relies on identifying a set of statements representing a range of possible opinions, while semantic

differential scales rely on choosing pairs of words that represent the range of possible opinions. Likert scales are more commonly used because identifying suitable statements that respondents will understand is easier than identifying semantic pairs that respondents interpret as intended.

Likert scales.

Likert scales are used for measuring opinions, attitudes, and beliefs, and consequently they are widely used for evaluating user satisfaction with products. For example, users' opinions about the use of color in a website could be evaluated with a Likert scale using a range of numbers, as in (1), or with words as in (2):

1. The use of color is excellent (where 1 represents strongly agree and 5 represents strongly disagree):

1	2	3	4	5
☐	☐	☐	☐	☐

2. The use of color is excellent:

strongly agree	agree	OK	disagree	strongly disagree
☐	☐	☐	☐	☐

In both cases, respondents could be asked to tick or ring the right box, number or phrase. Designing a Likert scale involves the following three steps:

1. Gather a pool of short statements about the subject to be investigated. For example, 'This control panel is clear' or 'The procedure for checking credit rating is too complex.' Abrainstorming session with peers in which you identify key aspects to be investigated is a good way of doing this.
2. Decide on the scale. There are three main issues to be addressed here: How many points does the scale need? Should the scale be discrete or continuous? How to represent the scale? See Box 7.3 for more on this topic.
3. Select items for the final questionnaire and reword as necessary to make them clear.

In the first example above, the scale is arranged with 1 as the highest choice on the left and 5 as the lowest choice on the right. While there is no absolute right or wrong way of ordering the numbers, some reseachers prefer to have 1 as the higher rating on the left and 5 as the lowest rating on the right. The logic for this is that first is the best place to be in a race and fifth would be the worst. Other researchers prefer to arrange the scales the other way around with 1 as the lowest on the left and 5 as the highest on the right. They argue that intuitively the higher number suggests the best choice and the lowest number suggests the worst choice. Another reason for going from lowest to highest is that when the results are reported, it is more intuitive for readers to see high numbers representing the best choices. The important things to remember are to decide which way around you will apply the scales, make sure your participants know, and then apply your scales consistently throughout your questionnaire.

Semantic differential scales.

Semantic differential scales explore a range of bipolar attitudes about a particular item. Each pair of attitudes is represented as a pair of adjectives. The participant is asked to place a cross in one of a number of positions between the two extremes to indicate agreement with the poles, as shown in Figure 7.6. The score for the investigation is found by summing the scores for each bipolar pair. Scores can then be computed across groups of participants. Notice that in this example the poles are mixed, so that good and bad features are distributed on the right and the left. In this example there are seven positions on the scale.

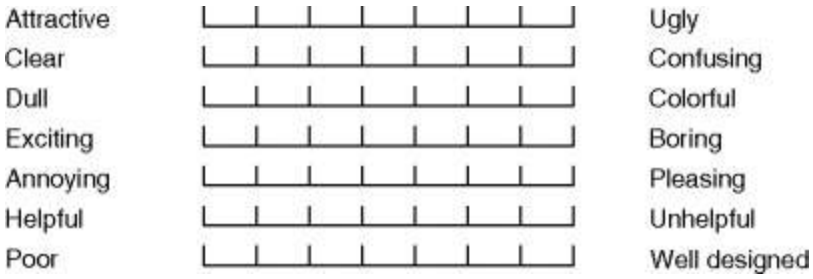


Figure 7.6 An example of a semantic differential scale

BOX 7.3

What Scales to Use: Three, Five, Seven, or More?

When designing Likert and semantic differential scales, issues that need to be addressed include: how many points are needed on the scale, how should they be presented, and in what form?

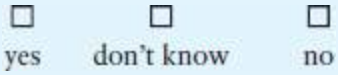
Many questionnaires use seven- or five-point scales and there are also three-point scales. Some even use 9-point scales. Arguments for the number of points go both ways. Advocates of long scales argue that they help to show discrimination. Rating features on an interface is more difficult for most people than, say, selecting among different flavors of ice cream, and when the task is difficult there is evidence to show that people 'hedge their bets.' Rather than selecting the poles of the scales if there is no right or wrong, respondents tend to select values nearer the center. The counter-argument is that people cannot be expected to discern accurately among points on a large scale, so any scale of more than five points is unnecessarily difficult to use.

Another aspect to consider is whether the scale should have an even or odd number of points. An odd number provides a clear central point. On the other hand, an even number forces participants to make a decision and prevents them from sitting on the fence.

We suggest the following guidelines:

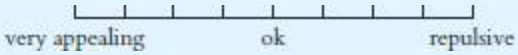
How many points on the scale?

Use a small number, e.g. three, when the possibilities are very limited, as in yes/no type answers:



Use a medium-sized range, e.g. five, when making judgments that involve like/dislike, agree/disagree statements:

Use a longer range, e.g. seven or nine, when asking respondents to make subtle judgments. For example, when asking about a user experience dimension such as 'level of appeal' of a character in a video game:



Discrete or continuous?

Use boxes for discrete choices and scales for finer judgments.

What order?

Decide which way you will order your scale and be consistent. For example, some people like to go from the strongest agreement to the weakest because they find it intuitive to order the scale that way:

–strongly agree

–slightly agree

–agree

–slightly disagree

–strongly disagree.

Activity 7.3

Spot four poorly designed features in the questionnaire in Figure 7.7.

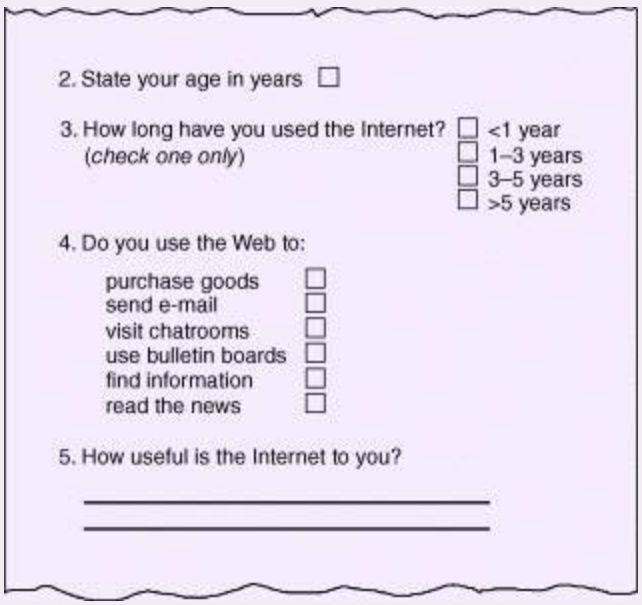


Figure 7.7 A questionnaire with poorly designed features

Comment

Some of the features that could be improved include:

- Question 2 requests exact age. Many people prefer not to give this information and would rather position themselves in a range.
- In question 3, years of experience is indicated with overlapping scales, i.e. 1–3, 3–5. How do you answer if you have 3 years of experience?
- For question 4, the questionnaire doesn't tell you whether you should check one, two, or as many boxes as you wish.
- The space left for people to answer the open-ended question 5 is too small, which will annoy some people and deter them from giving their opinions.

Many online survey tools have the added advantage that they prevent users from making some of these design errors. But it is important to be aware of such things because paper is still sometimes used! ■

7.5.3 Administering Questionnaires

Two important issues when using questionnaires are reaching a representative sample of participants and ensuring a reasonable response rate. For large surveys, potential respondents need to be selected using a sampling technique. However, interaction designers commonly use small numbers of participants, often fewer than 20 users. 100% completion rates are often achieved with these small samples, but with larger or more remote populations, ensuring that surveys are returned is a well-known problem. 40% return is generally acceptable for many surveys, but much lower rates are common. Depending on your audience you might want to consider offering incentives (see Section 7.2.3).

