CHAPTER 6
INTERFACES

Objectives

The main aims of this chapter are to:

- Provide an overview of the many different kinds of interfaces.
- Highlight the main design and research issues for each of the interfaces.
- Discuss the difference between graphical (GUIs) and natural user interfaces (NUIs).
- Consider which interface is best for a given application or activity.

6.1 Introduction

Until the mid-1990s, interaction designers concerned themselves largely with developing efficient and effective user interfaces for desktop computers aimed at the single user. This involved working out how best to present information on a screen such that users would be able to perform their tasks, including determining how to structure menus to make options easy to navigate, designing icons and other graphical elements to be easily recognized and distinguished from one another, and developing logical dialog boxes that are easy to fill in. Advances in graphical interfaces, speech, gesture and handwriting recognition, together with the arrival of the Internet, smartphones, wireless networks, sensor technologies, and an assortment of other new technologies providing large and small displays, have changed the face of human– computer interaction. During the last decade, designers have had many more opportunities for designing user experiences. The range of technological developments has encouraged different ways of thinking about interaction design and an expansion of research in the field. For example, innovative ways of controlling and interacting with digital information have been developed that include gesture-based, touch-based, and even brain–computer interaction. Researchers and developers have combined the physical and digital in novel ways, resulting in mixed realities, augmented realities, tangible interfaces, and wearable computing. A major thrust has been to design new interfaces that extend beyond the individual user: supporting small- and large-scale social interactions for people on the move, at home, and at work.

There is now a diversity of interfaces. The goal of this chapter is to consider how to design interfaces for different environments, people, places, and activities. We present a catalog of 20 interface types, starting with command-based and ending with brain–computer. For each one, we present an overview and outline the key research and design concerns. Some are only briefly touched upon while others – that are more established in interaction design – are described in more depth. It should be stressed that the chapter is not meant to be read from beginning to end but dipped into to find out about a particular type of interface.

6.2 Interface Types

Numerous adjectives have been used to describe the different kinds of interfaces that have been developed, including graphical, command, speech, multimodal, invisible, ambient, affective, mobile, intelligent, adaptive, smart, tangible, touchless, and natural. Some of the interface types are primarily concerned with a function (e.g. to be intelligent, to be adaptive, to be ambient, to be smart), while others focus on the interaction style used (e.g. command, graphical, multimedia), the input/output device used (e.g. pen-based, speech-based, gesture-based), or the platform being designed for (e.g. tablet, mobile, PC, wearable). Rather than cover every possible type that has been developed or described, we have chosen to select the main types that have emerged over the last 40 years. The interface types are loosely ordered in terms of when they were developed. They are numbered to make it easier to find a particular one (see Table 6.1 for complete set). It should be noted, however, that this classification is for convenience. The interface entries are not mutually exclusive since some products can appear in two categories. For example, a smartphone can be considered to have either a mobile or touch interface. Table 6.1 suggests which interfaces are related or have design issues in common.

Table 6.1 The types of interfaces covered in this chapter

Interface type	See also
1. Command-based	WIMP and web
2. WIMP and GUI	Augmented and mixed reality
3. Multimedia	Multimedia
4. Virtual reality	Mobile and multimedia
5. Information visualization and dashboards	Mobile
6. Web	Augmented and mixed reality
7. Consumer electronics and appliances	Shareable, touch
8. Mobile	Shareable, air-based gesture
9. Speech	Tangible
10. Pen	Multimodal
11. Touch	Speech, pen, touch, gesture, and haptic
12. Air-based gesture	Touch
13. Haptic	Virtual reality
14. Multimodal	
15. Shareable	
16. Tangible	
17. Augmented and mixed reality	
18. Wearable	
19. Robots and drones	
20. Brain–computer interaction (BCI)	

Table 6.2 A selection of classic HCI videos on the Internet that demonstrate pioneering interfaces

1. The Sketchpad – Ivan Sutherland (1963) describes the first interactive graphical interface http://youtu.be/USyoT_Ha_bA
2. The Mother of All Demos – Douglas Engelbart (1968) describes the first WIMP. http://youtu.be/yJDv-zdhzMY
3. Put that there (1979) – A short video from MIT demonstrating the first speech and gesture interface http://youtu.be/RyBEUyEtxQo
4. Unveiling the genius of multi-touch interface design – Jeff Han's TED talk (2007) http://youtu.be/ac0E6deG4AU
5. Intel's Future Technology Vision (2012) http://youtu.be/g_cauM3kccI

6.2.1 Command-Based

Early interfaces required the user to type in commands that were typically abbreviations (e.g. ls) at the prompt symbol appearing on the computer display, which the system responded to (e.g. by listing current files using a keyboard). Another way of issuing commands is through pressing certain combinations of keys (e.g. Shift+Alt+Ctrl). Some commands are also a fixed part of the keyboard, such as delete, enter, and undo, while other function keys can be programmed by the user as specific commands (e.g. F11 standing for print).

Command line interfaces have been largely superseded by graphical interfaces that incorporate commands such as menus, icons, keyboard shortcuts, and pop-up/predictable text commands as part of an application. Where command line interfaces continue to have an advantage is when users find them easier and faster to use than equivalent menu-based systems (Raskin, 2000) and for performing certain operations as part of a complex software package, such as for CAD environments, (e.g. Rhino3D and AutoCAD), to enable expert designers to be able to interact rapidly and precisely with the software. They also provide scripting for batch operations and are being increasingly used on the web, where the search bar acts as a general-purpose command line facility, e.g. www.yubnub.org. They have also been developed for visually impaired people to enable them to interact in virtual worlds, such as Second Life (see Box 6.1).

BOX 6.1

Command-Based Interfaces for Virtual Worlds

Virtual worlds such as Second Life have become popular places for learning and socializing. Unfortunately people who are visually impaired cannot join in. A command-based interface, called TextSL, was developed to enable them to participate by using a screen reader (Folmer et al, 2009). Commands can be issued to enable the user to move their avatar around, interact with others, and find out about the environment they are in. Figure 6.1 shows that the user has issued the command for their avatar to smile and say hello to other avatars sitting by a log fire. ■

Video demonstration of TextSL at http://youtu.be/0Ba_w7u44MM

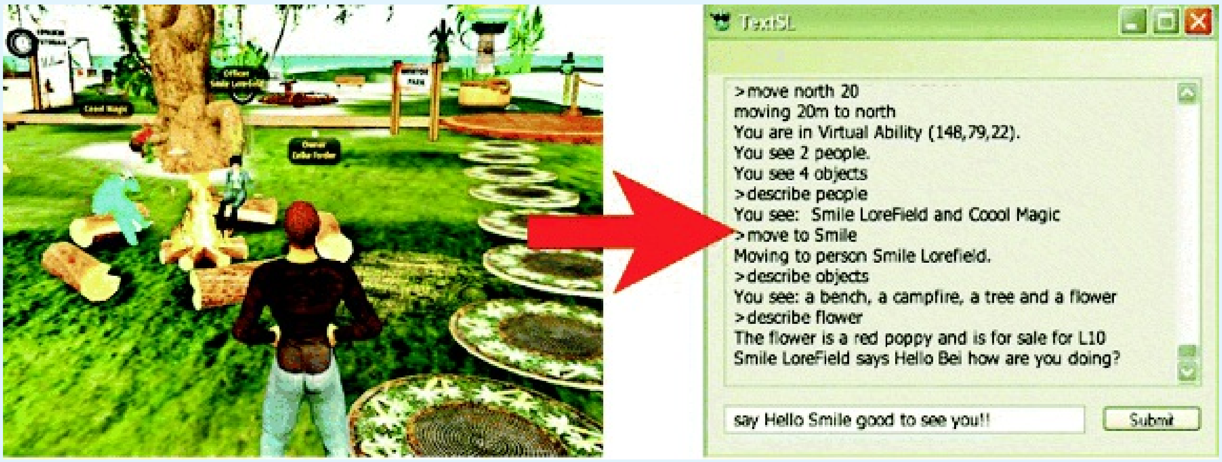


Figure 6.1 Second Life command-based interface for visually impaired users

Source: Reproduced with permission from http://www.eelke.com/images/textsl.jpg.

Research and Design Issues

In the 1980s, much research investigated ways of optimizing command-based interfaces. The form of the commands (e.g. use of abbreviations, full names, familiar names), syntax (e.g. how best to combine different commands), and organization (e.g. how to structure options) are examples of some of the main areas that have been investigated (Shneiderman, 1998). A further concern was which names to use as commands that would be the easiest to remember. A number of variables were tested, including how familiar users were with the chosen names. Findings from a number of studies, however, were inconclusive; some found specific names were better remembered than general ones (Barnard et al, 1982), others showed names selected by users themselves were preferable (e.g. Ledgard et al, 1981; Scapin, 1981), while yet others demonstrated that high-frequency words were better remembered than low-frequency ones (Gunther et al, 1986).

The most relevant design principle is consistency (see Chapter 1). The method used for labeling/naming the commands should be chosen to be as consistent as possible, e.g. always use first letters of operation when using abbreviations. ■

6.2.2 WIMP and GUI

The Xerox Star interface (described in Chapter 2) led to the birth of the WIMP and subsequently the GUI, opening up new possibilities for users to interact with a system and for information to be presented and represented at the interface. Specifically, new ways of visually designing the interface became possible, which included the use of color, typography, and imagery (Mullet and Sano, 1995). The original WIMP comprises:

- Windows (that could be scrolled, stretched, overlapped, opened, closed, and moved around the screen using the mouse).
- Icons (to represent applications, objects, commands, and tools that were opened or activated when clicked on).
- Menus (offering lists of options that could be scrolled through and selected in the way a menu is used in a restaurant).
- Pointing device (a mouse controlling the cursor as a point of entry to the windows, menus, and icons on the screen).

The first generation of WIMP interfaces was primarily boxy in design; user interaction took place through a combination of windows, scroll bars, checkboxes, panels, palettes, and dialog boxes that appeared on the screen in various forms (see Figure 6.2). Application programmers were largely constrained by the set of widgets available to them, of which the dialog box was most prominent. (A widget is a standardized display representation of a control, like a button or scroll bar, that can be manipulated by the user.) The challenge for software developers today is to design GUIs that are best suited for tablet, smartphone, and smartwatch interfaces. Instead of using a mouse and keyboard as input, the default for most users is to swipe and touch using a single finger when browsing and interacting with digital content (for more on this see sections on touch and mobile interfaces).

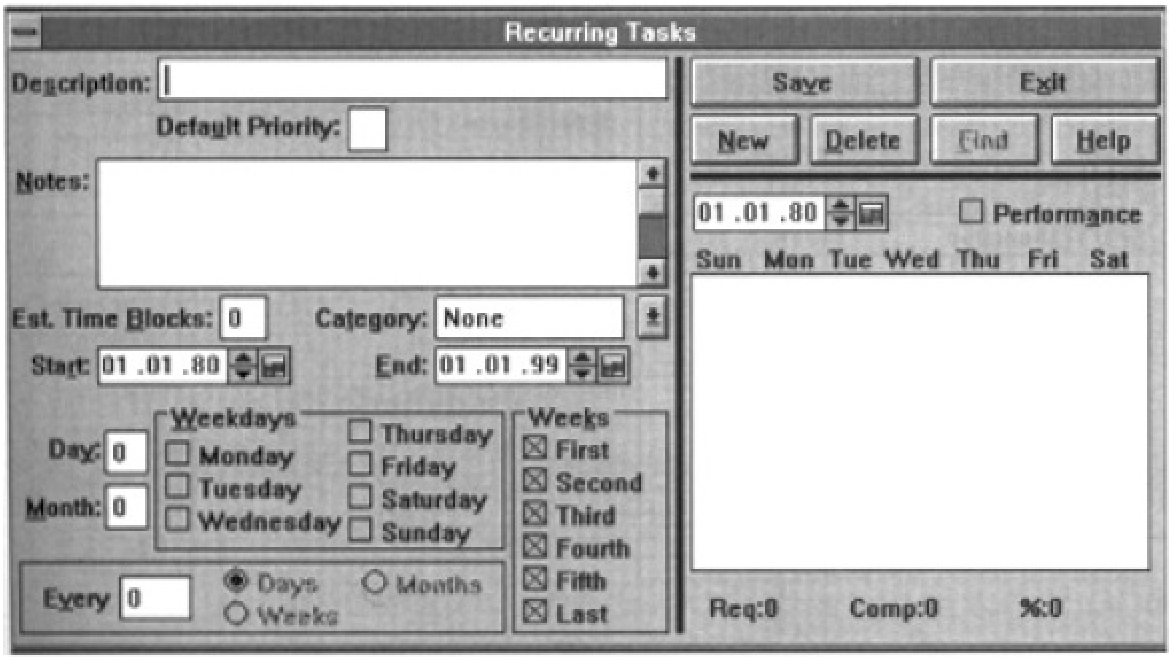


Figure 6.2 The boxy look of the fi rst generation of GUIs. The window presents several check boxes, notes boxes, and options as square buttons

Source: Mullet, Kevin; Sano, Darrell, Designing Visual Interfaces: Communication Oriented Techniques, 1st, © 1995. Reproduced by permission of Pearson Education, Inc., Upper Saddle River, New Jersey.

The basic building blocks of the WIMP are still part of the modern GUI used as part of a computer display, but have evolved into a number of different forms and types. For example, there are now many different types of icons and menus, including audio icons and audio menus, 3D animated icons, and 2D icon-based menus. Windows have also greatly expanded in terms of how they are used and what they are used for; for example, a variety of dialog boxes, interactive forms, and feedback/error message boxes have become pervasive. In addition, a number of graphical elements that were not part of the WIMP interface have been incorporated into the GUI. These include toolbars and docks (a row or column of available applications and icons of other objects such as open files) and rollovers (where text labels appear next to an icon or part of the screen as the mouse is rolled over it). Here, we give an overview of the design issues concerning the basic building blocks of the WIMP/GUI: windows, menus, and icons.

Window design.

Windows were invented to overcome the physical constraints of a computer display, enabling more information to be viewed and tasks to be performed at the same screen. Multiple windows can be opened at any one time, e.g. web pages and word processor documents, enabling the user to switch between them when needing to look or work on different documents, files, and applications. Scrolling bars within windows also enable more information to be viewed than is possible on one screen. Scroll bars can be placed vertically and horizontally in windows to enable upwards, downwards, and sideways movements through a document.

One of the disadvantages of having multiple windows open is that it can be difficult to find specific ones. Various techniques have been developed to help users locate a particular window, a common one being to provide a list as part of an application menu. Mac OS also provides a function that shrinks all windows that are open so they can be seen side by side on one screen. The user needs only to press one function key and then move the cursor over each one to see what they are called. This technique enables users to see at a glance what they have in their workspace and also enables them to easily select one to come to the forefront. Another option is to display all the windows open for a particular application, e.g. Word. The web browser, Safari, has an option of showing 12 shrunken web pages that you have recently visited (history) or most commonly visited (top sites) as a window pane that enables quick scanning (see Figure 6.3).

A particular kind of window that is commonly used in GUIs is the dialog box. Confirmations, error messages, checklists, and forms are presented through them. Information in the dialog boxes is often designed to guide user interaction, with the user following the sequence of options provided. Examples include a sequenced series of forms (i.e. Wizards) presenting the necessary and optional choices that need to be filled in when choosing a PowerPoint presentation or an Excel spreadsheet. The downside of this style of interaction is that there can be a tendency to cram too much information or data entry fields into one box, making the interface confusing, crowded, and difficult to read (Mullet and Sano, 1995).

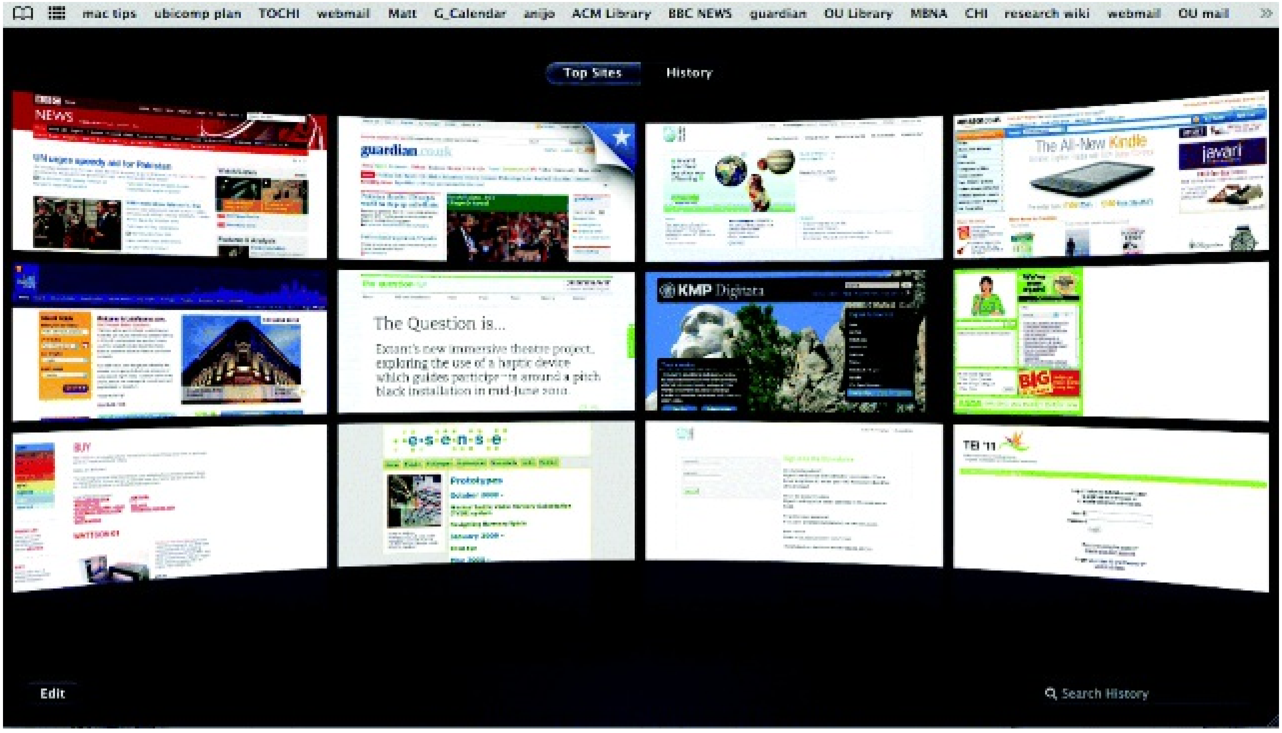


Figure 6.3 A window management technique provided in Safari: pressing the 4 × 3 icon in the top left corner of the bookmarks bar displays the 12 top sites visited, by shrinking them and placing them side by side. This enables the user to see them all at a glance and be able to rapidly switch between them

BOX 6.2

The Joys of Filling in Forms on the Web

For many of us, shopping on the Internet is generally an enjoyable experience. For example, choosing a book on Amazon or flowers from Interflora can be done at our leisure and convenience. The part we don't enjoy, however, is filling in the online form to give the company the necessary details to pay for the selected items. This can often be a frustrating and time-consuming experience. It starts with having to create an account and a new password. Once past this hurdle, a new interactive form pops up for the delivery address and credit card details. The standard online form has a fixed format, making it cumbersome and annoying to fill in, especially for people whose address does not fit within its constraints. Typically, boxes are provided (asterisked for where they must be filled in) for: address line 1 and address line 2, providing no extra lines for addresses that have more than two lines; a line for the town/city; and a line for the zip code (if the site is based in the USA) or other postal code (if based in another country). The format for the codes is different, making it difficult for non-US residents (and US residents for other country sites) to fill in this part. Further boxes are provided for home, work, and cell phone number, and email address (is it really necessary to provide all of these?) and credit card type, name of the owner, and credit card number.