While questionnaires are often web-based, paper questionnaires are used in situations where participants do not have Internet access, such as in airplanes, airports where people are on the move, and in remote areas of the world where the Internet is either not available or very expensive to use. Occasionally, short questionnaires are sent within the body of an email, but more often the advantages of the data being compiled and either partly or fully analyzed make web-based questionnaires attractive. In a recent study by Diaz de Rada and Dominguez-Alvarez (2014), in which the quality of the information collected from a survey given to citizens of Andalusia in Spain was analyzed, several advantages of using web-based versus paper-based questionnaires were identified. These included: a low number of unanswered questions, more detailed answers to open questions, and longer answers to questions in the web questionnaires than in the paper questionnaires. In the five open questions, respondents wrote 63 characters more on the web-based questionnaires than on the paper questionnaires. For the questions in which participants had to select from a drop-down menu, there was a better response rate than when the selection was presented on paper with blank spaces.

Web-based questionnaires are interactive and can include check boxes, radio buttons, pull-down and pop-up menus, help screens, graphics or videos, e.g. Figure 7.8. They can also provide immediate data validation, e.g. the entry must be a number between 1 and

20, and automatically skip questions that are irrelevant to some respondents, e.g. questions only aimed at teenagers. Other advantages of web-based questionnaires include faster response rates and automatic transfer of responses into a database for analysis (Sue and Ritter, 2012).

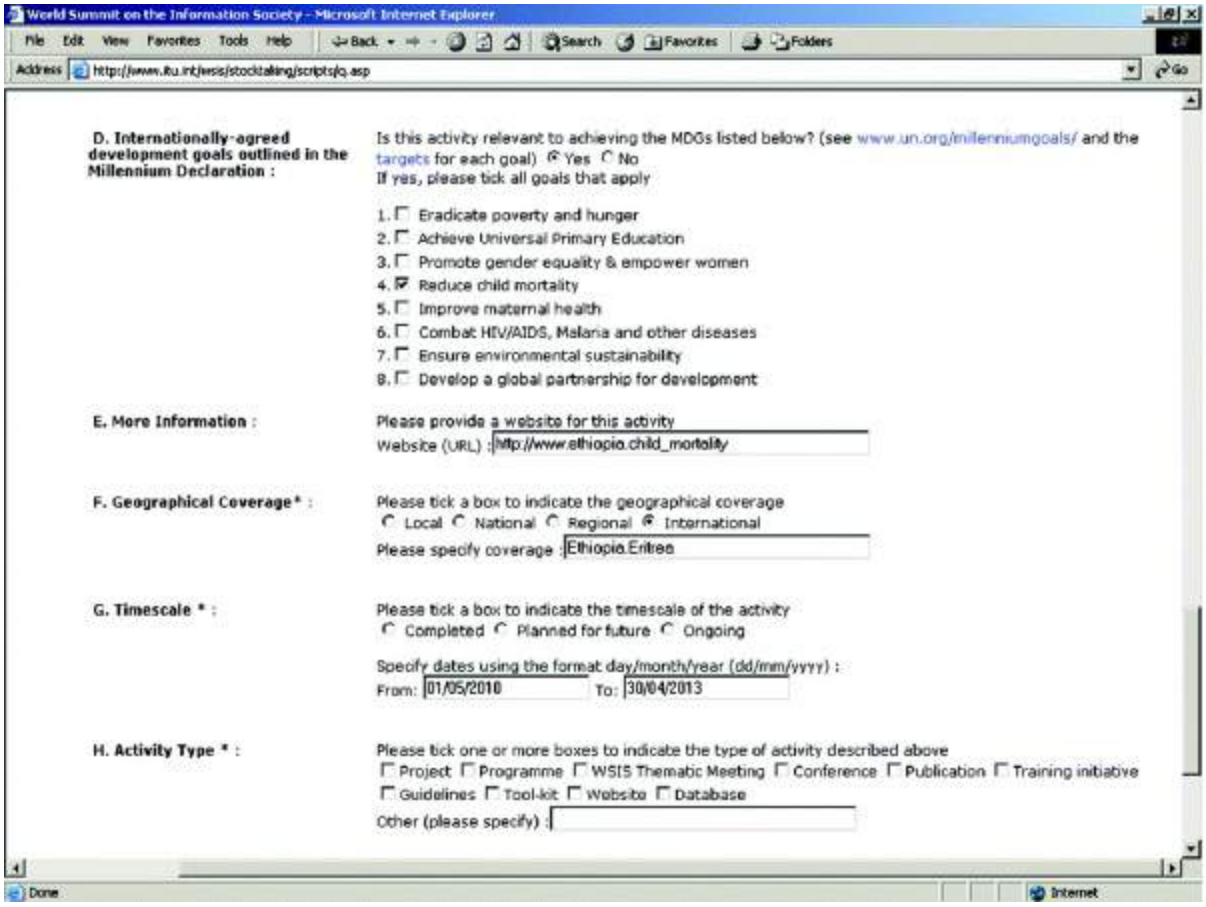


Figure 7.8 An excerpt from a web-based questionnaire showing check boxes, radio buttons, and pull-down menus

The main problem with web-based questionnaires is the difficulty of obtaining a random sample of respondents; web-based questionnaires usually rely on convenience sampling and hence their results cannot be generalized. In some countries, web- and smartphone-based questions are used in conjunction with television to elicit viewers' opinions of programs and political events, e.g. the television program Big Brother.

Deploying a web-based questionnaire involves the following steps (Andrews *et al*, 2003):

1. Design and implement an error-free interactive electronic questionnaire. It may be useful to embed feedback and pop-up help within the questionnaire.
2. Make sure information identifying each respondent can be captured and stored confidentially because the same person may submit several completed surveys. This can be done by recording the Internet domain name or the IP address of the respondent, which can then be transferred directly to a database. However, this action could infringe people's privacy and the legal situation should be checked. Another way is to access the transfer and referrer logs from the web server, which provide information about the domains from which the web-based questionnaire was accessed. Unfortunately, people can still send from different accounts with different IP addresses, so additional identifying information may also be needed.
3. Thoroughly pilot test the questionnaire. This may be achieved in four stages: the survey is reviewed by knowledgeable analysts; typical participants complete the survey using a think-aloud protocol (see below); a small version of the study is attempted; a final check to catch small errors is conducted.

There are many online questionnaire templates available on the web that provide a range of options, including different question types (e.g. open, multiple choice), rating scales (e.g. Likert, semantic differential), and answer types (e.g. radio buttons, check boxes, drop-down menus). The following activity asks you to make use of one of these templates to design a questionnaire for the web.

Activity 7.4

Go to questionpro.com or surveymonkey.com, or a similar survey site that allows you to design your own questionnaire using their set of widgets for a free trial period.

Create a web-based questionnaire for the set of questions you developed for Activity 7.2. For each question produce two different designs, for example radio buttons and drop-down menus for one question; for another question provide a ten-point semantic differential scale and a five-point scale.

What differences (if any) do you think your two designs will have on a respondent's behavior? Ask a number of people to answer one or other of your questions and see if the answers differ for the two designs.

Comment

You may have found that respondents use the response types in different ways. For example, they may select the end options more often from a drop-down menu than from a list of options that are chosen via radio buttons. Alternatively, you may find no difference and that people's opinions are not affected by the widget style used at the interface. Any differences found, of course, may be due to the variation between individual responses rather than being caused by features in the questionnaire design. To tease the effects apart you would need to ask a large number of participants (e.g. in the range 50–100) to respond to the questions for each design. ■

BOX 7.4

Do People Answer Online Questionnaires Differently to Paper and Pencil? If so, Why?

There has been much research examining how people respond to surveys when using a computer compared with the traditional paper and pencil method. Some studies suggest that people are more revealing and consistent in their responses when using a computer to report their habits and behaviors, such as eating, drinking, and amount of exercise, e.g. Luce *et al* (2003). Students have also been found to rate their instructors less favorably when online (Chang, 2004). One reason for this is that students may feel less social pressure when filling in a questionnaire at a computer and hence freer to write the truth than when sitting in a classroom, with others around them, filling out a paper-based version.

Another factor that can influence how people answer questions is the way the information is structured, such as the use of headers, the ordering, and the placement of questions. But the potential may be greater for web-based questionnaires since they provide more opportunities than paper ones for manipulating information (Smyth *et al*, 2004). For example, the use of drop-down menus, radio buttons, and jump-to options may influence how people read and navigate a questionnaire. But do these issues affect respondents' answers? Smyth *et al* (2005) have found that providing forced choice formats results in more options being selected. Another example is provided by Funcke *et al* (2011), who found that continuous sliders enabled researchers to collect more accurate data because they support continuous rather than discrete scales. They also encouraged higher response rates, but they were more challenging for participants who had not encountered continuous scales before and found the concept difficult to understand. ■

7.6 Observation

Observation is a useful data gathering technique at any stage during product development. Early in design, observation helps designers understand the users' context, tasks, and goals. Observation conducted later in development, e.g. in evaluation, may be used to investigate how well the developing prototype supports these tasks and goals.

Users may be observed directly by the investigator as they perform their activities, or indirectly through records of the activity that are studied afterwards. Observation may also take place in the field, or in a controlled environment. In the former case, individuals are observed as they go about their day-to-day tasks in the natural setting. In the latter case, individuals are observed performing specified tasks within a controlled environment such as a usability laboratory.

Activity 7.5

To appreciate the different merits of observation in the field and observation in a controlled environment, read the scenarios below and answer the questions that follow.

Scenario 1. Ausability consultant joins a group of tourists who have been given a wearable navigation device that fits onto a wrist strap to test on a visit to Stockholm. After sightseeing for the day, they use the device to find a list of restaurants within a two-kilometer radius of their current position. Several are listed and they find the telephone numbers of a couple, call them to ask about their menus, select one, make a booking, and head off to the restaurant. The usability consultant observes some difficulty operating the device, especially on the move. Discussion with the group supports the evaluator's impression that there are problems with the interface, but on balance the device is useful and the group is pleased to get a table at a good restaurant nearby.

Scenario 2. Ausability consultant observes how participants perform a pre-planned task using the wearable navigation device in a usability laboratory. The task requires the participants to find the telephone number of a restaurant called Matisse. It takes them several minutes to do this and they appear to have problems. The video recording and interaction log suggest that the interface is very fiddly and the audio interaction is of poor quality, and this is supported by participants' answers on a user satisfaction questionnaire.

1. What are the advantages and disadvantages of these two types of observation?
2. When might each type of observation be useful?

Comment

1. The advantages of the field study are that the observer saw how the device could be used in a real situation to solve a real problem. She experienced the delight expressed with the overall concept and the frustration with the interface. By watching how the group used the device on the move, she gained an understanding of what they liked and what was lacking. The disadvantage is that the observer was an insider in the group, so how objective could she be? The data is qualitative and while anecdotes can be very persuasive, how useful are they? Maybe she was having such a good time that her judgment was clouded and she missed hearing negative comments and didn't notice some people's annoyance. Another study could be done to find out more, but it is not possible to replicate the exact conditions of this study. The advantages of the laboratory study are that it is easier to replicate, so several users could perform the same task, specific usability problems can be identified, users' performance can be compared, and averages for such measures as the time it took to do a specific task and the number of errors can be calculated. The observer could also be more objective because she was an outsider. The disadvantage is that the study is artificial and says nothing about how the device would be used in the real environment.
2. Both types of study have merits. Which is better depends on the goals of the study. The laboratory study is useful for examining details of the interaction style to make sure that usability problems with the interface and button design are diagnosed and corrected. The field study reveals how the navigation device is used in a real-world context and how it integrates with or changes users' behavior. Without this study, it is possible that developers might not have discovered the enthusiasm for the device because the reward for doing laboratory tasks is not as compelling as a good meal! In fact, according to Kjeldskov and Skov (2014), there is no definitive answer to which kind of study is preferable for mobile devices. They suggest that the real question is when and how to engage with in the wild longitudinal studies. ■

7.6.1 Direct Observation in the Field

It can be very difficult for people to explain what they do or to even describe accurately how they achieve a task. So it is very unlikely that an interaction designer will get a full and true story by using interviews or questionnaires. Observation in the field can help fill in details about how users behave and use the technology, and nuances that are not elicited from the other forms of investigation may be observed. This understanding about the context for tasks provides important information about why activities happen the way they do. However, observation in the field can be complicated, and much more difficult to do well than at first appreciated. Observation can also result in a lot of data that is tedious to analyze and not very relevant, especially if the observation is not planned and carried out carefully.

All data gathering should have a clearly stated goal, but it is particularly important to have a focus for an observation session because there is always so much going on. On the other hand, it is also important to be able to respond to changing circumstances: for example, you may have planned one day to observe a particular person performing a task, but you are invited to an unexpected meeting which is relevant to your observation goal, and so it makes sense to attend the meeting instead. In observation there is a careful balance between being guided by goals and being open to modifying, shaping, or refocusing the study as you learn about the situation. Being able to keep this balance is a skill that develops with experience.

Structuring Frameworks for Observation in the Field

During an observation, events can be complex and rapidly changing. There is a lot for observers to think about, so many experts have a framework to structure and focus their observation. The framework can be quite simple. For example, this is a practitioner's framework for use in evaluation studies that focuses on just three easy-to-remember items to look for:

- The person: Who is using the technology at any particular time?
- The place: Where are they using it?
- The thing: What are they doing with it?

Even a simple framework such as this one based on who, where, and what can be surprisingly effective to help observers keep their goals and questions in sight. Experienced observers may, however, prefer more detailed frameworks, such as the one suggested by Robson (2011) which encourages observers to pay greater attention to the context of the activity:

- Space: What is the physical space like and how is it laid out?
- Actors: What are the names and relevant details of the people involved?
- Activities: What are the actors doing and why?
- Objects: What physical objects are present, such as furniture?
- Acts: What are specific individual actions?
- Events: Is what you observe part of a special event?
- Time: What is the sequence of events?
- Goals: What are the actors trying to accomplish?
- Feelings: What is the mood of the group and of individuals?

This framework was devised for any type of observation, so when used in the context of interaction design, it might need to be modified slightly. For example, if the focus is going to be on how some technology is used, the framework could be modified to ask:

- Objects: What physical objects, in addition to the technology being studied, are present, and do they impact on the technology use?

Other modifications might also be useful.

Activity 7.6

1. Find a small group of people who are using any kind of technology, e.g. computers, household or entertainment appliances, and try to answer the question, 'What are these people doing?' Watch for three to five minutes and write down what you observe. When you have finished, note down how you felt doing this, and any reactions in the group of people you observed.
2. If you were to observe the group again, how would you change what you did the first time?
3. Observe this group again for about 10 minutes using Robson's framework.