One of the biggest gripes about online registration forms is the country of residence box that opens up as a never-ending menu, listing all of the countries in the world in alphabetical order. Instead of typing in the country they live in, users are forced to select the one they are from, which is fine if they happen to live in Australia or Austria but not if they live in Venezuela or Zambia. Some menus place the host site country first, but this can be easily overlooked if the user is primed to look for the letter of their country (see Figure 6.4).

Figure 6.4 A scrolling menu

Source:Screenshot of Camino browser, ©The Camino Project.

This is an example of where the design principle of recognition over recall (see Chapter 3) does not apply and where the converse is true. A better design might be to have a predictive text option, where users need only to type in the first two or so letters of the country they are from to cause a narrowed-down list of choices to appear that they can then select from at the interface. ■

Activity 6.1

Go to the Interflora site (known as FTD in the US) and click on the international delivery option (its location varies across countries). How are the countries ordered? Is it an improvement to the scrolling pop-up menu?

Comment

Show/Hide

Research and Design Issues

A key research concern is window management – finding ways of enabling users to move fluidly between different windows (and monitors) and to be able to rapidly switch their attention between them to find the information they need or to work on the document/task within each of them – without getting distracted. Studies of how people use windows and multiple monitors have shown that window activation time (i.e. the time a window is open and interacted with) is relatively short, an average of 20 seconds, suggesting that people switch frequently between different documents and applications (Hutchings et al, 2004). Widgets like the taskbar and Jump List (see Figure 6.8) in the Windows environment are used as the main method of switching between windows.

To increase the legibility and ease of use of information presented in windows, the design principles of spacing, grouping, and simplicity should be used (discussed in Chapter 3). An early overview of window interfaces – that is still highly relevant today – is Myers's taxonomy of window manager interfaces (Myers, 1988). ■

Menu design.

Just like restaurant menus, interface menus offer users a structured way of choosing from the available set of options. Headings are used as part of the menu to make it easier for the user to scan through them and find what they want. Figure 6.6 presents two different styles of restaurant menu, designed to appeal to different cultures: Jamie's Italian one is organized into a number of categories including antipasti and sides, pasta, mains, and desserts, while the Japanese is presented in sequential categories: sushi and sashimi, sushi entrée, Japanese entrée platters, and Asian fusion chef speciality. Jamie's Italian menu uses enticing images to depict each category, while the Japanese menu uses a combination of text descriptions for the different choices and one photo of a representative dish for each category. Jamie's Italian requires you to select a category to find out more about the dishes available and their price, whereas the Japanese provides all the information you need to know to make an order.

Figure 6.6 Two different menu layouts

Source: (a) http://www.jamieoliver.com/italian/menu (b) http://www.tonysasianfusion.com/japanesemenu.html.

Interface menu designs have employed similar methods of categorizing and illustrating options available that have been adapted to the medium of the GUI. A difference is that interface menus are typically ordered across the top row or down the side of a screen using category headers as part of a menu bar. The contents of the menus are also for the large part invisible, only dropping down when the header is selected or rolled over with a mouse. The various options under each menu are typically ordered from top to bottom in terms of most frequently used options and grouped in terms of their similarity with one another, e.g. all formatting commands are placed together.

There are numerous menu interface styles, including flat lists, drop-down, pop-up, contextual, and expanding ones, e.g. scrolling and cascading. Flat menus are good at displaying a small number of options at the same time or where the size of the display is small, e.g. cell phones, cameras, MP3 players, smartwatches. However, they often have to nest the lists of options within each, requiring several steps to be taken by a user to get to the list with the desired option. Once deep down in a nested menu, the user then has to take the same number of steps to get back to the top of the menu. Moving through previous screens can be tedious.

Expanding menus enable more options to be shown on a single screen than is possible with a single flat menu list. This makes navigation more flexible, allowing for the selection of options to be done in the same window. However, as highlighted in Figure 6.4 it can be frustrating having to scroll through tens or even hundreds of options. To improve navigation through scrolling menus, a number of novel controls have been devised. For example, the original iPod provided a physical scrollpad that allows for clockwise and anti-clockwise movement, enabling long lists of tunes or artists to be rapidly scrolled through.

The most common type of expanding menu used as part of the PC interface is the cascading one (see Figure 6.7), which provides secondary and even tertiary menus to appear alongside the primary active drop-down menu, enabling further related options to be selected, e.g. selecting track changes from the tools menu leads to a secondary menu of three options by which to track changes in a Word document. The downside of using expanding menus, however, is that they require precise mouse control. Users can often end up making errors, namely overshooting or selecting the wrong options. In particular, cascading menus require users to move their mouse over the menu item, while holding the mouse pad or button down, and then when the cascading menu appears on the screen to move their cursor over to the next menu list and select the desired option. Most of us have experienced the frustration of under- or over-shooting a menu option that leads to the desired cascading menu and worse, losing it as we try to maneuver the mouse onto the secondary or tertiary menu. It is even worse for people who have poor motor control and find controlling a mouse difficult. Menus that are interacted with on smart TVs can also be difficult to navigate because the user has to use a remote button while sitting several feet away from the screen. It is easy to overshoot a letter when trying to type the name of a movie in a search box. Contextual menus provide access to often-used commands associated with a particular item, e.g. an icon. They provide appropriate commands that make sense in the context of a current task. They appear when the user presses the Control key while clicking on an interface element. For example, clicking on a photo in a website together with holding down the Control key results in a small set of relevant menu options appearing in an overlapping window, such as open it in a new window, save it, or copy it. The advantage of contextual menus is that they provide a limited number of options associated with an interface element, overcoming some of the navigation problems associated with cascading and expanding menus. Figure 6.8 shows another kind of contextual window that jumps up as a list.

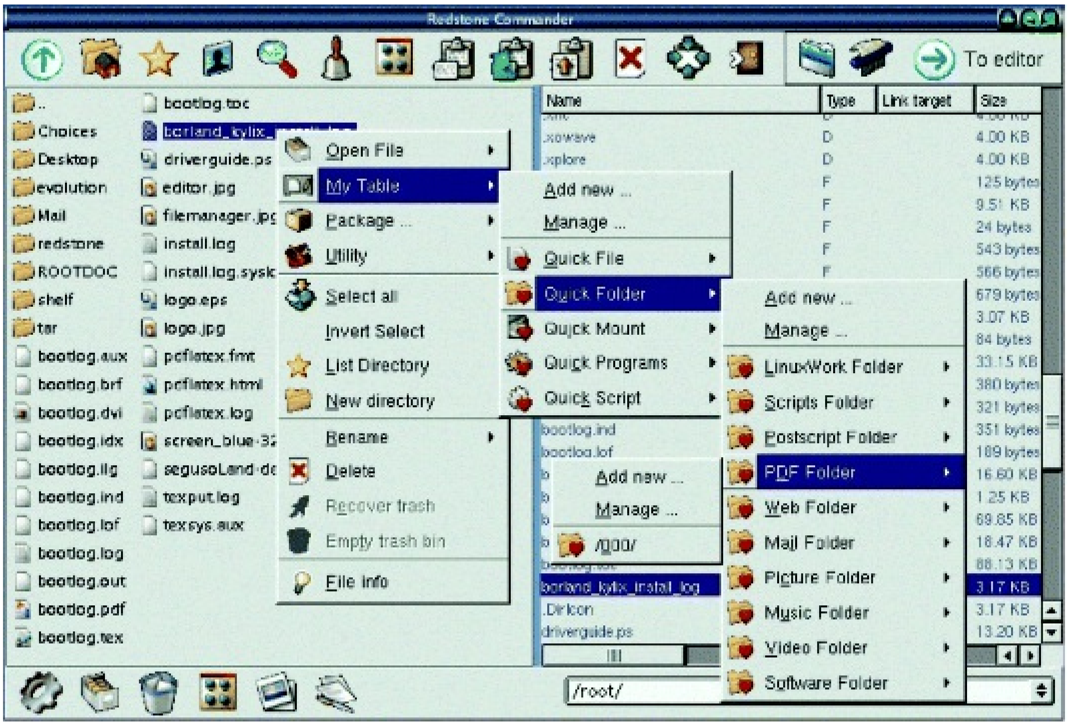


Figure 6.7 A cascading menu

Figure 6.8 Windows jump list

Source: http://windows.microsoft.com/en-US/windows7/products/features/jump-lists.

Activity 6.2

Open an application that you use frequently (e.g. word processor, email, web browser) on a PC/laptop or tablet and look at the menu header names (but do not open them just yet). For each one (e.g. File, Edit, Tools) write down what options you think are listed under each. Then look at the contents under each header. How many options were you able to remember and how many did you put in the wrong category? Now try to select the correct menu header for the following options (assuming they are included in the application): replace, save, spelling, and sort. Did you select the correct header each time or did you have to browse through a number of them?

Comment

Show/Hide

Research and Design Issues

Similar to command names, it is important to decide which are the best terms to use for menu options. Short phrases like ‘bring all to front’ can be more informative than single words like ‘front.’ However, the space for listing menu items is often restricted, such that menu names need to be short. They also need to be distinguishable, i.e. not easily confused with one another so that the user does not choose the wrong one by mistake. Operations such as quit and save should also be clearly separated to avoid the accidental loss of work.

The choice of which type of menu to use will often be determined by the application and type of system. Which is best will depend on the number of options that are on offer and the size of the display to present them in. Flat menus are best for displaying a small number of options at one time, while expanding menus are good for showing a large number of options, such as those available in file and document creation/editing applications.

Icon design.

The appearance of icons at the interface came about following the Xerox Star project. They were used to represent objects as part of the desktop metaphor, namely folders, documents, trashcans, and in- and out-trays. An assumption behind using icons instead of text labels is that they are easier to learn and remember, especially for non-expert computer users. They can also be designed to be compact and variably positioned on a screen.

Icons have become a pervasive feature of the interface. They now populate every application and operating system, and are used for all manner of functions besides representing desktop objects. These include depicting tools (e.g. paintbrush), applications (e.g. web browser), and a diversity of abstract operations (e.g. cut, paste, next, accept, change). They have also gone through many changes in their look and feel: black and white, color, shadowing, photorealistic images, 3D rendering, and animation have all been used.

While there was a period from the late 1980s into the early 1990s when it was easy to find poorly designed icons at the interface (see Figure 6.9), icon design has now come of age. Interface icons look quite different; many have been designed to be very detailed and animated, making them both visually attractive and informative. The result is the design of GUIs that are highly inviting and emotionally appealing, and that feel alive. For example, Figure 6.10 contrasts the simple and jaggy Mac icon designs of the early 1990s with those that were developed as part of the Aqua range for Mac OS X. Whereas early icon designers were constrained by the graphical display technology of the day, they now have more flexibility. For example, the use of anti-aliasing techniques enables curves and non-rectilinear lines to be drawn, enabling more photo-illustrative styles to be developed (anti-aliasing means adding pixels around a jagged border of an object to visually smooth its outline).

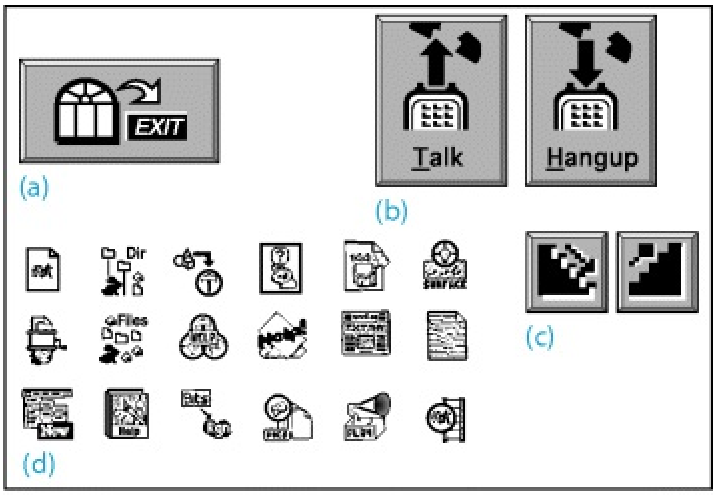


Figure 6.9 Poor icon set from the early 1990s. What do you think they mean and why are they so bad?

Source:K. Mullet and D. Sano: "Designing Visual Interfaces" Pearson 1995, reproduced with permission of Pearson Education.

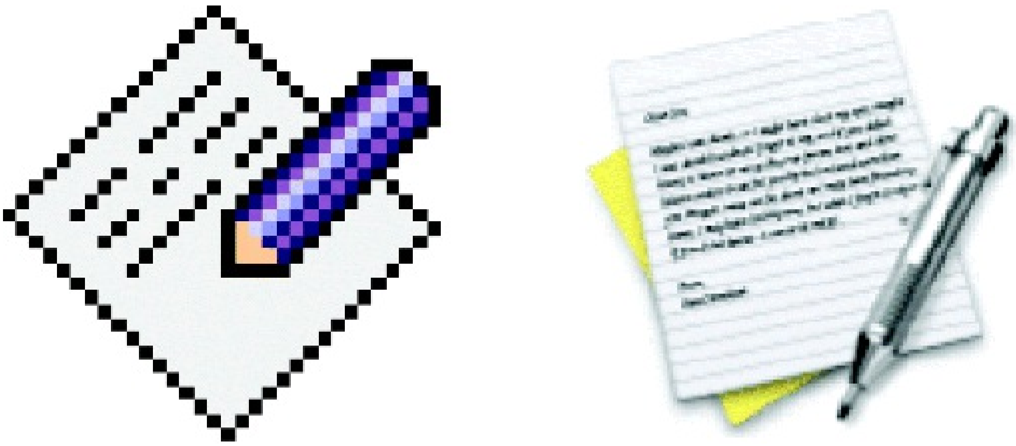


Figure 6.10 Early and more recent Mac icon designs for the TextEdit application

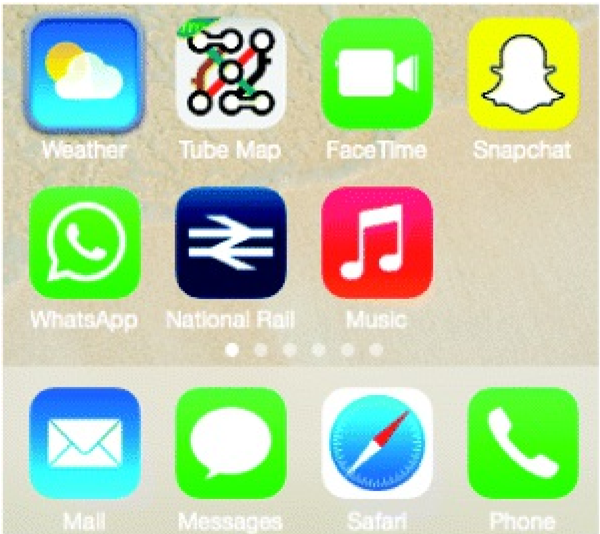
Icons can be designed to represent objects and operations at the interface using concrete objects and/or abstract symbols. The mapping between the representation and underlying referent can be similar (e.g. a picture of a file to represent the object file), analogical (e.g. a picture of a pair of scissors to represent cut), or arbitrary (e.g. the use of an X to represent delete). The most effective icons are generally those that are isomorphic since they have direct mapping between what is being represented and how it is represented. Many operations at the interface, however, are of actions to be performed on objects, making it more difficult to represent them using direct mapping. Instead, an effective technique is to use a combination of objects and symbols that capture the salient part of an action through using analogy, association, or convention (Rogers, 1989). For example, using a picture of a pair of scissors to represent cut in a word-processing application provides sufficient clues as long as the user understands the convention of cut for deleting text.