Comment

1. What problems did you encounter doing this exercise? Was it hard to watch everything and remember what happened? How do you think the people being watched felt? Did they know they were being watched? Perhaps some of them objected and walked away. If you didn't tell them, do you think you should have?
2. The initial goal of the observation, i.e. to find out what the people are doing, was vague, and the chances are that it was quite a frustrating experience not knowing what was significant for answering your question and what could be ignored. The questions used to guide observation need to be more focused. For example, you might ask: What are the people doing with the technology? Is everyone in the group using it? Are they looking pleased, frustrated, serious, happy? Does the technology appear to be central to the users' goals?
3. Hopefully you will have felt more confident this second time, partly because it is the second time you've done some observation, and partly because the framework provided you with a structure for what to look at. ■

Both of the frameworks introduced above are relatively general and could be used in many different types of study, and as a basis for developing your own frameworks.

Degree of Participation

Depending on the type of study, the degree of participation within the study environment varies across a spectrum, which can be characterized as insider at one end and outsider at the other. Where a particular study falls along this spectrum depends on its goal and on the practical and ethical issues that constrain and shape it.

An observer who adopts an approach right at the outsider end of the spectrum is called a passive observer and she will not take any part in the study environment at all. It is difficult to be a truly passive observer if you are in the field, simply because you can't avoid interacting with the activities happening around you. Passive observation is more appropriate in laboratory studies.

An observer who adopts an approach at the insider end of this spectrum is called a participant observer. This means that he attempts, at various levels depending on the type of study, to become a member of the group he is studying. This can be a difficult role to play since being an observer also requires a certain level of detachment, while being a participant assumes a different role. As a participant

observer it is important to keep the two roles clear and separate, so that observation notes are objective, while participation is also maintained. It may not be possible to take a full participant observer approach, for other reasons. For example, you may not be skilled enough in the task at hand, the organization/group may not be prepared for you to take part in their activities, or the timescale may not provide sufficient opportunity to become familiar enough with the task to participate fully. Similarly, if you wish to observe activity in a private place such as the home, full participation would be difficult. Bell, for example, emphasizes the importance of spending time with families and using a range of data gathering including observation (Bell, 2003; Bell *et al*, 2005).

Planning and Conducting an Observation in the Field

The frameworks introduced in the previous section are useful not only for providing focus but also for organizing the observation and data gathering activity. But although choosing a framework is important, it is only one aspect of planning an observation. Other decisions include: the level of participation to adopt; how to make a record of the data; how to gain acceptance in the group being studied; how to handle sensitive issues such as cultural differences or access to private spaces; and how to ensure that the study uses different perspectives (people, activities, job roles, etc.). One way to achieve this last point is to work as a team. This can have several benefits: each person can agree to focus on different people or different parts of the context, thereby covering more ground; observation and reflection can be interweaved more easily when there is more than one observer; more reliable data is likely to be generated because observations can be compared; and results will reflect different perspectives.

Once in the throes of an observation, there are other issues that need to be considered. For example, it will be easier to relate to some people than others and it will be tempting to pay attention to those who receive you well, but everyone in the group needs to be attended to. Observation is a fluid activity, and the study will need refocusing as you reflect upon what has been seen. Having observed for a while, interesting phenomena that seem relevant will start to emerge. Gradually ideas will sharpen into questions that guide further observation.

Observing is an intense and tiring activity, but however tired you are, it is important to check the notes and other records and to write up experiences and observations at the end of each day. If this is not done, then valuable information may be lost as the next day's events override your previous day's impressions. Writing a diary or private blog is one way of achieving this. Any documents or other artifacts that are collected or copied (e.g. minutes of a meeting, or discussion items) should be annotated, describing how they are used and at what stage of the activity. Some observers conducting an observation over several days or weeks take time out of each day to go through their notes and other records.

As notes are reviewed, personal opinion should be separated from observation of what happened, and suggestions of issues for further investigation should be clearly marked. It is also a good idea to check observations with an informant or members of the group to ensure that you have understood what is happening and that your interpretations are accurate.

Dilemma

When should I Stop Observing?

Knowing when to stop doing any type of data gathering can be difficult for novices, but it is particularly tricky in observational studies because there is no obvious ending. Schedules often dictate when your study ends. Otherwise, stop when you stop learning new things. Two indications of having done enough are when you start to see similar patterns of behavior being repeated, or when you have listened to all the main stakeholder groups and understand their perspectives. ■

Ethnography

Ethnography has traditionally been used in the social sciences to uncover the social organization of activities, and hence to understand work. Since the early 1990s it has gained credibility in interaction design, and particularly in the design of collaborative systems: see Box 7.5 and Crabtree (2003). A large part of most ethnographic studies is direct observation, but interviews, questionnaires, and studying artifacts used in the activities also feature in many ethnographic studies. The main distinguishing feature of ethnographic studies compared with other approaches to data gathering is that the aim is to observe a situation without imposing any *a priori* structure or framework upon it, and to view everything as ‘strange.’

BOX 7.5

Ethnography in Requirements

The MERboard is a tool to support scientists and engineers display, capture, annotate, and share information to support the operation of two Mars Exploration Rovers (MERs) on the surface of Mars. The MER (see Figure 7.9) acts like a human geological explorer by collecting samples, analyzing them, and transmitting results back to the scientists on Earth. The scientists and engineers collaboratively analyze the data received, decide what to study next, create plans of action, and send commands to the robots on the surface of Mars.

The requirements for MERboard were identified partly through ethnographic fieldwork, observations, and analysis (Trimble *et al*, 2002). The team of scientists and engineers ran a series of field tests that simulated the process of receiving data, analyzing it, creating plans, and transmitting them to the MERs. The main problems they identified stemmed from the scientists' limitations in displaying, sharing, and storing information (see Figure 7.10a).

Figure 7.9 Mars Exploration Rover

Source: **Reproduced by permission of NASA Jet Propulsion Laboratory (NASA-JPL).**

Figure 7.10 (a) The situation before MERboard; (b) A scientist using MERboard to present information

Source: **J. Trimble, R. Wales and R. Gossweiler (2002): "NASA position paper for the CSCW 2002 workshop on Public, Community and Situated Displays: Merboard".**

These observations led to the development of MERboard (see Figure 7.10b), which contains four core applications: a whiteboard for brainstorming and sketching, a browser for displaying information from the web, the capability to display personal information and information across several screens, and a file storage space linked specifically to MERboard. ■

Ethnography has become popular within interaction design because it allows designers to obtain a detailed and nuanced understanding of people's behavior and the use of technology that cannot be obtained by other methods of data gathering (Bell, 2001; Lazar *et al*, 2010a; Crabtree *et al*, 2009).

The observer in an ethnographic study adopts a participant observer (i.e. insider) role as much as possible (Fetterman, 2010). While participant observation is a hallmark of ethnographic studies, it can be used within other methodological frameworks as well such as within an action research program of study where one of the goals is to change and improve the situation.

Gathering ethnographic data is not hard. You gather what is available, what is 'ordinary,' what it is that people do, say, how they work.

The data collected therefore has many forms: documents, notes of your own, pictures, room layout sketches. Notebook notes may include snippets of conversation and descriptions of rooms, meetings, what someone did, or how people reacted to a situation. Data gathering is opportunistic in that you collect what you can collect and make the most of opportunities as they present themselves. Often, interesting phenomena do not reveal themselves immediately but only later on, so it is important to gather as much as possible within the framework of observation. Initially, time should be spent getting to know the people in the workplace and bonding with them. It is critical, from the very beginning, that they understand why you are there, what you hope to achieve, and how long you plan to be there. Going to lunch with them, buying coffee, and bringing small gifts, e.g. cookies, can greatly help this socialization process. Moreover, it may be during one of the informal gatherings that key information is revealed.

Always show interest in the stories, gripes, and explanations that are provided but be prepared to step back if the phone rings or someone else enters the workspace. Most people will stop mid-sentence if their attention is required elsewhere. Hence, you need to be prepared to switch in and out of their work cycles, moving into the shadow if something happens that needs the person's immediate attention.

A good tactic is to explain to one of the participants during a quiet moment what you think is happening and then let her correct you. It is important not to appear overly keen or obtrusive. Asking too many questions, taking pictures of everything, showing off your knowledge, and getting in their way can be very off-putting. Putting up cameras on tripods on the first day is not a good idea. Listening and watching while sitting on the sidelines and occasionally asking questions is a much better approach. When you have gained the trust and respect of the participants you can then ask if they mind you setting up a video camera, taking pictures, or using a recorder. Even taking pictures with a smartphone can be obtrusive.

The following is an illustrative list of materials that might be recorded and collected during an ethnographic study (adapted from Crabtree, 2003, p. 53):

- Activity or job descriptions.
- Rules and procedures (and so on) said to govern particular activities.
- Descriptions of activities observed.
- Recordings of the talk taking place between parties involved in observed activities.
- Informal interviews with participants explaining the detail of observed activities.
- Diagrams of the physical layout, including the position of artifacts.
- Photographs of artifacts (documents, diagrams, forms, computers, etc.) used in the course of observed activities.
- Videos of artifacts as used in the course of observed activities.
- Descriptions of artifacts used in the course of observed activities.
- Workflow diagrams showing the sequential order of tasks involved in observed activities.
- Process maps showing connections between activities.

Traditionally, ethnographic studies in this field aim to understand what people do and how they organize action and interaction within a particular context of interest to designers. However, recently there has been a trend towards studies that draw more on ethnography's anthropological roots and the study of culture. This trend has been brought about by the perceived need to use different approaches because the computers and other digital technologies, especially mobile devices, are embedded in everyday activity, and not just in the workplace as in the 1990s. Crabtree *et al* (2009) warn that using ethnography to study cultural aspects of a situation requires a different set of approaches and contributes differently to design.

BOX 7.6

Doing Ethnography Online

As collaboration and social activity have moved to having a large online presence, ethnographers have adapted their approach to study social media and the various forms of computer mediated communication (Rotman *et al*, 2012, 2013). This practice has various names, the most common of which are: online ethnography (Rotman *et al*, 2012), virtual ethnography (Hine, 2000), or netnography (Kozinets, 2010). Where a community or activity has both an online and offline presence, it is usual to incorporate both online and offline techniques within the data gathering program. However, where the community or activities of interest exist almost exclusively online, then only online techniques are used and virtual ethnography becomes central.

Why, you may ask, is it necessary to distinguish between online and face-to-face ethnography? Well it is important because interaction online is different from interaction in person. For example, communication in person is richer (through gesture, facial expression, tone of voice, and so on) than online communication, and anonymity is more easily achieved when communicating online. In addition, virtual worlds have a persistence, due to regular archiving, that does not typically occur in face-to-face situations. This makes characteristics of the communication different, which often includes how an ethnographer introduces herself to the community, how she acts within the community, and how she reports her findings.

For large social spaces such as digital libraries or Facebook, there are different ethical issues to consider. For example, it is probably unrealistic to ask everyone using a digital library to sign an informed consent form, yet you do need to make sure that participants understand your involvement in the study and the purpose of the study. Presentation of results will need to be modified too. Quotes from participants in the community, even if anonymized in the report, can easily be attributed by a simple search of the community archive or the IP address of the sender, so care is needed to protect their privacy. ■

7.6.2 Direct Observation in Controlled Environments

Observing users in a controlled environment may occur within a purpose-built usability laboratory, but portable laboratories that can be set up in any room are quite common and this avoids participants having to travel away from their normal environment, and reduces the expenses involved in creating and maintaining a purpose-built usability laboratory. Observation in a controlled environment inevitably takes on a more formal character than observation in the field, and the user is likely to feel apprehensive. As with interviews, discussed in Section 7.4, it is a good idea to prepare a script to guide how the participants will be greeted, be told about the goals of the study and how long it will last, and have their rights explained. Use of a script ensures that each participant will be treated in exactly the same way, which brings more credibility to the results obtained from the study.

The same basic data recording techniques are used for direct observation in the laboratory and field studies (i.e. capturing photographs, taking notes, collecting video, and so on), but the way in which these techniques are used is different. In the laboratory the emphasis is on the details of what individuals do, while in the field the context is important and the focus is on how people interact with each other, the technology, and their environment.

The arrangement of equipment with respect to the participant is important in a controlled study because details of the person's activity need to be captured. For example, one camera might record facial expressions, another might focus on mouse and keyboard activity, and another might record a broad view of the participant and capture body language. The stream of data from the cameras can be fed into a video editing and analysis suite where it is coordinated and time-stamped, annotated, and partially edited (see Chapters 13 and 14).

The Think-Aloud Technique

One of the problems with observation is that the observer doesn't know what users are thinking, and can only guess from what they see. Observation in the field should not be intrusive as this will disturb the very context you are trying to capture, so asking questions of the participant should be limited. However, in a controlled environment, the observer can afford to be a little more intrusive. The think-aloud technique is a useful way of understanding what is going on in a person's head.

Imagine observing someone who has been asked to evaluate the interface of the web search engine Lycos.com. The user, who does not have much experience of web searches, is told to look for a phone for a ten-year-old child. He is told to type 'www.lycos.com' and then proceed however he thinks best. He types the URL and gets a screen similar to the one in Figure 7.11.