The greater flexibility offered by current GUI interfaces has enabled developers to create icon sets that are distinguishable, identifiable, and memorable. For example, different graphical genres have been used to group and identify different categories of icons. Figure 6.11 shows how colorful photo-realistic images have been used, each slanting slightly to the left, for the category of user applications (e.g. email) whereas monochrome straight on and simple images have been used for the class of utility applications (e.g. printer setup). The former have a fun feel to them, whereas the latter have a more serious look about them.

Another approach that many smartphone designers use is flat 2D icons. These are very simple and use strong colors and pictograms or symbols. The effect is to make them easily recognizable and distinctive. Examples shown in Figure 6.12 include the white ghost on a yellow background (Snapchat), a white line bubble with a solid white phone handset on a lime-green background (WhatsApp), and the sun next to a cloud (weather).

Icons that appear in toolbars or palettes as part of an application or presented on small device displays (e.g. digital cameras, smartwatches) have much less screen real estate available. Because of this, they have been designed to be simple, emphasizing the outline form of an object or symbol and using only grayscale or one or two colors. They tend to convey the tool and action performed on them using a combination of concrete objects and abstract symbols, e.g. a blank piece of paper with a plus sign representing a new blank document, an open envelope with an arrow coming out of it indicating a new message has arrived.

[Figure 6.11](#) **Contrasting genres of Aqua icons used for the Mac. The top row of icons have been designed for user applications and the bottom row for utility applications**



[Figure 6.12](#) **Flat 2D icons designed for smartphone apps**

Activity 6.3

Sketch simple icons to represent the following operations to appear on a digital camera screen:

- Turn image 90 degrees sideways.
- Auto-enhance the image.
- Fix red-eye.
- Crop the image.

Show them to someone else, tell them that they are icons for a new digital camera intended to be really simple to use, and see if they can understand what each represents.

Comment

Show/Hide

Research and Design Issues

There are many icon libraries available that developers can download for free. Various online tutorials and books on how to design icons are also available (e.g. Hicks, 2012) together with sets of proprietary guidelines and style guides. For example, Apple provides its developers with style guides, explaining why certain designs are preferable to others and how to design icon sets. On its developers' website (developer.apple.com), advice is given on how and why certain graphical elements should be used when developing different types of icon. Among the various guidelines, it suggests that different categories of application (e.g. user, utility) should be represented by a different genre and recommends displaying a tool to communicate the nature of a task, e.g. a magnifying glass for searching, a camera for a photo editing tool. Microsoft also provides extensive guidance and step-by-step procedures on how to design icons for its applications on its website.

To help disambiguate the meaning of icons, text labels can be used under, above, or to the side of their icons. This method is effective for toolbars that have small icon sets, e.g. those appearing as part of a web browser, but is not as good for applications that have large icon sets, e.g. photo editing or word processing, since the screen can get very cluttered and busy making it sometimes harder and longer to find an icon. To prevent text/icon clutter at the interface, a rollover function can be used, where a text label appears adjacent to or above an icon after one second of the user holding the cursor over it and for as long as the user keeps the cursor on it. This method allows identifying information to be temporarily displayed when needed. ■

6.2.3 Multimedia

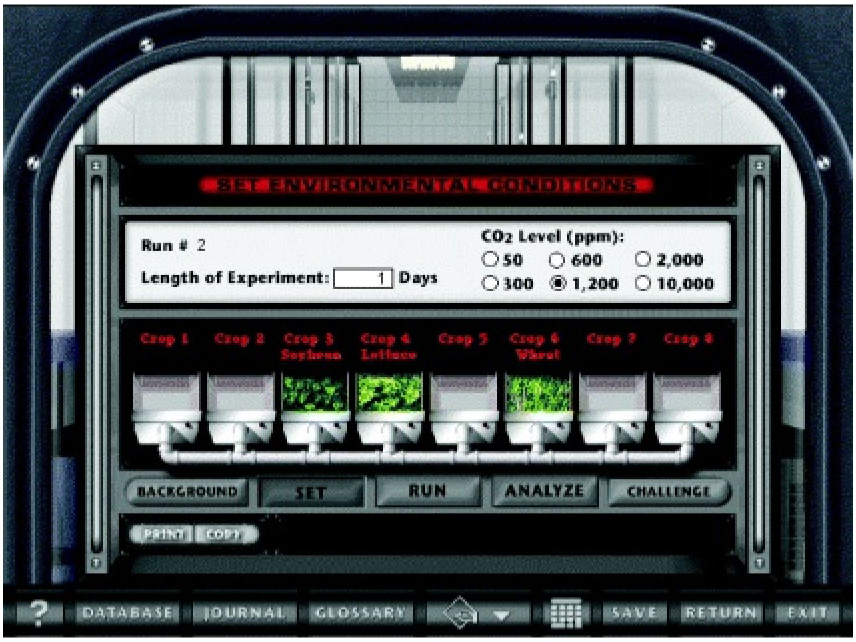
Multimedia, as the name implies, combines different media within a single interface, namely, graphics, text, video, sound, and animations, and links them with various forms of interactivity. It differs from previous forms of combined media, e.g. TV, in that the different media are interactive (Chapman and Chapman, 2004). Users can click on hotspots or links in an image or text appearing on one screen that leads them to another part of the program where, say, an animation or a video clip is played. From there they can return to where they were previously or move on to another place.

Many multimedia narratives and games have been developed that are designed to encourage users to explore different parts of the game or story by clicking on different parts of the screen. An assumption is that a combination of media and interactivity can provide better ways of presenting information than can either one alone. There is a general belief that more is more and the whole is greater than the sum of the parts (Lopuck, 1996). In addition, the added value assumed from being able to interact with multimedia in ways not possible with single media (i.e. books, audio, video) is easier learning, better understanding, more engagement, and more pleasure (see Scaife and Rogers, 1996).

One of the distinctive features of multimedia is its ability to facilitate rapid access to multiple representations of information. Many multimedia encyclopedias and digital libraries have been designed based on this multiplicity principle, providing an assortment of audio and visual materials on a given topic. For example, if you want to find out about the heart, a typical multimedia-based encyclopedia will provide you with:

- One or two video clips of a real live heart pumping and possibly a heart transplant operation.
- Audio recordings of the heart beating and perhaps an eminent physician talking about the cause of heart disease.
- Static diagrams and animations of the circulatory system, sometimes with narration.
- Several columns of hypertext, describing the structure and function of the heart.

Hands-on interactive simulations have also been incorporated as part of multimedia learning environments. An early example was the Cardiac Tutor, developed to teach students about cardiac resuscitation, that required students to save patients by selecting the correct set of procedures in the correct order from various options displayed on the computer screen (Eliot and Woolf, 1994). Several educational websites now provide multimedia educational content. For example, NASA has a multimedia section that provides simulation models based on their research to enable students to develop and test their own designs for a life support system for use on the Moon (see Figure 6.14). The learning environment provides a range of simulators that are combined with online resources.



[Figure 6.14](#) Screen dump from the multimedia environment BioBLAST

Source: Screenshot from BioBlast, ©Wheeling Jesuit University.

Multimedia has largely been developed for training, educational, and entertainment purposes. It is generally assumed that learning (e.g. reading and scientific inquiry skills) and playing can be enhanced through interacting with engaging multimedia interfaces. But what actually happens when users are given unlimited, easy access to multiple media and simulations? Do they systematically switch between the various media and 'read' all the multiple representations on a particular subject? Or, are they more selective in what they look at and listen to?

Activity 6.4

Watch the video of Don Norman appearing in his first multimedia CD-ROM book (1994), where he pops up every now and again in boxes or at the side of the page to illustrate the points being discussed on that page: http://vimeo.com/18687931

What do you think should be included in a modern-day interactive e-textbook? Do you think that as e-textbooks replace paper-based books, students will be tempted to jump to the interactivities – watching the videos, listening to the audios, doing the quizzes, and playing the animations – and maybe not even reading the text?

Comment

Show/Hide

BOX 6.3

Accessible Interactive TV Services for all

TV now provides many digital channels, of which sports, news, and movie channels are very popular. In addition, a range of interactive TV services are being offered that enable users to browse the web, customize their viewing choices, play interactive games, do their banking and shopping, and take an active part in a number of broadcast shows, e.g. voting. Besides offering a wide diversity of choices to the general public, there is much potential for empowering disabled and elderly users, by enabling them to access the services from the comfort of their own armchair. But it requires a sensitivity to interactive design, taking into account specific usability issues for those with impaired motor control, poor vision, and hearing difficulties (Newell, 2003). For example, remote controls need to be designed that can be manipulated with poor dexterity, text/icons need to be readable for those with poor eyesight, while navigation methods need to be straightforward for viewers who are not experienced with multimedia-based interfaces. ■

Research and Design Issues

A key research question is how to design multimedia to help users explore, keep track of, and integrate the multiple representations of information provided, be it a digital library, a game, or learning material. As mentioned above, one technique is to provide hands-on interactivities and simulations at the interface that require the user to complete a task, solve a problem, or explore different aspects of a topic. Specific examples include electronic notebooks that are integrated as part of the interface, where users can copy, download, or type in their own material; multiple-choice quizzes that give feedback about how well the user has done; interactive puzzles where the user has to select and position different pieces in the right combination; and simulation-type games where the user has to follow a set of procedures to achieve some goal for a given scenario. Another approach is to employ dynalinking, where information depicted in one window explicitly changes in relation to what happens in another. This can help users keep track of multiple representations and see the relationship between them (Scaife and Rogers, 1996).

Specific guidelines are available that recommend how best to combine multiple media in relation to different kinds of task, e.g. when to use audio with graphics, sound with animations, and so on for different learning tasks. For example, Alty (1991) suggests that audio information is good for stimulating imagination, movies for action information, text for conveying details, whilst diagrams are good at conveying ideas. From such generalizations it is possible to devise a presentation strategy for learning. This can be along the lines of: first, stimulate the imagination through playing an audio clip; then, present an idea in diagrammatic form; then, display further details about the concept through hypertext. ■

6.2.4 Virtual Reality

Virtual reality (VR) uses computer-generated graphical simulations to create “the illusion of participation in a synthetic environment rather than external observation of such an environment” (Gigante, 1993, p. 3). VR is a generic term that refers to the experience of interacting with an artificial environment, which makes it feel virtually real. The term virtual environment (VE) is used more specifically to describe what has been generated using computer technology (although both terms are used interchangeably). Images are displayed stereoscopically to the users – most commonly through shutter glasses – and objects within the field of vision can be interacted with via an input device like a joystick. The 3D graphics can be projected onto CAVE (Cave Automatic Virtual Environment) floor and wall surfaces, desktops, 3D TV, or large shared displays, e.g. IMAX screens. An early example of VR was the Virtual Zoo project. Allison et al (1997) found that people were highly engaged and very much enjoyed the experience of adopting the role of a gorilla, navigating the environment, and watching other gorillas respond to their movements and presence.

One of the main attractions of VR is that it can provide opportunities for new kinds of immersive experience, enabling users to interact with objects and navigate in 3D space in ways not possible in the physical world or a 2D graphical interface. The resulting user experience can be highly engaging; it can feel as if one really is flying around a virtual world. People can become completely absorbed by the experience (Kalawsky, 1993). The sense of presence can make the virtual setting seem convincing. By presence is meant “a state of consciousness, the (psychological) sense of being in the virtual environment” (Slater and Wilbur, 1997, p. 605), where someone behaves in a similar way to how they would if at an equivalent real event.

One of the advantages of VR is that simulations of the world can be constructed to have a higher level of fidelity with the objects they represent compared with other forms of graphical interface, e.g. multimedia. The illusion afforded by the technology can make virtual objects appear to be very life-like and behave according to the laws of physics. For example, landing and take-off terrains developed for flight simulators can appear to be very realistic. Moreover, it is assumed that learning and training applications can be improved through having a greater fidelity with the represented world.

Another distinguishing feature of VR is the different viewpoints it can offer. Players can have a first-person perspective, where their view of the game or environment is through their own eyes, or a third-person perspective, where they see the world through an avatar visually represented on the screen. An example of a first-person perspective is that experienced in first-person shooter games such as DOOM, where the player moves through the environment without seeing a representation of themselves. It requires the user to imagine what he might look like and decide how best to move around. An example of a third-person perspective is that experienced in Tomb Raider, where the player sees the virtual world above and behind the avatar of Lara Croft. The user controls Lara's interactions with the environment by controlling her movements, e.g. making her jump, run, or crouch. Avatars can be represented from behind or from the front, depending on how the user controls its movements. First-person perspectives are typically used for flying/driving simulations and games, e.g. car racing, where it is important to have direct and immediate control to steer the virtual vehicle. Third-person perspectives are more commonly used in games, learning environments, and simulations, where it is important to see a representation of self with respect to the environment and others in it. In some virtual environments it is possible to switch between the two perspectives, enabling the user to experience different viewpoints on the same game or training environment.

Early VR was developed using head-mounted displays. However, they were found to be uncomfortable to wear, sometimes causing motion sickness and disorientation. They were also expensive and difficult to program and maintain. VR technology has advanced considerably since the 1990s, with more affordable and comfortable VR headsets (e.g. Oculus Rift) that have more accurate head tracking that allows developers to create more compelling games, movies, and virtual environments (see Figure 4.10).

3D software toolkits are also available that make it much easier to program a virtual environment, e.g. Alice (www.alice.org/). Instead of moving in a physical space with a head-mounted display, users interact with a desktop virtual environment – as they would any other desktop application – using mice, keyboards, or joysticks as input devices. The desktop virtual environment can also be programmed to present a more realistic 3D effect (similar to that achieved in 3D movies shown at IMAX cinemas), requiring users to wear a pair of shutter glasses.

Research and Design Issues

VR has been developed to support learning and training for numerous skills. Researchers have designed applications to help people learn to drive a vehicle, fly a plane, and perform delicate surgical operations – where it is very expensive and potentially dangerous to start learning with the real thing. Others have investigated whether people can learn to find their way around a real building/place before visiting it by first navigating a virtual representation of it, e.g. Gabrielli et al (2000). VEs have also been designed to help people practice social and speaking skills, and confront their social phobias, e.g. Cobb et al (2002) and Slater et al (1999). An underlying assumption is that the environment can be designed as a safe place to help people gently overcome their fears (e.g. spiders, talking in public) by confronting them through different levels of closeness and unpleasantness (e.g. seeing a small virtual spider move far away, seeing a medium one sitting nearby, and then finally touching a large one). Studies have shown that people can readily suspend their disbelief, imagining a virtual spider to be a real one or a virtual audience to be a real audience. For example, Slater et al (1999) found that people rated themselves as being less anxious after speaking to a virtual audience that was programmed to respond to them in a positive fashion than after speaking to virtual audiences programmed to respond to them negatively.

Core design issues that need to be considered when developing virtual environments are: how to prevent users experiencing nausea; determining the most effective ways of enabling users to navigate through them, e.g. first versus third person; how to control their interactions and movements, e.g. use of head and body movements; how best to enable them to interact with information in them, e.g. use of keypads, pointing, joystick buttons; and how to enable users to collaborate and communicate with others in the virtual environment. A central concern is the level of realism to aim for. Is it necessary to design avatars and the environments they inhabit to be life-like, using rich graphics, or can simpler and more abstract forms be used, but which nonetheless are equally capable of engendering a sense of presence? For more on this topic see the dilemma box below. ■

Dilemma

Realism Versus Abstraction?

One of the challenges facing interaction designers is whether to use realism or abstraction when designing an interface. This means designing objects either to (i) give the illusion of behaving and looking like real-world counterparts or (ii) appear as abstractions of the objects being represented. This concern is particularly relevant when implementing conceptual models that are deliberately based on an analogy with some aspect of the real world. For example, is it preferable to design a desktop to look like a real desktop, a virtual house to look like a real house, or a virtual terrain to look like a real terrain? Or, alternatively, is it more effective to design representations as simple abstract renditions, depicting only a few salient features?

One of the main benefits of using realism at the interface is that it can enable people to feel more comfortable when first learning an application. The rationale behind this is that such representations can readily tap into people's understanding of the physical world. Hence, realistic interfaces can help users initially understand the underlying conceptual model. In contrast, overly schematic and abstract representations can appear to be off-putting to the newcomer. The advantage of more abstract interfaces, however, is that they can be more efficient to use. Furthermore, the more experienced users become, the more they may find comfortable interfaces no longer to their liking. A dilemma facing designers, therefore, is deciding between creating interfaces to make novice users feel comfortable (but experienced users are less comfortable) and designing interfaces to be effective for more experienced users (but maybe harder to learn by novices).

Mullet and Sano (1995) pointed out how early 3D graphical renditions of objects such as a desk suffered from both an unnatural point of view and an awkward rendering style. One reason for this is that these kinds of 3D depictions conflict with the effective use of display space, especially when 2D editing tasks need to be performed. As can be seen in Figure 6.15, these kinds of task were represented as flat buttons that appear to be floating in front of the desk, e.g. mail, program manager, task manager.