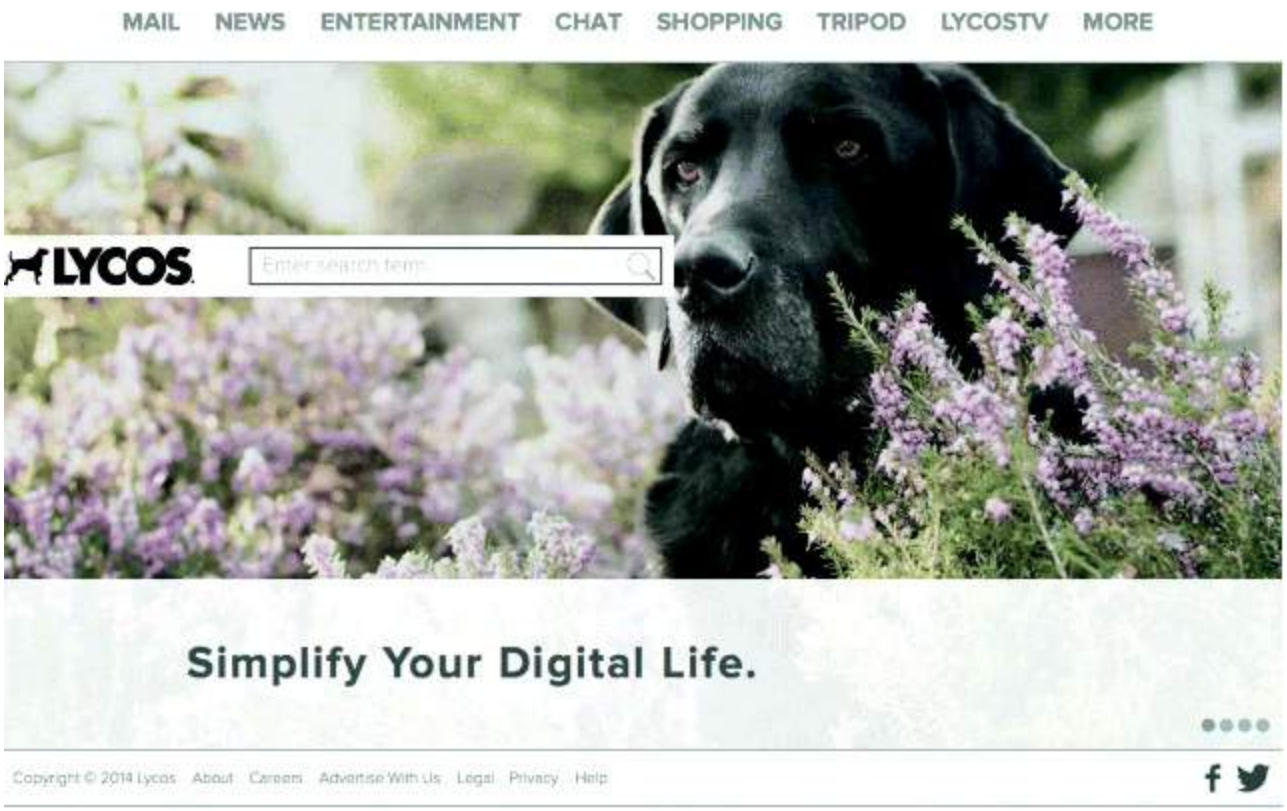


Figure 7.11 Home page of Lycos search engine

Next he types 'child's phone' in the search box. He gets a screen similar to the one in Figure 7.12. He is silent. What is going on, you wonder? What is he thinking? One way around the problem of knowing what he is doing is to collect a think-aloud protocol, a technique developed by Erikson and Simon (1985) for examining people's problem-solving strategies. The technique requires people to say out loud everything that they are thinking and trying to do, so that their thought processes are externalized.

Figure 7.12 The screen that appears in response to choosing 'child's phone'

So, let's imagine an action replay of the situation just described, but this time the user has been instructed to think aloud:

'I'm typing in www.lycos.com, as you told me.' <types>

'Now I am typing 'child's phone' and then clicking on the search button.

<pause and silence>

'It's taking a few seconds to respond.'

'Oh! Now I have a choice of other websites to go to. Hmm, I wonder which one I should select. Well, it's for a young child so I want "child safe phone".' <He clicks to select Smarter.com>

'Gosh, there's a lot more models to select from than I expected! Hmm, some of these are for older children. I wonder what I do next to find one for a ten-year-old.'

<pauses and looks at the screen>

'I guess I should scroll through them and identify those that might be appropriate.'

<silence . . . >

Now you know more about what the user is trying to achieve but he is silent again. You can see that he is looking at the screen. What you don't know is what he is thinking now or what he is looking at.

The occurrence of these silences is one of the biggest problems with the think-aloud technique.

Activity 7.7

Try a think-aloud exercise yourself. Go to a website, such as Amazon or eBay, and look for something that you want to buy. Think aloud as you search and notice how you feel and behave.

Afterwards, reflect on the experience. Did you find it difficult to keep speaking all the way through the task? Did you feel awkward? Did you stop when you got stuck?

Comment

You probably felt self-conscious and awkward doing this. Some people say they feel really embarrassed. At times you may also have started to forget to speak out loud. You may also have found it difficult to think aloud when the task got difficult. In fact, you probably stopped speaking when the task became demanding, and that is exactly the time when an observer is most eager to hear your comments.

If a user is silent during a think-aloud protocol, the observer could interrupt and remind him to think out loud, but that would be intrusive. Another solution is to have two people work together so that they talk to each other. Working with another person (called constructive interaction (Miyake, 1986)) is often more natural and revealing because participants talk in order to help each other along. This technique has been found particularly successful with children and it also avoids possible cultural influences on concurrent verbalization (Clemmensen *et al*, 2008). ■

7.6.3 Indirect Observation: Tracking Users' Activities

Sometimes direct observation is not possible because it is obtrusive or observers cannot be present over the duration of the study, and so activities are tracked indirectly. Diaries and interaction logs are two techniques for doing this.

Diaries

Participants are asked to write a diary of their activities on a regular basis, e.g. what they did, when they did it, what they found hard or easy, and what their reactions were to the situation. For example, Sohn *et al* (2008) asked 20 participants to record their mobile information needs through text messages, and then to use these messages as prompts to help them answer six questions through a website at the end of each day. From the data collected, they identified 16 categories of mobile information needs, the most frequent of which was 'trivia.'

Diaries are useful when participants are scattered and unreachable in person, for example as in many web-based projects. Diaries have several advantages: they do not take up much researcher time to collect data; they do not require special equipment or expertise; and they are suitable for long-term studies. In addition, templates, like those used in open-ended online questionnaires, can be created online to standardize the data entry format so that the data can be entered directly into a database for analysis. However, diary studies rely on participants being reliable and remembering to complete them at the assigned time and as instructed, so incentives may be needed and the process has to be straightforward and quick. Furthermore, studies lasting longer than two weeks are less likely to be successful. Another problem is that the participants' memories of events are often exaggerated or detail is forgotten, e.g. remembering them as better or worse than they really were, or taking more or less time than they actually did take.

The use of multiple media in diaries (e.g. photographs including selfies, audio and video clips, and so on) has been explored by several researchers. Carter and Mankoff (2005) considered whether capturing events through pictures, audio, or artifacts related to the event affects the results of the diary study. They found that images resulted in more specific recall than other media, but audio was useful for capturing events when taking a picture was too awkward. Tangible artifacts, such as those in Figure 7.13, also encouraged

discussion about wider beliefs and attitudes. Several researchers note that collecting diary data from mobile technology users can be particularly tricky when users are constantly on the move (Palen and Salzman, 2002).

Figure 7.13 Some tangible objects collected by participants involved in a study about a jazz festival

Source: **S. Carter and J. Mankoff (2005): "When participants do the capturing: the role of media in diary studies" CHI 2005 pp. 899–908 ©2005 Association for Computing Machinery, Inc. Reprinted by permission.**

The experience sampling method (ESM) is similar to a diary in that it relies on participants recording information about their everyday activities. However, it differs from more traditional diary studies because participants are prompted at random times using a pager, smartphone, or similar device to answer specific questions about their context, feelings, and actions (Hektner *et al*, 2006). These prompts have the benefit of encouraging immediate data capture. For example, Mancini *et al* (2009) used a combination of experience sampling and deferred contextual interviews when investigating mobile privacy. A simple multiple-choice questionnaire was sent electronically to the participants' smartphones, and participants also answered the questions through their smartphones. Interviews about the recorded events were based on the questionnaire answers given at the time.

Interaction Logs and Web Analytics

Interaction logging involves installing software on a device that is being used to record users' activity in a log that can be examined later. A variety of actions may be recorded, from key presses, and mouse or other device movements, to time spent searching a web page, to looking at help systems, and task flow through software modules. If used in a usability evaluation, then gathering of the data is often synchronized with video and audio logs to help evaluators analyze users' behavior and understand how users worked on the tasks they set. Typically usability labs provide this facility.

A key advantage of logging activity is that it is unobtrusive provided system performance is not affected, but it also raises ethical concerns about observing participants if this is done without their knowledge (see the Dilemma box that follows). Another advantage is that large volumes of data can be logged automatically. However, visualization tools are needed to explore and analyze this data quantitatively and qualitatively. Examples of visualizations to help with data analysis and interpretation are in Figures 6.17, 6.18, 8.6, and 8.7.

Web analytics is one form of interaction logging that has become very popular. This involves collecting, analyzing, and reporting data that tracks a user's behavior when interacting with a website. Logging the number of visitors to a website has been common for many years. This kind of data can be used to monitor changes in the number of website visitors after making modifications. Web analytics data tracks users' behavior much more closely, such as how long people stay on a web page, where they spend most of their time, which other sites they came from, what adverts they looked at and for how long, and so on. Web analytics can be used to assess whether users' goals are being met, to support usability studies and to inform future design. They are a powerful tool for business and market research, and can benefit a range of projects. Khoo *et al* (2008) discuss the use of web metrics for digital libraries. They focus particularly on session length as a useful metric, but warn that it is important for any such metrics to be triangulated with other research. This project is discussed further in Box 8.5.

What Are Web Analytics Used for?

Web analytics are a system of tools and techniques for measuring, collecting, analyzing, and reporting web data to understand and optimize web usage. Web analytics help gauge traffic and popularity trends by providing information about the number of website

visitors and number of page views. As well as measuring web traffic, analytics are used in business and market research to assess and improve website effectiveness. Web analytics can further help companies measure the result of print or media advertising campaigns by estimating how traffic to a site changes after launching a campaign.

There are two categories of web analytics: on-site and off-site analytics. On-site analytics are used by website owners to measure visitor behavior and the performance of their website in a commercial context. Data is compared against key performance indicators and used to improve a website or marketing campaign's audience response. Unlike on-site web analytics, off-site analytics measure the performance of a website's potential audience (opportunity), share of voice (visibility), and buzz (comments) on the Internet.

Historically, web analytics has referred to on-site visitor measurement, but in recent years the line between off-site and on-site analytics has blurred because vendors are producing tools spanning both categories. Additional data sources may be conjointly used to augment website behavior data. For instance, email and click-through rates, direct mail campaign data, sales, and history may be paired with web traffic data to provide further insights into user behavior.

Google Analytics is the most widely used on-site web analytics and statistics service, used by over 50% of the 10,000 most popular websites (Empson, 2012). The tool is designed to help Internet marketers and small business owners understand website traffic patterns, sources, and behaviors. The service tracks visitors from referring sites, search engines, social networks, and user visits, and tracks email marketing, pay-per-click networks, and display advertising.

Figure 7.14 shows segments from the Google Analytics dashboard for the accompanying website for this book, id-book.com, for a month in August–September 2014. The first segment shows information about who accessed the site and the second gives some information about the mobile devices used to view the website. These show only a fraction of the information that analytics packages like this can provide. Activity 7.8 asks you to investigate the information shown here.

(a)

	Screen Resolution	Acquisition			Behaviour		
		Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration
		688	74.27%	511	68.46%	2.32	00:01:47
	1. 768x1024	173	60.12%	104	58.38%	2.66	00:02:17
	2. 320x568	137	84.67%	116	79.56%	1.53	00:00:45
	3. 360x640	40	87.50%	35	72.50%	2.00	00:01:50
	4. 320x480	34	82.35%	28	82.35%	1.35	00:00:07
	5. 360x592	19	94.74%	18	78.95%	1.68	00:00:13
	6. 480x800	19	52.63%	10	57.89%	7.05	00:06:24
	7. 720x1280	13	69.23%	9	69.23%	2.77	00:01:35
	8. 1366x768	10	70.00%	7	40.00%	4.20	00:05:52
	9. 1536x864	10	50.00%	5	30.00%	5.20	00:07:35
	10. 1280x800	9	44.44%	4	66.67%	2.00	00:02:23

(b)

Figure 7.14 Segments of the Google Analytics dashboard for id-book.com in September 2014 (a) audience overview, (b) screen resolution of mobile devices used to view the website

Activity 7.8

Consider the two screenshot segments shown in Figure 7.14 from the Google Analytics for id-book.com. Study this information and answer the following questions.

1. How many people visited the site during this period?
2. What do you think someone might look at in 2.32 minutes (the average time they spent at the site)?
3. 'Bounce rate' refers to the percentage of visitors who view just one page of your site. What is the bounce rate for this period? Why do you think this might be a useful metric to capture for any website?
4. Which screen resolution has the highest bounce rate, and which has the lowest? Why do you think that might be?

Comment

1. 4441 users visited the site in this period. Notice that some users must have had more than one session since the number of users is not the same as the number of sessions.
2. The bounce rate is 56.87%. This is a useful metric because it represents a simple but significant characteristic of users' behavior. If the bounce rate is high (40–60% is about average while >65% is high and <35% is low), it deserves further investigation to see if there is a problem with the website.
3. The number of pages per session is about 3, so on average users looked at 3 pages in 2.32 minutes. In 2.32 minutes a user probably won't have played any of the videos on the site, nor read the case studies in any great detail. They may have downloaded PowerPoint slides, however, or looked at the additional resources and then moved on to one of the pointers given for extra information.
4. The highest bounce rate is 82.35% for screen resolutions of 320 × 480. The lowest bounce rate is 30% for screen resolutions of 1536 × 864. Visitors using a lower screen resolution tended to look at only one page on the website, hence they had a high bounce rate. On the other hand, visitors using high resolution screens viewed more pages. This may be because the site is difficult to read or navigate with a low resolution screen, or that the graphics are more appealing on a high resolution screen. ■

In President Obama's 2012 re-election campaign, quick and easy access to actionable data was essential to understanding and responding to voters. Nate Lubin, the Director of Digital Marketing for 'Obama for America', and his team used Google Analytics to inform them when making key decisions quickly during the re-election campaign. The ability to do rapid, real-time optimization, Lubin explained, was particularly important during the presidential debates. Research had shown that 64% of voters used the Internet to verify or 'fact check' a claim made by a candidate, and that voters researched issues discussed during presidential debates in real time. In order to speak to supporters and persuade voters who were researching online during debates, Lubin's team used real-time reports in Google Analytics to understand voters' questions and concerns, allowing them to deliver answers directly from the campaign.

You can view tutorials showing you how to install Google Analytics and more, as detailed below. You can also read a case study about how analytics were used in the wine trade in Chapter 15.

Video on 'Google Analytics Tutorial – Install' to manually install Google Analytics 2013 on your website, at http://youtu.be/P_l4oc6tbYk

Video on 'Google Analytics Tutorial Step-By-Step' for a comprehensive tutorial that describes the statistics included in Google Analytics, and provides insight into how the analytics may be used to improve user traffic, at

http://youtu.be/mm78xlsADgc

For an overview of different dashboards that can be customized in Google Analytics, see Poulter, N. (2013), 6 Google Analytics Custom Dashboards to Save You Time NOW! Retrieved from http://www.stateofdigital.com/google-analytics-dashboards/.

BOX 7.7

Other Analytics Tools

As well as Google Analytics, new tools are emerging that provide additional layers of information, better access control options, and raw and real-time data collection.

- Moz Analytics – Tracks search marketing, social media marketing, brand activity, links and content marketing, and is particularly useful for link management and analysis: www.moz.com
- *TruSocialMetrics* – Tracks social media metrics and helps calculate social media marketing return on investment: www.truesocialmetrics.com
- *Clicky* – Comprehensive and real-time analytics tool that shows individual visitors and the actions they take, and helps define what people from different demographics find interesting: www.clicky.com
- *KISSmetrics* – Detailed analytics tool that displays what website visitors are doing on your website before, during, and after they buy: www.kissmetrics.com
- *Crazy Egg* – Tracks visitor clicks based on where they are specifically clicking, and creates click heat maps useful for website design, usability, and conversion: www.crazyegg.com
- *ClickTale* – Records website visitor actions and uses meta-statistics to create visual heat map reports on customer mouse movement, scrolling, and other visitor behaviors: www.clicktale.com ■

Link to more on this topic, described by Dubois (2014), 11 Best Web Analytics Tools at http://www.inc.com/guides/12/2010/11-best-web-analytics-tools.html

Dilemma

They Don't Know We Are Watching. Shall We Tell Them?