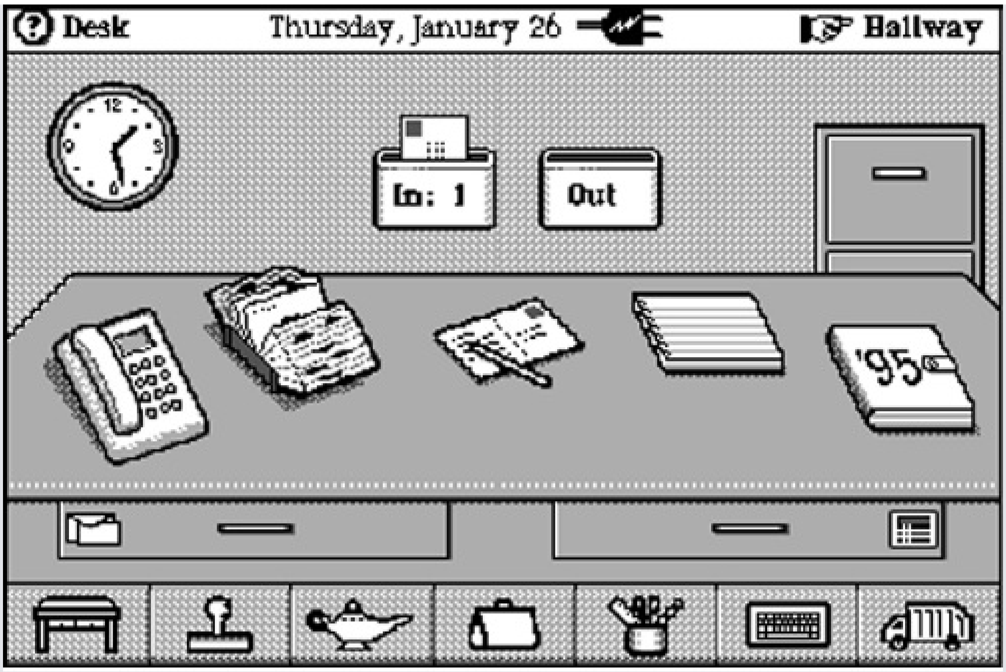


Figure 6.15 Magic Cap's 3D desktop interface

Source: Reprinted by permission of General Magic Inc.

For certain kinds of applications, using realism can be very effective for both novices and experienced users. Computer-based games fall into this category, especially those where users have to react rapidly to dynamic events that happen in a virtual world in real time, say flying a plane or playing a game of virtual football. Making the characters in the game resemble humans in the way they look, move, dress, and behave also makes them seem more convincing and lifelike, enhancing the enjoyment and fun factor. ■

Activity 6.5

Many games have been ported from the PC platform to cell and smartphones. Because of the screen size, however, they tend to be simpler and sometimes more abstract. To what extent does this adaptation of the interface affect the experience of playing the same game?

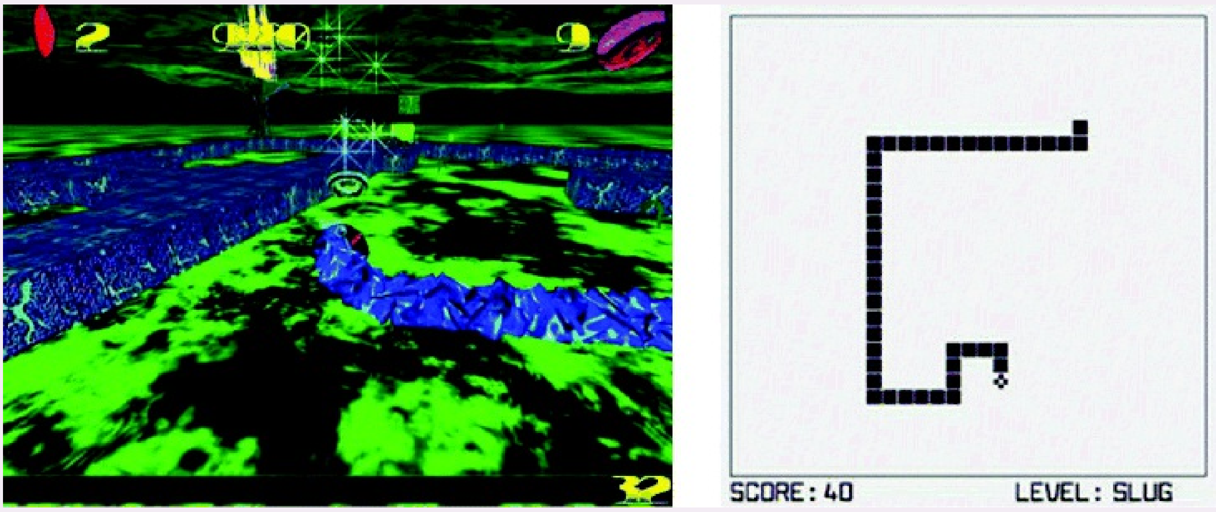


Figure 6.16 Two screenshots from the game Snake – the one on the left is played on a PC and the one on the right on a cell phone. In both games, the goal is to move the snake (the blue thing and the black squares, respectively) towards targets that pop up on the screen (e.g. the bridge, the star) and to avoid obstacles (e.g. a flower, the end of the snake's tail). When a player successfully moves his snake head over or under a target, the snake increases its length by one blob or block. The longer the snake gets, the harder it is to avoid obstacles. If the snake hits an obstacle, the game is over. On the PC version there are lots of extra features that make the game more complicated, including more obstacles and ways of moving. The cell phone version has a simple 2D bird's eye representation, whereas the PC version adopts a 3D third-person avatar perspective

Comment

Show/Hide

6.2.5 Information Visualization and Dashboards

Information visualizations (infoviz) are computer-generated graphics of complex data that are typically interactive and dynamic. The goal is to amplify human cognition, enabling users to see patterns, trends, and anomalies in the visualization and from this to gain insight (Card et al, 1999). Specific objectives are to enhance discovery, decision-making, and explanation of phenomena. Most interactive visualizations have been developed for use by experts to enable them to understand and make sense of vast amounts of dynamically changing domain data or information, e.g. satellite images or research findings, that take much longer to achieve if using only text-based information.

Common techniques that are used for depicting information and data are 3D interactive maps that can be zoomed in and out of and which present data via webs, trees, clusters, scatterplot diagrams, and interconnected nodes (Bederson and Shneiderman, 2003; Chen, 2004). Hierarchical and networked structures, color, labeling, tiling, and stacking are also used to convey different features and their spatial relationships. Figure 6.17 shows a typical treemap, called MillionVis, that depicts one million items all on one screen using the graphical techniques of 2D stacking, tiling, and color (Fekete and Plaisant, 2002). The idea is that viewers can zoom in to parts of the visualization to find out more about certain data points, while also being able to see the overall structure of an entire data set. The treemap (Shneiderman, 1992) has been used to visualize file systems, enabling users to understand why they are running out of disk space, how much space different applications are using, and also for viewing large image repositories that contain terabytes of satellite images. Similar visualizations have been used to represent changes in stocks and shares over time, using rollovers to show additional information, e.g. Marketmap on SmartMoney.com.

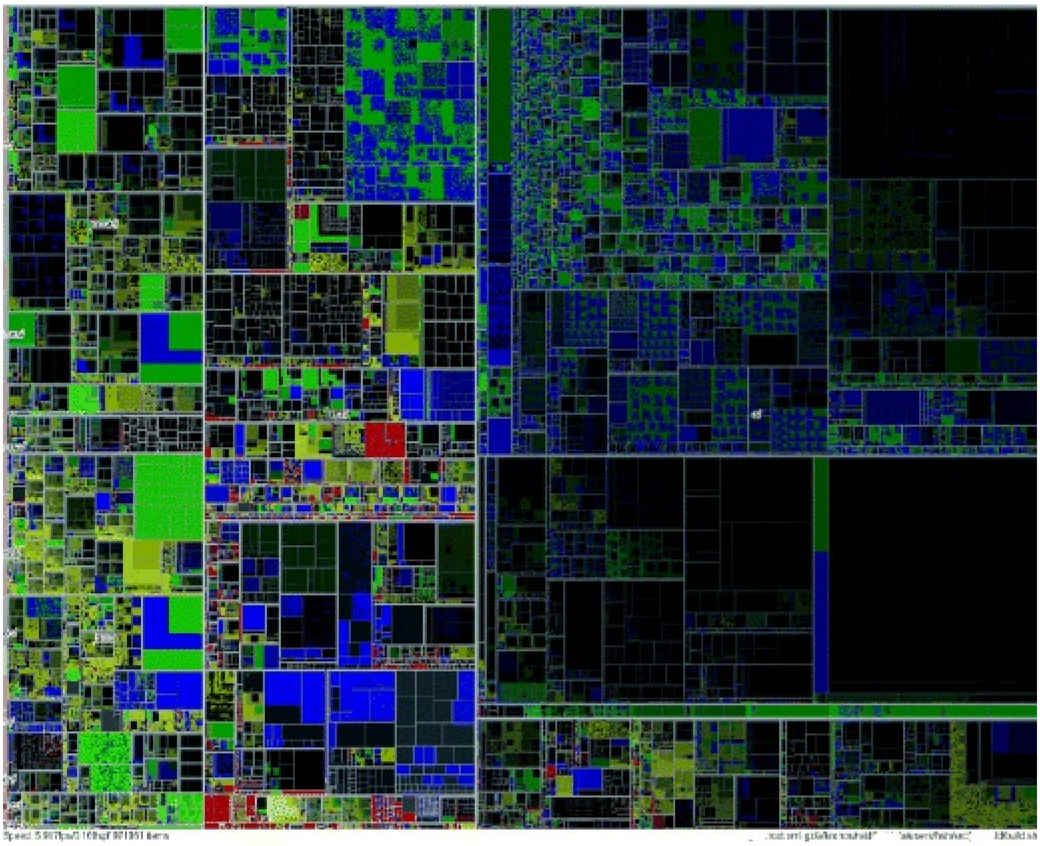


Figure 6.17 An info visualization, using flat colored blocks

Source: Reproduced with permission from Fekete, J.D., Plaisant, C., Interactive Information Visualization of a Million Items, Proc. IEEE Symposium on Information Visualization (2002), 117–124. www.cs.umd.edu/hcil/millionvis.

Dashboards have become an increasingly popular form of visualizing information. They show screenshots of data updated over periods of time, intended to be read at a glance. Unlike other kinds of information visualizations, they tend not to be interactive; the slices of data are intended to depict the current state of a system or process. However, many commercial dashboards have been constructed with poor visual design by software vendors who “focus their marketing efforts on dazzle that subverts the goals of clear communication” (Few, 2013, p. 2). The result can be a hotch-potch of dials, gauges, and graphs, making it difficult to know where to find something or how to compare data over time or across dimensions. Dashboards should be designed to provide digestible and legible information so that users can home in on what is important to them. This requires considering how best to design the spatial layout of a dashboard so that it is intuitive to read when first looking at it. It also needs to be designed in order to direct a user's attention to anomalies or unexpected deviations. This involves working out how best to combine and contrast different elements.

Activity 6.6

(a)

(b)

Figure 6.18 Screenshots from two dashboards: (a) British Airways frequent flier club that shows how much a member has flown since joining them, and (b) London City that provides various information feeds. Which is the easier to read and most informative?

Comment

Show/Hide

Research and Design Issues

Key design issues include whether to use animation and/or interactivity, what form of coding to use, e.g. color or text labels, whether to use a 2D or 3D representational format, what forms of navigation, e.g. zooming or panning, and what kinds of and how much additional information, e.g. rollovers or tables of text, to provide. The type of metaphor to be used is also an important concern, e.g. one based on flying over a geographical terrain or one that represents documents as part of an urban setting. An overriding principle is to design a visualization that is easy to comprehend and easy to make inferences from. If too many variables are depicted in the same visualization, it can make it much more difficult for viewers to read and make sense of what is being represented. ■

6.2.6 Web

Early websites were largely text-based, providing hyperlinks to different places or pages of text. Much of the design effort was concerned with how best to structure information at the interface to enable users to navigate and access it easily and quickly. For example, Nielsen (2000) adapted his and Mohlich's usability guidelines (Nielsen and Mohlich, 1990) to make them applicable to website design, focusing on simplicity, feedback, speed, legibility, and ease of use. He also stressed how critical download time was to the success of a website. Simply, users who have to wait too long for a page to appear are likely to move on somewhere else.

Since the 1990s, many web designers have tried to develop sites that are aesthetically pleasing, usable, and easy to maintain. Graphical design was viewed as a top priority. A goal was to make web pages distinctive, striking, and pleasurable for the user when they first view them and also to make them readily recognizable on their return. Sometimes, they were able to meet all three criteria while at other times they have managed to make a website look good but terrible to navigate and even worse to update content. Other times, they managed to design easy to navigate sites that looked dreadful. Krug (2014) characterized the debate on usability versus attractiveness in terms of the difference between how designers create websites and how users actually view them. He argues that web designers create sites as if the user was going to pore over each page, reading the finely crafted text word for word, looking at the use of images, color, icons, etc., examining how the various items have been organized on the site, and then contemplating their options before they finally select a link. Users, however, often behave quite differently. They will glance at a new page, scan part of it, and click on the first link that catches their interest or looks like it might lead them to what they want. Much of the content on a web page is not read. In his words, web designers are “thinking great literature” (or at least “product brochure”) while the user's reality is much closer to a “billboard going by at 60 miles an hour” (Krug, 2014, p. 21). While somewhat of a caricature of web designers and users, his depiction highlights the discrepancy between the meticulous ways designers create their websites and the rapid and less than systematic approach that users take to look at them.

Website design took off in a big way in the early 2000s when user-centered editing tools (e.g. Dreamweaver) and programming languages (e.g. php, Flash and XML) emerged, providing opportunities for both designers and the general public to create websites to look and behave more like multimedia environments. HTML5 and web development techniques, such as Ajax, enable applications to be built that are largely executed on a user's computer, allowing the development of web applications that mimic desktop apps. Wikis and blogs also became very popular, enabling any number of interlinked web pages to be created and edited easily. WordPress then became very popular, with an easy-to-use interface that provided over 200 free themes (i.e. templates) for users to get started creating their own blog or website. Customized web pages started to be developed for smartphone browsers that listlinked (i.e. provided scrolling lists of articles, games, tunes that could be clicked on) rather than hyperlinked pages.

Swiping became the new form of interaction that very young children learn as naturally as clicking on buttons and turning dials. To the extent sometimes that they think that is how you interact with the real world as the video below suggests!

> Video of ‘A Magazine Is an iPad That Does Not Work’ at http://youtu.be/aXV-yaFmQNk

With the universal uptake of tablet computers in the late 2000s, web traffic shot up and began to overtake that of smartphones. Web designers started to rethink how to design web browsers and websites – not just for PC/laptops and small displays – but for touch-screen tablets. More and more people were looking at online content by scrolling and flicking through it with their fingers rather than using a mouse. The standard desktop interface was found to not work as well on a tablet. In particular, the typical fonts, buttons, and menu tabs were too small and fiddly to select when using a finger. Instead of double-clicking on interface elements – as users do with a mouse or trackpad – tablet screens afford finger tapping. The main ways of navigating are by swiping and pinching. A new style of website emerged that mapped better onto this kind of interaction style but which could also be interacted with easily when using a mouse and trackpad. Responsive websites were developed that change their layout, graphic design, font and appearance depending on the screen size (smartphone, tablet, PC) it was being displayed on.

If you look at the design of many websites, you will see that the front page presents a banner at the top, a short promotional video about the company/product/service, arrows to the left or right to indicate where to flick to move through pages and further details appearing beneath the home page that the user can scroll through. Navigation is largely done through swiping of pages from left to right (and right to left) or scrolling up and down.

Link to some tips on designing websites for tablets at http://css-tricks.com/a-couple-of-best-practices-for-tablet-friendly-design/http://webdesign.tutsplus.com/articles/how-the-ipad-and-tablets-aredriving-new-web-design-trends--webdesign-2428

BOX 6.4

In-Your-Face Web Ads

Web advertising has become pervasive and invasive. Advertisers realized how effective flashing and animated ads were for promoting their products, taking inspiration from the animated neon light adverts used in city centers, such as London's Piccadilly Circus. But since the banner ads emerged in the 1990s, they have become even more cunning and aggressive in their tactics. In addition to designing even flashier banner ads, more intrusive kinds of web ads have begun to appear on our screens. Short movies and garish cartoon animations – often with sound – now pop up in floating windows that zoom into view or are tagged on at the front end of an online newspaper or magazine news videoclip. Moreover, this new breed of in-your-face web ads often requires the user to either wait till it ends or find a check box to close the window down. This can be really annoying, especially when multiple ad windows open up. Sites that provide free services, such as Facebook, YouTube, and Gmail, have also become populated with web ads. Many people choose to ignore them or simply put up with them. However, as advertisers get even more aggressive in their tactics there will be a point where that will become harder to do. The problem is that advertisers pay significant revenues to online companies to have their adverts placed on their websites, entitling them to say where, what, and how they should appear. ■

Research and Design Issues

There are numerous classic books on web design and usability that have been updated as new editions (e.g. Krug (2014); Cooper et al (2014)). In addition, there are many good online sites offering guidelines and tips, together with pointers to examples of bad websites. Key design issues for websites are captured very well by three questions proposed by Keith Instone (quoted in Veen, 2001): Where am I? What's here? Where can I go?

These three fundamental questions are still very relevant for today's websites. Web Content Accessibility Guidelines (WCAG) are available to enable developers to know how to design websites that are inclusive. The latest version, WCAG 2.0, is available at http://www.w3.org/WAI/intro/wcag. These guidelines include designing websites for:

- Users who may not be able to see, hear, or move, or may not be able to process some types of information easily or at all.
- Users who have difficulty reading or comprehending text.
- Users who may not have or be able to use a keyboard or mouse.
- Users who may have a text-only screen, a small screen, or a slow Internet connection.

Website content also needs to be designed for:

- Users who may not speak or understand fluently the language in which the document is written.
- Users who are in a setting where their eyes, ears, or hands are busy or interfered with, e.g. driving to work.
- Users who may have an early version of a browser, a different browser entirely, a voice browser, or a different operating system. ■

Activity 6.7

Look at a fashion brand's website, such as Nike or Levi's, and describe the kind of interface used. How does it contravene the design principles outlined by Veen? Does it matter? What kind of user experience is it providing for? What was your experience of engaging with it?

Comment

Show/Hide

6.2.7 Consumer Electronics and Appliances

Consumer electronics and appliances include machines for everyday use in the home, public place, or car (e.g. washing machines, DVD players, vending machines, remotes, photocopiers, printers, and navigation systems) and personal devices (e.g. MP3 player, digital clock, and digital camera). What they have in common is that most people using them will be trying to get something specific done in a short period of time, such as putting the washing on, watching a program, buying a ticket, changing the time, or taking a snapshot. They are unlikely to be interested in spending time exploring the interface or spending time looking through a manual to see how to use the appliance.

Research and Design Issues

Cooper et al (2014) suggest that appliance interfaces require the designer to view them as transient interfaces, where the interaction is short. All too often, however, designers provide full-screen control panels or an unnecessary array of physical buttons that serve to frustrate and confuse the user and where only a few in a structured way would be much better. Here, the two fundamental design principles of simplicity and visibility are paramount. Status information, such as what the photocopier is doing, what the ticket machine is doing, and how much longer the washing is going to take should be provided in a very simple form and at a prominent place on the interface. A key design question is: as soft displays, e.g. LCD and touch screens, increasingly become part of an appliance interface, what are the trade-offs with replacing the traditional physical controls, e.g. dials, buttons, knobs? ■

Activity 6.8

Look at the controls on your toaster (or the one in Figure 6.19 if you don't have one nearby) and describe what each does. Consider how these might be replaced with an LCD screen. What would be gained and lost from changing the interface in this way?

Comment

Show/Hide

6.2.8 Mobile

Mobile devices have become pervasive, with people increasingly using them in all aspects of their everyday and working lives. They have become business tools to clinch important deals; a remote control for the real world, helping people cope with daily travel delay frustrations; and a relationship appliance to say goodnight to loved ones when away from home (Jones and Marsden, 2006). The Android app, Locket, monitored how many times its 150,000 users checked their phone and found during a six-month period in 2013 that the average person checks their phone 110 times a day. This varies greatly across the day, but they also found it increases considerably in the evening. How does this compare with your usage?

Handheld devices, such as smartphones and iPods, differ from PCs and laptops, in terms of their size, portability, and interaction style. They can be kept in someone's pocket or purse. Early cell phones provided hard-wired small physical keyboards, where letters were pressed. Most smartphones are now touch based, with virtual keyboards that pop up when needed, and are interacted with by finger and thumb tapping. They are increasingly being used by people in mobile settings where they need access to real-time data or information whilst walking around. For example, they are now commonly used in restaurants to take orders, car rentals to check in car returns, supermarkets for checking stock, and on the streets for multiplayer gaming. Tablets are also being used in work settings. For example, many airlines provide their flight attendants with one so they can use their customized flight apps while airborne and at airports; sales and marketing people also use them to demonstrate their goods or collect public opinions. Tablets and smartphones are also being increasingly used in classrooms.

The introduction of Apple's iPhone in 2008 introduced the world to the app – a new user experience that was designed primarily for people to enjoy. There are now over one million apps available with many new ones appearing each day for many different categories, including games, entertainment, social networking, music, productivity, lifestyle, travel, and navigation. Healthy lifestyle and well-being apps (e.g. FitBit, Jawbone Up), which combine a wearable device such as a wristband or headband with a smartphone mobile app (see Chapter 5), are becoming more popular. These can be used on the go and while asleep to monitor and track someone's behaviors and bodily functions. They make use of sensors embedded in the wearable device, such as an accelerometer to detect movement, a thermometer to measure temperature, and galvanic skin response to measure changes in sweat level on someone's skin. Other apps may not be designed for any need, want, or use but purely for idle moments to have some fun. An example of an early highly successful fun app was iBeer (see Figure 6.20), developed by magician Steve Sheraton. Within months of release, hundreds of thousands of people had downloaded the app, then showed their friends who also then downloaded it and showed it to their friends. It became an instant hit, a must have, a party piece – quite unlike any other kind of software. Moreover, a magician created it – rather than an interaction designer – who really understood what captivates people. Part of its success was due to the ingenious use of the accelerometer that is inside the phone. It detects the tilting of the iPhone and uses this information to mimic a glass of beer being drunk. The graphics and sounds are also very enticing; the color of the beer together with frothy bubbles and accompanying sound effects give the illusion of virtual beer being swished around a virtual glass. The beer can be drained if the phone is tilted enough, followed by a belch sound when it has been finished.

Figure 6.20 The iBeer smartphone app

Source: iBeer™ Photo ©2010 HOTTRIX® Reproduced with permission.

Smartphones can also be used to download contextual information by scanning barcodes in the physical world. Consumers can instantly download product information by scanning barcodes using their iPhone when walking around a supermarket, including allergens, such as nuts, gluten, and dairy. For example, the GoodGuide app enables shoppers to scan products in a store by taking a photo of their barcode to see how they rate for healthiness and impact on the environment. Another method that provides quick access to relevant information is the use of QR (quick response) codes that store URLs and look like black and white chequered squares. They can appear in magazines (see Figure 6.21), on billboards, business cards, clothing, food and drink packaging, trains, and so on. They work by people taking a picture using their camera phone which then instantly takes them to a particular website. However, despite their universal appeal to companies as a way of providing additional information or special offers, not many people actually use them in practice. One of the reasons is that they can be slow, fiddly, and cumbersome to use in situ. People have to download a QR reader app first, open it, and then try to hold it over the QR code to take a photo that can take time to open up a webpage (if the WiFi reception is poor).

Figure 6.21 QR code appearing on a magazine page

Research and Design Issues

Mobile interfaces typically have a small screen and limited control space. Designers have to think carefully about what type of dedicated controls (i.e. hard wired) to include, where to place them on the device, and then how to map them onto the software. Applications designed for mobile interfaces need to take into account that navigation will be restricted and text input entry slow, whether using touch, pen, or keypad input. The use of vertical and horizontal scrolling provides a rapid way of scanning though images, menus, and lists. A number of mobile browsers have also been developed that allow users to view and navigate the Internet, magazines, or other media, in a more streamlined way. For example, Edge Browser was one of the first cell phone browser apps to not have an address bar or navigation buttons. The trade-off, however, is it makes it less obvious how to perform the functions that are no longer visible on the screen. A key concern is the hit area. This is an area on the phone display that the user touches to make something happen, such as a key, an icon, a button, or an app. The space needs to be big enough for fat fingers to accurately press. If the space is too small, the user may accidentally press the wrong key, which can be very annoying. The average fingertip is between one and two centimeters wide. Apple, Nokia, and Microsoft each recommend slightly different sizes to accommodate these hit areas to account for the nature of their touch screens (see Chapter 15 for more about how to determine the best size and location of buttons and touch area).

A number of guidelines exist providing advice on how to design interfaces for mobile devices (e.g. Weiss, 2002). Android, Windows Phone, and Apple also provide extensive guidelines for developing smartphone interfaces and apps. Case study 11.1 describes how prototyping can be used for developing mobile interfaces, while case study 11.2 explores the effect of different form factors. ■

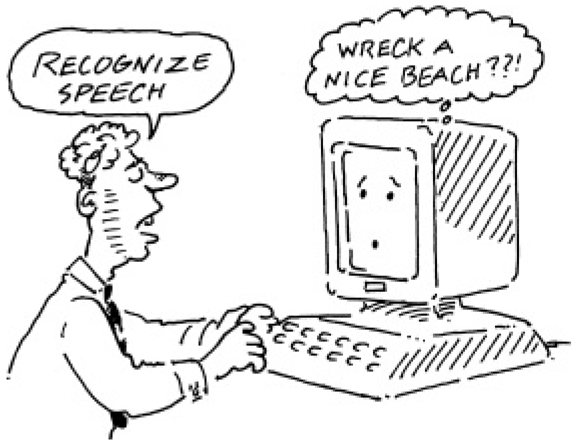
Link to website where Elaine McVicar has provided a number of online easy-to-read and nicely illustrated tutorials as part of the UX Booth for mobile interaction design that include 1) information architecture and 2) interaction techniques, including logging on, page flipping, swiping, pinching, and form design, which can be found at

1. http://tinyurl.com/cmw54vj
2. http://tinyurl.com/c32ns6d

6.2.9 Speech

A speech or voice user interface is where a person talks with a system that has a spoken language application, like a train timetable, a travel planner, or a phone service. It is most commonly used for inquiring about specific information (e.g. flight times) or to perform a transaction (e.g. buy a ticket or top up a smartphone account). It is a specific form of natural language interaction that is based on the interaction type of conversing (see Chapter 2), where users speak and listen to an interface. There are many commercially available speech-based applications that are now being used by corporations, especially for offering their services over the phone. Speech-to-text systems have also become popular, such as Dragon Dictate. Speech technology has also advanced applications that can be used by people with disabilities, including speech recognition word processors, page scanners, web readers, and speech recognition software for operating home control systems, including lights, TV, stereo, and other home appliances.

Technically, speech interfaces have come of age, being much more sophisticated and accurate than the first generation of speech systems in the early 1990s, which earned a reputation for mishearing all too often what a person said (see cartoon). Actors are increasingly used to record the messages and prompts provided that are much friendlier, more convincing, and pleasant than the artificially sounding synthesized speech that was typically used in the early systems.



One of the most popular applications of speech technology is call routing, where companies use an automated speech system to enable users to reach one of their services. Many companies are replacing the frustrating and unwieldy touchtone technology for navigating their services (which was restricted to 10 numbers and the # and * symbols) with the use of caller-led speech. Callers can now state their needs in their own words (rather than pressing a series of arbitrary numbers); for example, 'I'm having problems with my voice mail,' and in response are automatically forwarded to the appropriate service (Cohen et al, 2004).

In human conversations we often interrupt each other, especially if we know what we want, rather than waiting for someone to go through a series of options. For example, at a restaurant we may stop the waitress in mid-flow when describing the specials if we know what we want rather than let her go through the whole list. Similarly, speech technology has been designed with a feature called barge-in that allows callers to interrupt a system message and provide their request or response before the message has finished playing. This can be very useful if the system has numerous options for the caller to choose from and the chooser knows already what he wants.

There are several ways a dialog can be structured. The most common is a directed dialog where the system is in control of the conversation, asking specific questions and requiring specific responses, similar to filling in a form (Cohen et al, 2004):

System: Which city do you want to fly to?

Caller: London

System: Which airport – Gatwick, Heathrow, Luton, Stansted, or City?

Caller: Gatwick

System: What day do you want to depart?

Caller: Monday week

System: Is that Monday 5th May?

Caller: Yes

Other systems are more flexible, allowing the user to take more initiative and specify more information in one sentence (e.g. ‘I'd like to go to Paris next Monday for two weeks'). The problem with this approach is that there is more chance of error, since the caller might assume that the system can follow all of her needs in one go as a real travel agent can (e.g. ‘I'd like to go to Paris next Monday for two weeks and would like the cheapest possible flight, preferably leaving Stansted airport and definitely no stop-overs . . .'). The list is simply too long and would overwhelm the system's parser. Carefully guided prompts can be used to get callers back on track and help them speak appropriately (e.g. ‘Sorry, I did not get all that. Did you say you wanted to fly next Monday?').

A number of speech-based phone apps exist that enable people to use them while mobile, making them more convenient to use than text-based entry. For example, people can speak their queries into their phone using Google Voice or Apple Siri rather than entering text manually. Mobile translators are also coming into their own, allowing people to communicate in real time with others who speak a different language, by letting a software app on their phone do the talking (e.g. Google translate). People speak in their own language using their own phone while the software translates what each person is saying into the language of the other one. Potentially, that means people from all over the world (there are over 6000 languages) can talk to one another without ever having to learn another language.

Research and Design Issues

Key research questions are concerned with how to design systems that can recognize speech and keep the conversation on track. Some researchers focus on making it appear natural (i.e. like human conversations) while others are concerned more with how to help people navigate efficiently through a menu system, by enabling them to recover easily from errors (their own or the system's), be able to escape and go back to the main menu (cf. to the undo button of a GUI), and to guide those who are vague or ambiguous in their requests for information or services using prompts. The type of voice actor, e.g. male, female, neutral, or dialect and form of pronunciation are also topics of research. Do people prefer to listen to and are more patient with a female or male voice? What about one that is jolly or one that is serious?

Cohen et al (2004) discuss the pros and cons of using different techniques for structuring the dialog and managing the flow of voice interactions, the different ways of expressing errors, and the use of conversational etiquette. A number of commercial guidelines are available for voice interfaces and for the visually impaired. ■

6.2.10 Pen

Pen-based devices enable people to write, draw, select, and move objects at an interface using lightpens or styluses that capitalize on the well-honed drawing and writing skills that are developed from childhood. They have been used to interact with tablets and large displays, instead of mouse or keyboard input, for selecting items and supporting freehand sketching. Digital ink, such as Anoto, uses a combination of an ordinary ink pen with a digital camera that digitally records everything written with the pen on special paper. The pen works by recognizing a special non-repeating dot pattern that is printed on the paper. The non-repeating nature of the pattern means that the pen is able to determine which page is being written on, and where on the page the pen is. When writing on the digital paper with a digital pen, infrared light from the pen illuminates the dot pattern, which is then picked up by a tiny sensor. The pen decodes the dot pattern as the pen moves across the paper and stores the data temporarily in the pen. The digital pen can transfer data that has been stored in the pen via Bluetooth or USB port to a PC. Handwritten notes can also be converted and saved as standard typeface text.

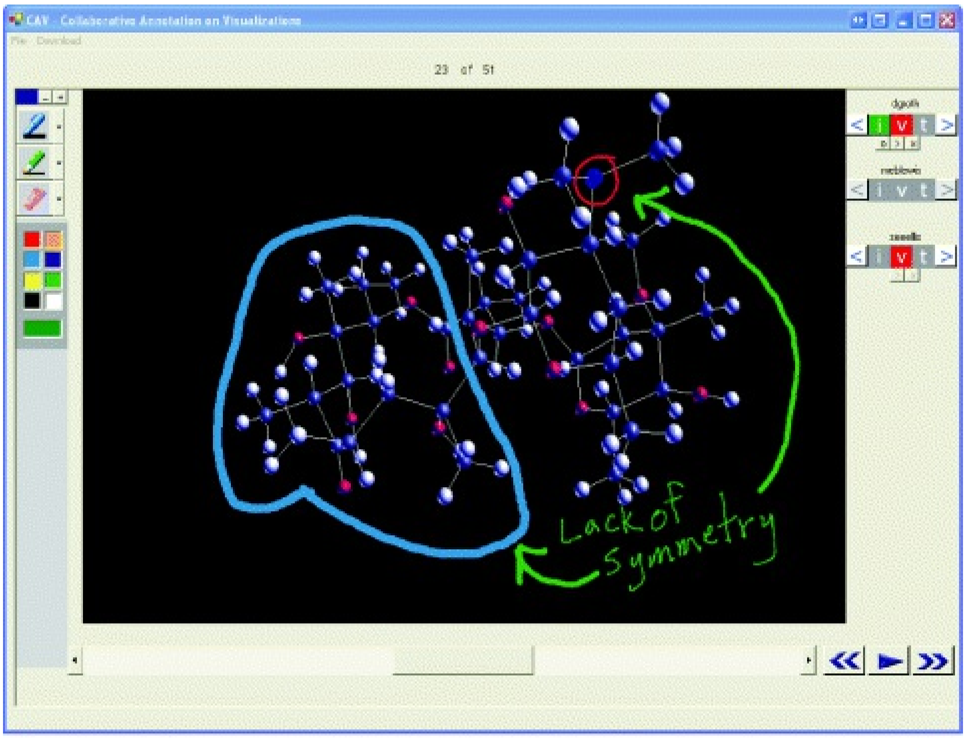


Figure 6.22 Microsoft's digital ink in action showing how it can be used to annotate a scientific diagram

Source: Reproduced by permission of Dennis Groth.

Another advantage of digital pens is that they allow users to quickly and easily annotate existing documents, such as spreadsheets, presentations, and diagrams (see Figure 6.22) – in a similar way to how they would do when using paper-based versions. A number of usability studies have been carried out comparing different ways of entering text using pen input, for children and adults. For example, a study by Read (2005) compared three methods for text input using digital ink technologies; handwriting with a stylus on a Tablet PC, handwriting with a graphics tablet and pen on a standard PC, and handwriting with a digital pen on digital paper. The user group was made up of children aged between 7 and 8, and 12 and 13. The findings showed that the older children were able to use the digital pens best but that both sets of children were able to use the stylus with the Tablet PC without making many errors.

A problem with using pen-based interactions on small screens, such as PDAs, is that sometimes it can be difficult to see options on the screen because a user's hand can occlude part of it when writing.

BOX 6.5

Electronic Ink

Digital ink is not to be confused with the term electronic ink (or e-ink). Electronic ink is a display technology designed to mimic the appearance of ordinary ink on paper used in e-readers. The display used reflects light like ordinary paper. ■

6.2.11 Touch

Touch screens, such as walk-up kiosks (e.g. ticket machines, museum guides), ATMs, and till machines (e.g. restaurants), have been around for some time. They work by detecting the presence and location of a person's touch on the display; options are selected by tapping on the screen. More recently, multitouch surfaces have been developed as the interface for tabletops and smartphones that support a range of more dynamic finger tip actions, such as swiping, flicking, pinching, pushing, and tapping. These have been mapped onto specific kinds of operations, e.g. zooming in and out of maps, moving photos, selecting letters from a virtual keyboard when writing, and scrolling through lists. Two hands can also be used together to stretch and move objects on a tabletop surface, similar to how both hands are used to stretch an elastic band or scoop together a set of objects.

The flexibility of interacting with digital content afforded by finger gestures has resulted in new ways of experiencing digital content. Most notable are the richer ways of reading, scanning, and searching interactive magazines and books on tablets. Wired magazine, for example, was the first to enhance reading through accompanied experiencing of its online version. Similar to the idea behind multimedia, the idea is to enable the reader to readily switch between reading about something (e.g. the history of Mars landings) and experiencing it (e.g. by exploring a virtual simulation of the planet) – only rather than through mouse clicking on hyperlinks, to do it by deft finger movements. A new conceptual model has also been used; content is organized using cards, carousels, and stacks to support rapid finger-flicking navigation, allowing readers to go directly to specific stories while still maintaining a sense of place.