If you have appropriate algorithms and sufficient computer storage, large quantities of data about Internet usage can be collected and users need never know. This information could be very valuable for many different reasons. For example, Google, Facebook, Amazon, and other companies do this so that they can serve advertisements about products or services to their users at appropriate times. This enables the companies to sell more products and services directly or to collect more revenue from advertising companies. Have you ever wondered why, for instance, after searching for information about something recently, such as buying sports gear or a bike, other related products often appear at the side of a future search in the form of advertisements? Are these advertisements helpful or an annoying intrusion? Should users be told that their interactions online are being logged? Knowing this, users will likely change their behavior, which will make their logged data less useful to the company collecting it. What is reasonable? It depends on the context, how much personal information is collected, and how the information will be used. Many companies now tell you that your computer activity and phone calls may be logged for quality assurance and other purposes. Most people do not object to this practice. However, should we be concerned about logging personal information (e.g. discussions about health or financial information)? Should users be worried? How can we exploit the ability to log user behavior when visiting websites without overstepping a person's ethical and civil rights? Where should we draw the line? ■

7.7 Choosing and Combining Techniques

It is desirable to combine data gathering techniques for a single data gathering program; the benefit is to provide multiple perspectives. However, it can be time-consuming and costly. Choosing which data gathering techniques to use depends on a variety of factors related to your goals. There is no right technique or combination of techniques, but some will undoubtedly be more appropriate than others. The decision about which to use will need to be made after taking all the factors into account. Table 7.1 below provides some information to help you choose a set of techniques for a specific project. It lists the kind of information you can get (e.g. answers to specific questions) and the type of data it yields (e.g. mostly qualitative or mostly quantitative). It also includes some advantages and disadvantages for each technique (for a discussion of qualitative and quantitative data, see Section 8.2).

Table 7.1 Overview of data gathering techniques and their use

Technique	Good for	Kind of data	Advantages	Disadvantages
Interviews	Exploring issues	Some quantitative but mostly qualitative	Interviewer can guide interviewee if necessary. Encourages contact between developers and users	Time-consuming. Artificial environment may intimidate interviewee
Focus groups	Collecting multiple viewpoints	Some quantitative but mostly qualitative	Highlights areas of consensus and conflict. Encourages contact between developers and users	Possibility of dominant characters
Questionnaires	Answering specific questions	Quantitative and qualitative	Can reach many people with low resource	The design is crucial. Response rates may be low. Unless carefully designed, the responses may not provide suitable data
Direct observation in the field	Understanding context of user activity	Mostly qualitative	Observing gives insights that other techniques don't give	Very time-consuming. Huge amounts of data are produced
Direct observation in a controlled environment	Capturing the detail of what individuals do	Quantitative and qualitative	Can focus on the details of a task without interruption	Results may have limited use in the normal environment because the conditions were artificial
Indirect observation	Observing users without disturbing their activity; data captured automatically	Quantitative (logging) and qualitative (diary)	User doesn't get distracted by the data gathering; automatic recording means that it can extend over long periods of time	A large amount of quantitative data needs tool support to analyze (logging); participants' memories may exaggerate (diary)

The Focus of the Study

The techniques used must be compatible with the goal of the study, i.e. they must be able to gather appropriate data. For example, the data to be collected may be implicit knowledge or it may be explicit, observable behavior; it may be opinion or it may be facts; it may be formal documented rules or it may be informal work-arounds and heuristics; it may be publicly accessible information or it may be confidential, and so on. The kind of data you want will probably be influenced by where you are in the development cycle. For example, at the beginning of the project you may not have any specific questions that need answering, so it is better to spend time exploring issues through interviews and observation rather than sending out questionnaires.

The activity being investigated will also have dimensions that influence the techniques to use. For example, Olson and Moran (1996) suggest a task can be characterized along three dimensions: (i) is it a set of sequential steps or a rapid overlapping series of subtasks; (ii) does it involve a lot of information and complex displays, or little information and simple representations; and (iii) is the task to be performed by a lay-person or by a trained professional?

The Participants Involved

The characteristics of the target user group for the product will affect the kind of data gathering technique used. For example, techniques used for data gathering from young children may be very different from those used with adults (see Box 7.2). If the participants are in a hurry to catch a plane, they will not be receptive to a long interview; if their job involves interacting with people then they may be comfortable in a focus group, and so on.

The location and accessibility of participants also needs to be considered. It may be attractive to run a focus group for a large set of stakeholders, but if they are spread across a wide geographical area, a face-to-face meeting is unlikely to be practical. Similarly, the time participants need to give their undivided attention to the session is significant, e.g. an interview requires a higher level of active engagement while an observation allows the participant to continue with her normal activity.

Depending on what is motivating the participants to take part, it may be better to conduct interviews rather than to issue a questionnaire. It may also be better to conduct a focus group in order to widen consultation and participation, thereby enhancing feelings of ownership and expectations of the users.

The Nature of the Technique

We have already mentioned the issue of participants' time and the kind of data to be collected, but there is also the issue of whether the technique requires specialist equipment or training, and whether the available investigators have the appropriate knowledge and experience. For example, how experienced is the investigator at conducting ethnographic studies, or in handling video data?

Available Resources

The resources available will influence the choice of techniques, too. For example, sending out questionnaires nationwide requires sufficient time, money, and people to do a good design, pilot it, adapt the questionnaire based on the findings from the pilot study and distribute it, collate the data, and analyze them. If there is very little time and no one on the team has designed a questionnaire before, then the team may run into problems that result in poor data collection.

Activity 7.9

For each of the situations below, consider what kinds of data gathering would be appropriate and how you might use the different techniques introduced above. You should assume that you are at the beginning of product development and that you have sufficient time and resources to use any of the techniques.

1. You are developing a new software system to support a small organic produce shop. There is a system running already with which the users are reasonably happy, but it is looking dated and needs upgrading.
2. You are looking to develop an innovative device for diabetes sufferers to help them record and monitor their blood sugar levels. There are some products already on the market, but they tend to be a bit large and unwieldy. Many diabetes sufferers still rely on manual recording and monitoring methods involving a ritual with a needle or needle-like device, some chemicals, and a written or visual scale.
3. You are developing a website for a young persons' fashion e-commerce site.

Comment

1. As this is a small shop, there are likely to be few stakeholders. Some period of observation is always important to understand the context of the new and the old system. Interviewing the staff rather than giving them questionnaires is likely to be appropriate because there aren't very many of them, and this will yield richer data and give the developers a chance to meet the users. Organic produce is regulated by a variety of laws and so it is important to look at this documentation in order to understand any legal constraints that have to be taken into account. Therefore, we would suggest a series of interviews with the main users to understand the positive and negative features of the existing system, a short observation session to understand the context of the system, and a study of documentation surrounding the regulations.
2. In this case, your user group is spread about, so talking to all of them is not feasible. However, it is important to interview some, possibly at a local diabetic clinic, making sure that you have a representative sample of potential users. And you would need to observe the existing manual operation to understand what is required. Afurther group of stakeholders would be those who use or have used the other products on the market. These stakeholders can be questioned to find out the problems with the existing devices so that the new device can improve on them. Aquestionnaire sent to a wider group in order to back up the findings from the interviews would be appropriate, as might a focus group where possible.
3. Again, you are not going to be able to interview all your users. In fact, the user group may not be very well defined. Interviews backed up by questionnaires and focus groups would be appropriate. In this case, identifying similar or competing sites and evaluating them will help provide information for an improved product. ■