Research and Design Issues

A research question is whether finger-flicking, swiping, stroking, and touching a screen rather than pointing, dragging, and clicking with a mouse will result in new ways of consuming, reading, creating, and searching digital content. On the one hand, it can be much faster to scroll through wheels, carousels, and bars of thumbnail images or lists of options by finger flicking. On the other, it can be more cumbersome, error-prone, and slower to type using a virtual keyboard on a touch display than using a physical keyboard. A novel typing method that has been developed for touch displays is to allow people to swipe their fingers across a virtual keyboard rather than tap at it, such as Swype (see Figure 6.23 and video). Swyping allows the user to move their finger from letter to letter on a virtual keyboard without lifting it. The software senses which are the intended letters by where the user pauses and changes direction. One of the benefits of typing by sliding your fingers across a screen rather than pecking at the keys is that it can make typing faster while also reducing error rate. Another approach, mentioned in Chapter 1, is Minuum's new keyboard that provides a staggered line keyboard for selecting characters. This layout provides a way of fanning out the alphanumeric characters, thereby expanding the hit area. This can be effective for small devices, such as smartwatches, where the screen is relatively small. ■

Figure 6.23 The Swype interface developed for mobile touch displays

Source: Reproduced from http://www.geek.com/articles/mobile/nuances-t9-trace-virtual-keyboard-allows-you-to-swipe-rather-than-type-20100323/technology/.

Video of Swype demo at
http://youtu.be/2xA64e3Txe8

6.2.12 Air-Based Gestures

Camera capture, sensor, and computer vision techniques have advanced such that it is now possible to fairly accurately recognize people's body, arm, and hand gestures in a room. An early commercial application that used gesture interaction was Sony's EyeToy, which used a motion-sensitive camera that sat on top of a TV monitor and plugged into the back of a Sony PlayStation. It could be used to play various video games. The camera filmed the player when standing in front of the TV, projected her image onto the screen, and made her the central character of the video game. The game could be played by anyone, regardless of age or computer experience, simply by moving her legs, arms, head, or any part of the body.

Sony then introduced a motion-sensing wand, called the Move, that uses the Playstation Eye camera to track players' movements using light recognition technology. Nintendo's Wii gaming console also introduced the Wii Remote (Wiimote) controller as a novel input device. It uses accelerometers for gesture recognition. The sensors enable the player to directly input by waving the controller in front of a display, such as the TV. The movements are mapped onto a variety of gaming motions, such as swinging, bowling, hitting, and punching. The player is represented on the screen as an avatar that shows him hitting the ball or swinging the bat against the backdrop of a tennis court, bowling alley, or boxing ring. Like Sony's EyeToy, it was designed to appeal to anyone, from young children to grandparents, and from professional gamers to technophobes, to play games such as tennis or golf, together in their living room. The Wiimote also plays sound and has force feedback, allowing the player to experience rumbles that are meant to enhance the experience when playing the game. The Nunchuk controller can also be used in conjunction with the Wiimote to provide further input control. The analog stick can be held in one hand to move an avatar or characters on the screen while the Wiimote is held in the other to perform a specific action, such as throwing a pass in football.

In late 2010, Microsoft introduced another gesture-based gaming input system for the Xbox: the Kinect (see Figure 6.24). It is more similar to the EyeToy than the Wii in that it does not use a sensor-controller for gesture recognition but camera technology together with a depth sensor and a multiarray microphone (this enables speech commands). An RGB camera sits on the TV, and works by looking for your body; on finding it, it locks onto it, and measures the three-dimensional positioning of the key joints in your body. The feedback provided on the TV screen in response to the various air-gestures has proven to be remarkably effective. Many people readily see themselves as the avatar and learn how to play games in this more physical manner. However, sometimes the gesture/body tracking can misinterpret a player's movements, and make the ball or bat move in the wrong direction. This can be disconcerting, especially for expert gamers.

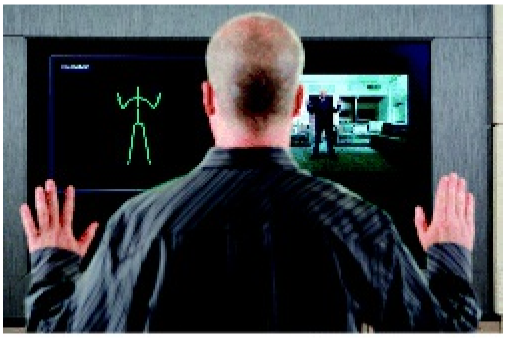


Figure 6.24 Microsoft's Xbox Kinect comprising an RGB camera for facial recognition plus video capturing, a depth sensor (an infrared projector paired with a monochrome camera) for movement tracking, and downward-facing mics for voice recognition

Source: ©PA Images.

A number of air-based gesture systems were developed for controlling home appliances. Early systems used computer vision techniques to detect certain gesture types (e.g. location of hand, movement of arm) that were then converted into system commands. Other systems then began using sensor technologies to detect touch, bend, and speed of movement of the hand and/or arm. Ubi-Finger was developed to allow users to point at an object, e.g. a switch, using his/her index finger and then control it by an appropriate gesture, e.g. pushing the finger down as if flicking on the switch (Tsukada and Yasumura, 2002). Sign language applications have also been built to enable hearing-impaired people to communicate with others without needing a sign language interpreter (Sagawa et al, 1997).

A recent application of air-based gesture interaction is in the operating theater. Surgeons need to keep their hands sterile during operations but also need to be able to look at X-rays and scans during an operation. However, after being scrubbed and gloved, they need to avoid touching any keyboards, phones, and other non-sterile surfaces. A far from ideal workaround is to pull their surgical gown over their hands and manipulate a mouse through the gown. As an alternative, O'Hara et al (2013) have developed a touchless gesture-based system, using Microsoft's Kinect technology, which can recognize a range of gestures that surgeons can use to interact with and manipulate MRI or CT images, including single-handed gestures for moving forward or backward through images, and two-handed gestures for zooming and panning.

Figure 6.25 Touchless gesturing in the operating theater

Source: Courtesy of Kenton O'Hara, Microsoft.

Research and Design Issues

A key design concern for using air-based gestural input is to consider how a computer system recognizes and delineates the user's gestures. In particular, how does it determine the start and end point of a hand or arm movement and how does it know the difference between a deictic gesture (a deliberate pointing movement) and hand waving (an unconscious gesticulation) that is used to emphasize what is being said verbally? Another key design issue is whether holding a control device feels more intuitive for the game or other activity than controller- free gestures. Sometimes it clearly is better to be holding something – when for example, hitting a ball with a bat. Other times it may be better hands-free such as when dancing or doing aerobics. ■

6.2.13 Haptic

Haptic interfaces provide tactile feedback, by applying vibration and forces to the person, using actuators that are embedded in their clothing or a device they are carrying, such as a smartphone or smartwatch. We have already mentioned above how the Wiimote provides rumbles as a form of haptic feedback. Other gaming consoles have also employed vibration to enrich the experience. For example, car steering wheels that are used with driving simulators can vibrate in various ways to provide a feel of the road. As the driver makes a turn, the steering wheel can be programmed to feel like it is resisting – in the way a real steering wheel does.

Vibrotactile feedback can also be used to simulate the sense of touch between remote people who want to communicate. Actuators embedded in clothing can be designed to recreate the sensation of a hug or a stroke through being buzzed on various parts of the body (see Huggy Pajama in Chapter 4). Another use of haptics is to provide feedback to guide people when learning a musical instrument, such as a violin or drums. For example, the MusicJacket (van der Linden et al, 2011) was developed to help novice violin players learn how to hold their instrument correctly and develop good bowing action. Vibrotactile feedback is provided via the jacket to give nudges at key places on the arm and torso to inform the student when either they are holding their violin incorrectly or their bowing trajectory has deviated from a desired path (see Figure 6.26). A user study with novice players showed that players were able to react to the vibrotactile feedback, and adjust their bowing or their posture in response.

Figure 6.26 The MusicJacket prototype with embedded actuators that nudge the player

Research and Design Issues

Haptics are now commonly used in gaming consoles and controllers to heighten the experience. Haptic feedback is also being developed in clothing and other wearables as a way of simulating being touched, stroked, prodded, or buzzed. A promising application area is sensory-motor skills, such as in sports training and learning to play a musical instrument. For example, patterns of vibrations have been placed across snowboarders' bodies to indicate which moves to take whilst snowboarding. A study reported faster reaction times than when the same instructions were given verbally (Spelmezan et al, 2009). A key design question is where best to place the actuators on the body, whether to use a single or a sequence of touches, when to activate, and at what intensity and how often to use them to make the feeling of being touched convincing (e.g. Jones and Sarter, 2008). Providing continuous haptic feedback would be simply too annoying. People would also habituate too quickly to the feedback. Intermittent buzzes can be effective at key moments when a person needs to attend to something but not necessarily tell them what to do. For example, a study by Johnson et al (2010) of a commercially available haptic device, intended to improve posture through giving people a vibrotactile buzz whenever they slouched, found that while the buzzing did not show them how to improve their posture it did improve their body awareness. Different kinds of buzzes can also be used to indicate different tactile experiences that map onto events; for example, a smartphone could transmit feelings of slow tapping to feel like water dropping, which is meant to indicate it is about to rain, and transmit the sensation of heavy tapping to indicate a thunderstorm is looming. ■

6.2.14 Multimodal

Multimodal interfaces are intended to provide enriched and complex user experiences by multiplying the way information is experienced and controlled at the interface through using different modalities, i.e. touch, sight, sound, speech (Bouchet and Nigay, 2004). Interface techniques that have been combined for this purpose include speech and gesture, eye-gaze and gesture, and pen input and speech (Dumas et al, 2009). An assumption is that multimodal interfaces can support more flexible, efficient, and expressive means of human–computer interaction, that are more akin to the multimodal experiences humans experience in the physical world (Oviatt, 2002).

Different input/outputs may be used at the same time, e.g. using voice commands and gestures simultaneously to move through a virtual environment, or alternately using speech commands followed by gesturing. The most common combination of technologies used for multimodal interfaces is speech and vision processing (Deng and Huang, 2004), such as used by Microsoft's Kinect.

Speech-based mobile devices that allow people to interact with information via a combination of speech and touch are beginning to emerge. An example is SpeechWork's multimodal interface developed for one of Ford's SUV concept cars, which allows the occupants to operate on-board systems including entertainment, navigation, cell phone, and climate control by speech.

Research and Design Issues

Multimodal systems rely on recognizing aspects of a user's behavior – be it her handwriting, speech, gestures, eye movements, or other body movements. In many ways, this is much harder to accomplish and calibrate than single modality systems that are programmed to recognize one aspect of a user's behavior. The most researched modes of interaction are speech, gesture, and eye-gaze tracking. A key research question is what is actually gained from combining different input and outputs and whether talking and gesturing as humans do with other humans is a natural way of interacting with a computer (see Chapter 4). Guidelines for multimodal design can be found in Reeves et al (2004). ■

6.2.15 Shareable

Shareable interfaces are designed for more than one person to use. Unlike PCs, laptops, and mobile devices – that are aimed at single users – they typically provide multiple inputs and sometimes allow simultaneous input by collocated groups. These include large wall displays, e.g. SmartBoards (see Figure 6.27a), where people use their own pens or gestures, and interactive tabletops, where small groups can interact with information being displayed on the surface using their fingertips. Examples of interactive tabletops include Microsoft's Surface, Smart's SmartTable, and Circle Twelve's DiamondTouch (Dietz and Leigh, 2001, see Figure 6.27b). The DiamondTouch tabletop is unique in that it can distinguish between different users touching the surface concurrently. An array of antennae is embedded in the touch surface and each one transmits a unique signal. Each user has their own receiver embedded in a mat they stand on or a chair they sit on. When a user touches the tabletop, very small signals are sent through the user's body to their receiver, which identifies which antenna has been touched and sends this to the computer. Multiple users can touch the screen at the same time.

Video of ‘Circle Twelve's’ demonstration of Diamond Touch tabletop at http://youtu.be/S9QRdXITndU

An advantage of shareable interfaces is that they provide a large interactional space that can support flexible group working, enabling groups to create content together at the same time. Compared with a collocated group trying to work around a single-user PC or laptop – where typically one person takes control, making it more difficult for others to take part – large displays have the potential of being interacted with by multiple users, who can point to and touch the information being displayed, while simultaneously viewing the interactions and having the same shared point of reference (Rogers et al, 2009).

Roomware designed a number of integrated interactive furniture pieces, including walls, table, and chairs, that can be networked and positioned together so they can be used in unison to augment and complement existing ways of collaborating (see Figure 6.28). An underlying premise is that the natural way people work together is by congregating around tables, huddling, and chatting besides walls and around tables. The Roomware furniture was designed to augment these kinds of informal collaborative activities, allowing people to engage with digital content that is pervasively embedded at these different locations.

(a)

(b)

Figure 6.27 (a) A SmartBoard in use during a meeting and (b) Mitsubishi's interactive tabletop interface, where collocated users can interact simultaneously with digital content using their fingertips

Source: (a) ©2006 SMART Technologies Inc. Used with permission. (b) Image courtesy of Mitsubishi Electric Research Labs.

Figure 6.28 Roomware furniture

Source: By permission of AMBIENTE.

Research and Design Issues

Early research on shareable interfaces focused largely on interactional issues, such as how to support electronically based handwriting and drawing, and the selecting and moving of objects around the display (Elrod et al, 1992). The PARCTAB system (Schilit et al, 1993) investigated how information could be communicated between palm-sized, A4-sized, and whiteboard-sized displays using shared software tools, such as Tivoli (Rønby-Pedersen et al, 1993). Another concern was how to develop fluid and direct styles of interaction with large displays, both wall-based and tabletop, involving freehand and pen-based gestures (e.g. Shen et al, 2003). Ecologies of devices have been developed where groups can share and create content across multiple devices, such as tabletops and wall displays.

A key research issue is whether shareable surfaces can facilitate new and enhanced forms of collaborative interaction compared with what is possible when groups work together using their own devices, like laptops and PCs (see Chapter 4). One benefit is easier sharing and more equitable participation. For example, tabletops have been designed to support more effective joint browsing, sharing, and manipulation of images during decision-making and design activities (Shen et al, 2002;

Yuill and Rogers, 2012). Core design concerns include whether size, orientation, and shape of the display have an effect on collaboration. User studies have shown that horizontal surfaces compared with vertical ones support more turn-taking and collaborative working in collocated groups (Rogers and Lindley, 2004), while providing larger-sized tabletops does not necessarily improve group working but can encourage more division of labor (Ryall et al, 2004). The need for both personal and shared spaces has been investigated to see how best to enable users to move between working on their own and together as a group. Several researchers have begun to investigate the pros and cons of providing users with complementary devices, such as iPods, digital pens, and other wall displays that are used in conjunction with the shareable surface. Tangible devices (see Section 6.2.16), such as blocks, pucks, and paper models, have also been designed to be used in conjunction with tabletops. An example of this mixed form of interface (described in Chapter 4) is the Reactable, which is an interactive tool for computer music performers. Design guidelines and summaries of empirical research on tabletops and multitouch can be found in Scott et al (2003), O'Hara et al (2003), and Müller-Tomfelde (2010). ■

6.2.16 Tangible

Tangible interfaces use sensor-based interaction, where physical objects, e.g. bricks, balls, and cubes, are coupled with digital representations (Ishii and Ullmer, 1997). When a person manipulates the physical object(s), it is detected by a computer system via the sensing mechanism embedded in the physical object, causing a digital effect to occur, such as a sound, animation, or vibration (Fishkin, 2004). The digital effects can take place in a number of media and places, or they can be embedded in the physical object itself. For example, Zuckerman and Resnick's (2005) Flow Blocks depict changing numbers and lights that are embedded in the blocks, depending on how they are connected together. The flow blocks are designed to simulate real-life dynamic behavior and react when arranged in certain sequences. Another type of tangible interface is where a physical model, e.g. a puck, a piece of clay, or a model, is superimposed on a digital desktop. Moving one of the physical pieces around the tabletop causes digital events to take place on the tabletop. For example, one of the earliest tangible interfaces, called Urp, was built to facilitate urban planning; miniature physical models of buildings could be moved around on the tabletop and used in combination with tokens for wind and shadow-generating tools, causing digital shadows surrounding them to change over time and visualizations of airflow to vary.

The technologies that have been used to create tangibles include RFID tags and sensors embedded in physical objects and digital tabletops that sense the movements of objects and subsequently provide visualizations surrounding the physical objects. Many tangible systems have been built with the aim of encouraging learning, design activities, playfulness, and collaboration. These include planning tools for landscape and urban planning (e.g. Hornecker, 2005; Underkoffler and Ishii, 1998). Another example is Tinkersheets, which combines tangible models of shelving with paper forms for exploring and solving warehouse logistics problems (Zufferey et al, 2009). The underlying simulation allows students to set parameters by placing small magnets on the form.

Tangible computing (Dourish, 2001) has been described as having no single locus of control or interaction. Instead of just one input device such as a mouse, there is a coordinated interplay of different devices and objects. There is also no enforced sequencing of actions and no modal interaction. Moreover, the design of the interface objects exploits their affordances to guide the user in how to interact with them. Tangible interfaces differ from the other approaches insofar as the representations are artifacts in their own right that the user can directly act upon, lift up, rearrange, sort, and manipulate.

What are the benefits of using tangible interfaces compared with other interfaces, like GUI, gesture, or pen-based? One advantage is that physical objects and digital representations can be positioned, combined, and explored in creative ways, enabling dynamic information to be presented in different ways. Physical objects can also be held in both hands and combined and manipulated in ways not possible using other interfaces. This allows for more than one person to explore the interface together and for objects to be placed on top of each other, beside each other, and inside each other; the different configurations encourage different ways of representing and exploring a problem space. In so doing, people are able to see and understand situations differently, which can lead to greater insight, learning, and problem-solving than with other kinds of interfaces (Marshall et al, 2003).

BOX 6.6

VoxBox – a Tangible Questionnaire Machine

Traditional methods for gathering public opinions, such as surveys, involve approaching people in situ but can disrupt the positive experience they are having. VoxBox is a tangible system designed to gather opinions on a range of topics in situ at an event through playful and engaging interaction (Golsteijn et al, 2015). It is intended to encourage wider participation by grouping similar questions, encouraging completion, gathering answers to open and closed questions, and connecting answers and results. It was designed as a large physical system that provides a range of tangible input mechanisms through which people give their opinions, instead of using, for example, text messages or social media input. The various input mechanisms include sliders, buttons, knobs, and spinners – which people are all familiar with. In addition, the system has a transparent tube at the side that drops a ball step by step as sets of questions are completed – to act as an incentive for completion and as a progress indicator. The results of the selections are aggregated and presented as simple digital visualizations on the other side (e.g. 95% are engaged; 5% are bored). VoxBox has been used at a number of events drawing in the crowds, who become completely absorbed in answering questions in this tangible format. ■

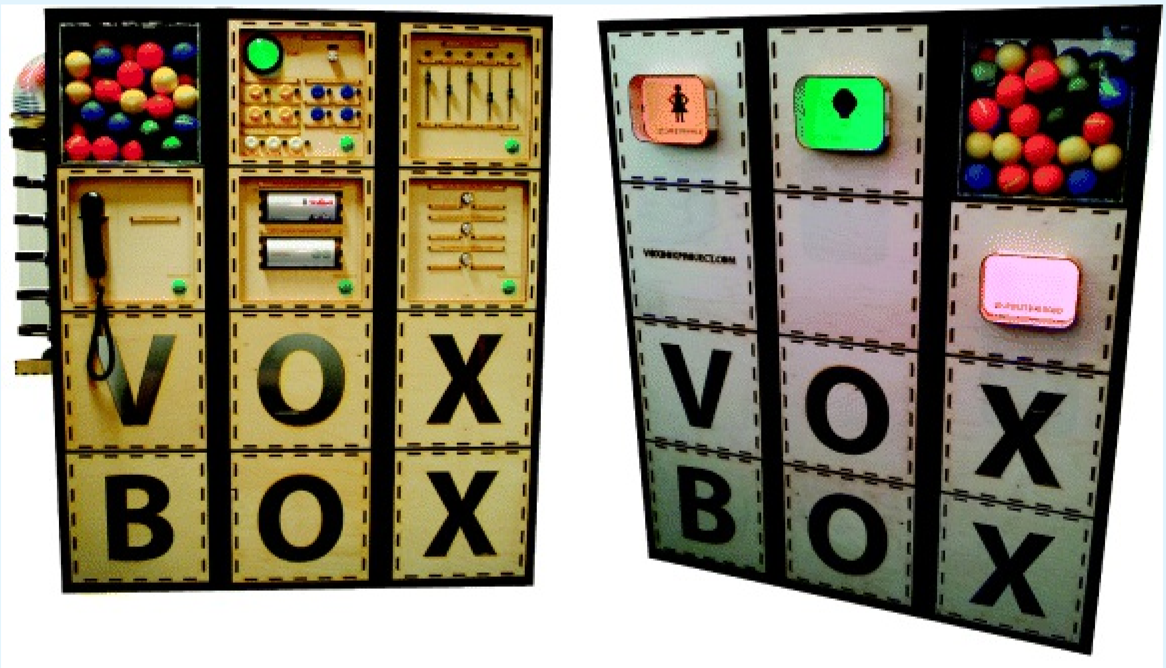


Figure 6.29 VoxBox – Front and back of the tangible machine questionnaire

Source: Golsteijn, C., Gallacher, S., Koeman, L., Wall, L., Andberg, S., Rogers, Y. and Capra, L. (2015) VoxBox: a Tangible Machine that Gathers Opinions from the Public at Events. In Proc. of TEI' 2015. ACM.

Research and Design Issues

Because tangible interfaces are quite different from GUI-based ones, researchers have developed alternative conceptual frameworks that identify their novel and specific features, e.g. Fishkin (2004) and Ullmar et al (2005). A key design concern is what kind of coupling to use between the physical action and effect. This includes determining where the digital feedback is provided in relation to the physical artifact that has been manipulated: for example, should it appear on top of the object, beside it, or some other place. The type and placement of the digital media will depend to a large extent on the purpose of using a tangible interface. If it is to support learning then an explicit mapping between action and effect is critical. In contrast, if it is for entertainment purposes, e.g. playing music or storytelling, then it may be better to design them to be more implicit and unexpected. Another key design question is what kind of physical artifact to use to enable the user to carry out an activity in a natural way. Bricks, cubes, and other component sets are most commonly used because of their flexibility and simplicity, enabling people to hold them in both hands and to construct new structures that can be easily added to or changed. Sticky notes and cardboard tokens can also be used for placing material onto a surface that is transformed or attached to digital content, e.g. Klemmer et al (2001) and Rogers et al (2006). An extensive overview about tangible user interfaces, outlining the important research and design questions, has been written by Shaer and Hornecker (2010). ■

6.2.17 Augmented and Mixed Reality

Other ways that the physical and digital worlds have been bridged include augmented reality, where virtual representations are superimposed on physical devices and objects, and mixed reality, where views of the real world are combined with views of a virtual environment (Drascic and Milgram, 1996). One of the precursors of this work was the Digital Desk (Wellner, 1993). Physical office tools, like books, documents, and paper, were integrated with virtual representations, using projectors and video cameras. Both virtual and real documents were combined.

To begin with, augmented reality was mostly experimented with in medicine, where virtual objects, e.g. X-rays and scans, were overlaid on part of a patient's body to aid the physician's understanding of what was being examined or operated on. It was then used to aid controllers and operators in rapid decision-making. One example is air traffic control, where controllers are provided with dynamic information about the aircraft in their section that is overlaid on a video screen showing the real planes landing, taking off, and taxiing. The additional information enables the controllers to easily identify planes that are difficult to make out – something especially useful in poor weather conditions. Similarly, head-up displays (HUDs) are used in military and civil planes to aid pilots when landing during poor weather conditions. A HUD provides electronic directional markers on a fold-down display that appears directly in the field of view of the pilot. Instructions for building or repairing complex equipment, such as photocopiers and car engines, have also been designed to replace paper-based manuals, where drawings are superimposed upon the machinery itself, telling the mechanic what to do and where to do it.

Everyday graphical representations, e.g. maps, can be overlaid with additional dynamic information. Such augmentations can complement the properties of the printed information in that they enable the user to interact with embedded information in novel ways. An early application is the augmentation of paper-based maps with photographs and video footage to enable emergency workers to assess the effects of flooding and traffic (Reitmayr et al, 2005). A camera mounted above the map tracks the map's locations on the surface while a projector augments the maps with projected information from overhead. Figure 6.30 shows areas of flooding that have been superimposed on a map of Cambridge (UK), together with images of the city center captured by cameras.

Figure 6.30 **An augmented map showing the flooded areas at high water level overlaid on the paper map. The handheld device is used to interact with entities referenced on the map**

Source: Reproduced with permission.

There are many augmented reality apps available now for a range of contexts, from education to car navigation, where digital content is overlaid on geographic locations and objects. To reveal the digital information, users open the AR app on a smartphone or tablet and the content appears superimposed on what is viewed through the screen. An example is of Top Gear presenter, James May, appearing as a 3D character (see Figure 6.31) to act as personal tour guide at the Science Museum in London. Other AR apps have been developed to aid people walking in a city or town. Directions (in the form of a pointing hand or arrow) and local information (e.g. the nearest McDonald's) are overlaid on a picture of the street the person holding the phone is walking in. Real-estate apps have also been developed that combine an image of a residential property with its price per square meter. The directions and information change as the person walks or drives up the street.

Figure 6.31 James May appearing in 3D Augmented Reality

Source: http://www.wired.com/2012/04/top-gear-host-narrates-museum-exhibits-as-augmented-reality-avatar/. Roberto Baldwin/Wired/©Conde Nast

Link to app of James May, appearing as a 3D character to act as personal tour guide at the Science Museum in London, can be seen at http://www.sciencemuseum.org.uk/visitmuseum_old/jamesmay.aspx

Research and Design Issues

A key research concern when designing mixed reality environments and augmented reality is what form the digital augmentation should take and when and where it should appear in the physical environment (Rogers et al, 2005). The information needs to stand out but not distract the person from his ongoing activity in the physical world. For example, ambient sounds need to be designed to be distinct from naturally occurring sounds so that they draw a person's attention without distracting him and then allow him to return to what he was doing. Information that is superimposed on the physical world, e.g. digital information overlaying video footage of a runway to identify vehicles and planes, needs to be simple and easy to align with the real-world objects.

It is important to understand how designing for playful learning experiences is very different from designing for military or medical applications. Ambiguity and uncertainty may be exploited to good effect in mixed reality games but could be disastrous in the latter categories. The type of technology will also determine what guidance will be of relevance. A guideline for the use of an optical see-through display, e.g. shutter glasses or head-mounted display, may not be relevant for a video see-through display. Likewise, a guideline for a mobile augmented reality solution may not be relevant for a fixed display application. Published design guidelines include Cawood and Fiala (2008) and Wetzel et al (2008). ■

6.2.18 Wearables

Imagine being at a party and being able to access the Facebook of a person whom you have just met, while or after talking to her, to find out more about her. The possibility of having instant information before one's very own eyes that is contextually relevant to an ongoing activity and that can be viewed surreptitiously (i.e. without having to physically pull out a smartphone) is very appealing. Since the early experimental days of wearable computing, where Steve Mann (1997) donned head and eye cameras to enable him to record what he saw while also accessing digital information on the move, there have been many innovations and inventions including the latest Google Glass.

Dilemma

Google Glass: Seeing too much?

Google Glass is a wearable that went on sale in 2014 in various fashion styles (see Figure 6.32). It was designed to look like a pair of glasses, but with one lens of the glass being an interactive display with an embedded camera that can be controlled with speech input. It allows the wearer to take photos and video on the move and look at digital content, such as emails, texts, and maps. The wearer can also search the web using voice commands and the results come back on the screen. A number of applications have been developed besides everyday use, including WatchMeTalk that provides live captions that help the hearing-impaired in their day-to-day conversations and Preview for Glass that enables a wearer to watch a movie trailer the moment they look at a movie poster.

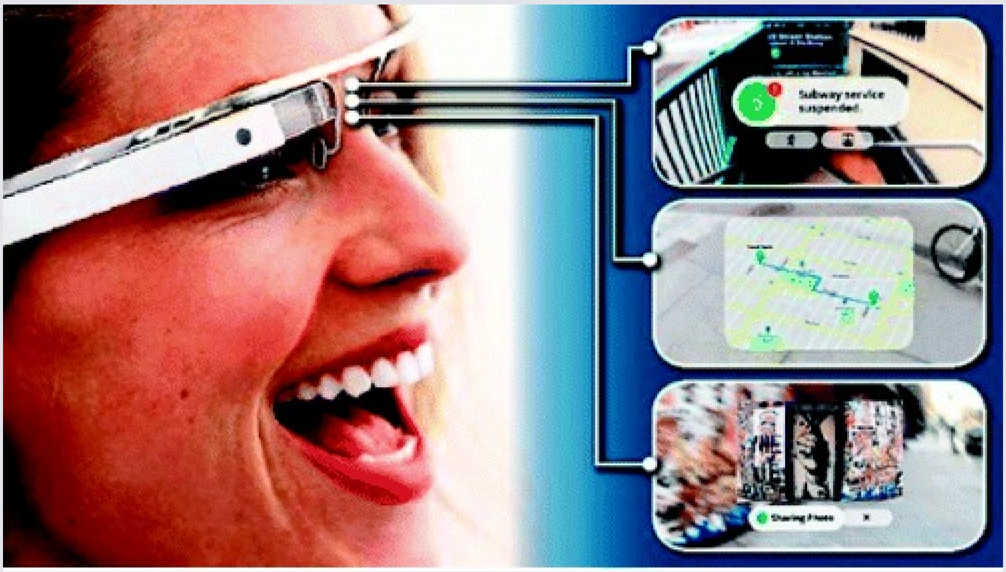


Figure 6.32 Google Glass

Source: https://www.google.co.uk/intl/en/glass/start/. Google and the Google logo are registered trademarks of Google Inc., used with permission.

Video of 'London through Google Glass' at http://youtu.be/Z3AIdnzZUsE

However, it can be slightly unnerving when in the company of someone wearing Google Glass as they look up and to the right to view what is on the glass screen rather than at you and into your eyes. As a result, you might see more of the whites of their eyes than the usual interested dilated pupils. Could this be the end of eye contact as we know it? One of the criticisms of early wearers of Google Glass was that it made them appear to be staring into the distance.

Others are worried that those wearing Google Glass are recording everything that is happening in front of them. As a reaction, a number of bars and restaurants in San Francisco and other cities have implemented a ‘no Glass’ policy to prevent customers from recording other patrons. There has also been much debate in the press about the latest developments in facial recognition. There are apps developed for Google Glass that take a picture of the person you are talking with and then check their online profile, providing a cloud of personal information about them, presumably mined from Facebook and other social media apps. So you can find out more about someone on the go while talking to them – for example, what music they like, what films they have just seen, where they have just been on vacation, and so on – all in a digestible précis surrounded by a halo of photos. One could imagine that if this way of meeting up with others actually takes off, we might find ourselves in the situation where we won't need to talk to each other anymore. Just as text messaging has largely taken over from making phone calls for many people, ‘cloud talk’ could start taking over our initial encounters with people when we meet them at parties, at conferences, on trains, etc. We might nod and smile in acknowledgement of each other but we won't ever have to have those awkward conversations anymore, such as about where you come from or what you do for work. A panacea for the shy? But how will we know what each other is looking at? You might think I am reading your blog or tweets when in your presence, but really I might just be watching the latest updates of the football results and pretending to ‘meet you’. ■

New flexible display technologies, e-textiles, and physical computing (e.g. Arduino) provide opportunities for thinking about how to embed such technologies on people in the clothes they wear. Jewelry, head-mounted caps, glasses, shoes, and jackets have all been experimented with to provide the user with a means of interacting with digital information while on the move in the physical world. An early motivation was to enable people to carry out tasks (e.g. selecting music) while moving without having to take out and control a handheld device. Examples include a ski jacket with integrated MP3 player controls that enable wearers to simply touch a button on their arm with their glove to change a track and automatic diaries that keep users up-to-date on what is happening and what they need to do throughout the day. More recent applications have focused on embedding various textile, display, and haptic technologies to promote new forms of communication and have been motivated by aesthetics and playfulness. For example, CuteCircuit develops fashion clothing, such as the KineticDress, which is embedded with sensors that follow the body of the wearer to capture their movements and interaction with others. These are then displayed through electroluminescent embroidery that covers the external skirt section of the dress. Depending on the amount and speed of the wearer's movement it will change pattern, displaying the wearer's mood to the audience and creating a magic halo around her. CuteCircuit also developed the Hug Shirt (see Chapter 4).

Video of the ‘Talking Shoe’ concept at http://youtu.be/VcaSwxbRkcE

Research and Design Issues

A core design concern – that is specific to wearable interfaces – is comfort. Users need to feel comfortable wearing clothing that is embedded with technology. It needs to be light, small, not get in the way, fashionable, and (with the exception of the displays) preferably hidden in the clothing. Another related issue is hygiene – is it possible to wash or clean the clothing once worn? How easy is it to remove the electronic gadgetry and replace it? Where are the batteries going to be placed and how long is their lifetime? A key usability concern is how does the user control the devices that are embedded in his clothing – is touch, speech, or more conventional buttons and dials preferable? ■

Activity 6.9

Smartwatches, such those made by Android, Apple, Pebble, and Samsung, have become popular wearables, providing a multitude of functions including fitness tracking and beaming out messages, Facebook updates, and the latest tweets. Samsung's even has a fingerprint scanner to enable payments to be made simply by touching the watch. Smartwatches are also context and location aware. On detecting the wearer's presence, promotional offers may be pinged to a person wearing a smartwatch from nearby stores, tempting them in to buy. How do you feel about this?

Comment

Show/Hide

6.2.19 Robots and Drones

Robots have been with us for some time, most notably as characters in science fiction movies, but also playing an important role as part of manufacturing assembly lines, as remote investigators of hazardous locations (e.g. nuclear power stations and bomb disposal), and as search and rescue helpers in disasters (e.g. fires) or far-away places (e.g. Mars). Console interfaces have been developed to enable humans to control and navigate robots in remote terrains, using a combination of joysticks and keyboard controls together with camera and sensor-based interactions (Baker et al, 2004). The focus has been on designing interfaces that enable users to effectively steer and move a remote robot with the aid of live video and dynamic maps.

Domestic robots that help with the cleaning and gardening have become popular. Robots are also being developed to help the elderly and disabled with certain activities, such as picking up objects and cooking meals. Pet robots, in the guise of human companions, are being commercialized. A somewhat controversial idea is that sociable robots should be able to collaborate with humans and socialize with them – as if they were our peers (Breazeal, 2005).

Several research teams have taken the ‘cute and cuddly’ approach to designing robots, signaling to humans that the robots are more pet-like than human-like. For example, Mitsubishi has developed Mel the penguin (Sidner and Lee, 2005) whose role is to host events, while the Japanese inventor Takanori Shibata developed Paro in 2004, a baby harp seal that looks like a cute furry cartoon animal, and whose role was as a companion (see Figure 6.33). Sensors have been embedded in the pet robots, enabling them to detect certain human behaviors and respond accordingly. For example, they can open, close, and move their eyes, giggle, and raise their flippers. The robots afford cuddling and talking to – as if they were pets or animals. The appeal of pet robots is thought to be partially due to their therapeutic qualities, being able to reduce stress and loneliness among the elderly and infirm (see Chapter 5 for more on cuddly robot pets). Paro has since been used in the UK to help patients with dementia to make them feel more at ease and comforted (Griffiths, 2014). Specifically, it has been used to encourage social behavior amongst patients who often anthropomorphize it. For example, they might say as a joke “it's farted on me!”, which makes them and others around them laugh, leading to further laughter and joking. This form of encouraging of social interaction is thought to be therapeutic.

Figure 6.33 Left: Mel, the penguin robot, designed to host activities; right: Japan's Paro, an interactive seal, designed as a companion, primarily for the elderly and sick children

Source: (left) Image courtesy of Mitsubishi Electric Research Labs. (right) Courtesy of Parorobots.com.

Video of 'Robot Pets of the Future' at http://youtu.be/wBFws1lhuv0

Drones are a form of unmanned aircraft that are controlled remotely. They were first used by hobbyists and then by the military. Since, they have become more affordable, accessible, and easier to fly, and as a result have begun to be used in a wider range of contexts. These include entertainment, such as carrying drinks and food to people at festivals and parties; agricultural applications, such as flying them over vineyards and fields to collect data that is useful to farmers; and helping to track poachers in wildlife parks in Africa. Compared with other forms of data collecting, they can fly low and stream photos to a ground station, where the images can be stitched together into maps and then used to determine the health of a crop or when it is the best time to harvest the crop.

> Video of OppiKoppi, a drone that drops beer to festival goers at http://youtu.be/janur7RJwm0

Figure 6.34 A drone being used to survey the state of a vineyard

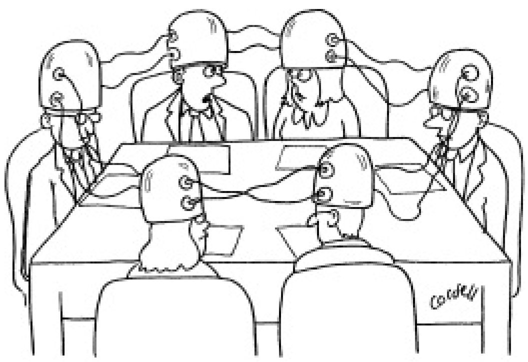
Source: Courtesy of Discover Sonoma County Wine http://www.latimes.com/business/la-fi-drones-agriculture-20140913-story.html#page=1.

Research and Design Issues

An ethical concern is whether it is acceptable to create robots that exhibit behaviors that humans will consider to be human- or animal-like. While this form of attribution also occurs for PC-based agent interfaces (see Chapter 2), having a physical embodiment – as robots do – can make people suspend their disbelief even more, viewing the robots as pets or humans. This raises the moral question as to whether such anthropomorphism should be encouraged. Should robots be designed to be as human-like as possible, looking like us with human features, e.g. eyes and mouth, behaving like us, communicating like us, and emotionally responding like us? Or should they be designed to look like robots and behave like robots, e.g. vacuum cleaner robots that serve a clearly defined purpose? Likewise, should the interaction be designed to enable people to interact with the robot as if it were another human being, e.g. talking, gesturing, holding its hand, and smiling at it, or should the interaction be designed to be more like human–computer interaction, e.g. pressing buttons, knobs, and dials to issue commands?

For many people, the cute pet approach to robotic interfaces seems preferable to one that aims to design them to be more like fully fledged human beings. Humans know where they stand with pets and are less likely to be unnerved by them and, paradoxically, are more likely to suspend their disbelief in the companionship they provide.

Another ethical concern is whether it is acceptable to use unmanned drones to take a series of images or videos of fields, towns, and private property without permission or people knowing what is happening. ■



"Frankly, I'm not sure this whole idea-sharing thing is working."

6.2.20 Brain–Computer Interfaces

Brain–computer interfaces (BCI) provide a communication pathway between a person's brain waves and an external device, such as a cursor on a screen or a tangible puck that moves via airflow). The person is trained to concentrate on the task (e.g. moving the cursor or the puck). Several research projects have investigated how this technique can be used to assist and augment human cognitive or sensory-motor functions. The way BCIs work is through detecting changes in the neural functioning in the brain. Our brains are filled with neurons that comprise individual nerve cells connected to one another by dendrites and axons. Every time we think, move, feel, or remember something, these neurons become active. Small electric signals rapidly move from neuron to neuron – that can to a certain extent be detected by electrodes that are placed on a person's scalp. The electrodes are embedded in specialized headsets, hairnets, or caps (see Figure 6.35). Tan Le, in her 2010 TED talk, demonstrated how it is possible, using the Emotiv Systems headset, for a participant to move virtual objects, such as a cube, on a screen.

Figure 6.35 The Brainball game using a brain–computer interface

Source: "Brainball" from The Interactive Institute. Reproduced with permission.

Video demonstrating brain–computer interaction at www.ted.com/talks/tan_le_a_headset_that_reads_your_brainwaves.html

Brain–computer interfaces have also been developed to control various games. For example, Brainball was developed as a game to be controlled by players' brain waves in which they compete to control a ball's movement across a table by becoming more relaxed and focused.

Other possibilities include controlling a robot and being able to fly a virtual plane by thinking of lifting the mind.

Pioneering medical research conducted by the BrainGate research group at Brown University has started using brain–computer interfaces to enable people who are paralyzed to control robots. For example, a robotic arm controlled by a tethered BCI has enabled patients who are paralyzed to feed themselves (see video).

> Video of a woman who is paralyzed moving a robot with her mind at http://youtu.be/ogBX18maUiM

6.3 Natural User Interfaces and Beyond

As we have seen, there are many kinds of interface that can be used to design for user experiences. The staple for many years was the GUI (graphical user interface), which without doubt has been very versatile in supporting all manner of computer-based activities, from sending email to managing process control plants. But is its time up? Will NUIs (short for natural user interfaces) begin to overtake them?

But what exactly are NUIs? A NUI is one that enables people to interact with a computer in the same ways they interact with the physical world, through using their voice, hands, and bodies. Instead of using a keyboard and a mouse (as is the case with GUIs), a natural user interface allows users to speak to machines, stroke their surfaces, gesture at them in the air, dance on mats that detect feet movements, smile at them to get a reaction, and so on. The naturalness refers to the way they exploit the everyday skills we have learned, such as talking, writing, gesturing, walking, and picking up objects. In theory, they should be easier to learn and map more readily onto how people interact with the world than compared with learning to use a GUI. For example, as Steve Ballmer, a former CEO of Microsoft, noted when the idea of NUIs first came to the fore:

> **I believe we will look back on 2010 as the year we expanded beyond the mouse and keyboard and started incorporating more natural forms of interaction such as touch, speech, gestures, handwriting, and vision – what computer scientists call the 'NUI' or natural user interface. (Ballmer, 2010)**

Instead of having to remember which function keys to press to open a file, a NUI means a person only has to raise their arm or say 'open'. But how natural are NUIs? Is it more natural to say 'open' than to flick a switch when wanting to open a door? And is it more natural to raise both arms to change a channel on the TV than to press a button on the remote? Whether a NUI is more natural than a GUI will depend on a number of factors, including how much learning is required, the complexity of the application/device's interface, and whether accuracy and speed are needed (Norman, 2010). Sometimes a gesture is worth a thousand words. Other times, a word is worth a thousand gestures. It depends on how many functions the system supports.

Consider the sensor-based faucets that were described in Chapter 1. The gesture-based interface works mostly (with the exception of people wearing black clothing that cannot be detected) because there are only two functions: (i) turning on by waving one's hands under the tap, and (ii) turning off by removing them from the sink. Now think about other functions that faucets usually provide, such as controlling water temperature and flow. What kind of a gesture would be most appropriate for changing the temperature and then the flow? Would one decide on the temperature first by raising one's left hand and the flow by raising one's right hand? How would we know when to stop raising our hand to get the right temperature? We would need to put a hand under the tap to check. If we put our right hand under that might have the effect of decreasing the flow. And when does the system know that the desired temperature and flow has been reached? Would it require having both hands suspended in mid-air for a few seconds to register that was the desired state? We would all need to become water conductors. It is hardly surprising that such a system of control does not exist – since it simply would not work. Hence, the reason why sensor-based faucets in public toilets all have their temperature and flow set to a default.

This caricature illustrates how it can be more difficult to design even a small set of gestures to map onto a set of control functions, which can be accurately recognized by the system while also readily learned and remembered by the general public. It also highlights how gestural, speech, and other kinds of NUIs will not replace GUIs as the new face of interaction design. However, it does not mean they will not be useful. They are proving to be effective and enjoyable to use when controlling and manipulating digital content in a number of tasks and activities. For example, using gestures and whole body movement has proven to be highly enjoyable as a form of input for many computer games and physical exercises, such as those that have been developed for the Wii and Kinect systems. Furthermore, new kinds of gesture, speech, and touch interfaces have proven to be very empowering for people who are visually impaired and who have previously had to use specialized tools to interface with GUIs. For example, the iPhone's VoiceOver control features enable visually impaired people to send email, use the web, play music, and so on, without having to buy an expensive customized phone or screen reader. Moreover, being able to purchase a regular phone means not being singled out for special treatment. And while some gestures may feel cumbersome for sighted people to learn and use, they may not be for blind or visually impaired people. The VoiceOver press and guess feature that reads out what you tap on the screen (e.g. ‘messages,’ ‘calendar,’ ‘mail: 5 new items') can open up new ways of exploring an application while a three-finger tap can become a natural way to turn the screen off.

An emerging class of human–computer interfaces are those that rely largely on subtle, gradual, continuous changes triggered by information obtained implicitly from the user. They are connected with lightweight, ambient, context aware, affective, and augmented cognition interfaces and are especially found in high-performance tasks such as gaming apps (Solovey et al, 2014). Using brain, body, behavioral, and environmental sensors, it is now possible to capture subtle changes in people's cognitive and emotional states in real time. This opens up new doors in human–computer interaction. In particular, it allows for information to be used as both continuous and discrete input, potentially enabling new outputs to match and be updated with what people might want and need at any given time. However, brain, body, and other sensor data are different from GUIs. Future research needs to consider how best to exploit this more subtle class of input in order to achieve new interfaces.

6.4 Which Interface?

In this chapter we have given an overview of the diversity of interfaces that is now available or currently being researched. There are many opportunities to design for user experiences that are a far cry from those originally developed using command-based interfaces in the 1980s. An obvious question this raises is: but which one and how do you design it? In many contexts, the requirements for the user experience that have been identified during the design process will determine what kind of interface might be appropriate and what features to include. For example, if a healthcare application is being developed to enable patients to monitor their dietary intake, then a mobile device – that has the ability to scan barcodes and/or take pictures of food items that can be compared with a database – would appear to be a good interface to use, enabling mobility, effective object recognition, and ease of use. If the goal is to design a work environment to support collocated group decision-making activities then combining shareable technologies and personal devices that enable people to move fluidly between them would be a good choice.

But how do we decide which interface is preferable for a given task or activity? For example, is multimedia better than tangible interfaces for learning? Is speech effective as a command-based interface? Is a multimodal interface more effective than a single media interface? Are wearable interfaces better than mobile interfaces for helping people find information in foreign cities? Are virtual environments the ultimate interface for playing games? Or will mixed reality or tangible environments prove to be more challenging and captivating? Will shareable interfaces, such as interactive furniture, be better at supporting communication and collaboration compared with using networked desktop technologies? And so forth. These questions are currently being researched. In practice, which interface is most appropriate, most useful, most efficient, most engaging, most supportive, etc., will depend on the interplay of a number of factors, including reliability, social acceptability, privacy, ethical, and location concerns.

Assignment

In Activity 6.4 we asked you to compare the experience of playing the game of Snake on a PC with playing on a cell/smart phone. For this assignment, we want you to consider the pros and cons of playing the same game using different interfaces. Select three interfaces, other than the GUI and mobile ones (e.g. tangible, wearable, and shareable) and describe how the game could be redesigned for each of these, taking into account the user group being targeted. For example, the tangible game could be designed for young children, the wearable interface for young adults, and the shareable interface for elderly people.

a. Go through the research and design issues for each interface and consider whether they are relevant for the game setting and what issues they raise. For the wearable interface, issues to do with comfort and hygiene are important when designing the game.

b. Describe a hypothetical scenario of how the game would be played for each of the three interfaces.

c. Consider specific design issues that will need to be addressed. For example, for the shareable surface would it be best to have a tabletop or a wall-based surface? How will the users interact with the snake for each of the different interfaces; by using a pen, fingertips, or other input device? Is it best to have a representation of a snake for each player or one they take turns to play with? If multiple snakes are used, what will happen if one person tries to move another person's snake? Would you add any other rules? And so on.

d. Compare the pros and cons of designing the Snake game using the three different interfaces with respect to how it is played on the cell phone and the PC.

Take a Quickvote on Chapter 6:
www.id-book.com/quickvotes/chapter6

Summary

This chapter has given an overview of the diversity of interfaces that can be designed for user experiences, identifying key design issues and research questions that need to be addressed. It has highlighted the opportunities and challenges that lie ahead for designers and researchers who are experimenting with and developing innovative interfaces. It has also explicated some of the assumptions behind the benefits of different interfaces – some that are supported, others that are still unsubstantiated. It has presented a number of interaction techniques that are particularly suited (or not) for a given interface type. It has also discussed the dilemmas facing designers when using a particular kind of interface, e.g. abstract versus realism, menu selection versus free-form text input, human-like versus non-human-like. Finally, it has presented pointers to specific design guidelines and exemplary systems that have been designed using a given interface.

Key points

- Many interfaces have emerged post the WIMP/GUI era, including speech, wearable, mobile, tangible, brain–computer, robots, and drones.
- A range of design and research questions need to be considered when deciding which interface to use and what features to include.
- So-called natural user interfaces may not be as natural as graphical user interfaces – it depends on the task, user, and context.
- An important concern that underlies the design of any kind of interface is how information is represented to the user (be it speech, multimedia, virtual reality, augmented reality), so that they can make sense of it with respect to their ongoing activity, e.g. playing a game, shopping online, or interacting with a pet robot.
- Increasingly, new interfaces that are context-aware or monitor people raise ethical issues concerned with what data is being collected and what it is used for.

Further Reading

Many of the best books on designing interfaces have been developed for the practitioner market. They are often written in a humorous and highly accessible way, replete with cartoons, worldly prescriptions, and figures. They also use modern fonts that make the text very appealing. We recommend:

GOOGLE (2014) Material Design http://www.google.com/design/spec/material-design/introduction.html. This online resource provides a living online document that visually illustrates essential interface design principles. It is beautifully laid out and very informative to click through all the interactive examples it provides. It shows how to add some physical properties to the digital world to make it feel more intuitive to use across platforms.

JOHNSON, J. (2007) GUI Bloopers. 2.0: Common user interface design don'ts and dos, (2nd edn). Morgan Kaufmann. This second edition of a classic has been updated to reflect the bloopers that are common across the design of a range of interfaces. It is full of the author's amusing anecdotes and other designer howlers.

There are also many good practical guides on web usability and interaction design that have been published by New Riders. Some are updated on a regular basis while others are new. These include:

KRUG, S. (2014) Don't Make Me Think! (3rd edn). New Riders Press.

NIELSEN, J. and LORANGER, H. (2006) Prioritizing Web Usability. New Riders Press.

VEEN, J. (2001) The Art and Science of Web Design. New Riders Press.

And finally, a thought-provoking essay that everyone should read (a shorter version is also available on Don Norman's website):

NORMAN, D. (2010) Natural interfaces are not natural, interactions, May/June, 6–10.

Interview with Leah Beuchley

Leah Buechley is an independent designer, engineer, and educator. She has a PhD in Computer Science and a degree in physics. She began her studies as a dance major and has also been deeply engaged in theater, art, and design over the years. She was the founder and director of the high-low tech group at the MIT media lab from 2009 to 2014. She has always blended the sciences and the arts in her education and her career – as witnessed by her current work, comprising computer science, industrial design, interaction design, art, and electrical engineering.

Why did you call your MIT media lab research group high-low tech?

Technology is made from a limited palette of physical materials, designed and built by a small subset of people, and interacted with in a very constrained manner. The name high-low tech is meant to evoke an alternate vision of technology – technology that is handcrafted by different people to fit their own personal needs. More specifically, I was interested in expanding the technology space to encompass a broader palette of materials (including materials like fabrics, ceramics, paper, and wood), a more diverse group of designers and engineers, and an expanded conception of interface.

Can you give me some examples of how you mesh the digital with physical materials?

I've been working on a project called LilyPad Arduino (or LilyPad) for almost 10 years. LilyPad is a construction kit that enables people to embed computers and electronics into their clothes. It's a set of sewable electronic pieces – including microcontrollers, sensors, and LEDs – that are stitched together with conductive thread. People can use the kit to make singing pillows, glow in the dark handbags, and interactive ball gowns. I recently co-authored a book with my former student Kanjun Qiu, Sew Electric, that introduces electronics and programming via LilyPad.

Another example is the work my former students and I have done in paper-based computing. My former student Jie Qi just developed a kit called ‘circuit stickers’ that lets you build interactive paper-based projects. Based on her years of research in high-low tech, the kit is a set of flexible peel-and-stick electronic stickers. You can connect ultra-thin LEDs, microcontrollers, and sensors with conductive ink, tape, or thread to quickly make beautiful electronic sketches.

Why would anyone want to wear a computer in their clothing?

Computers open up new creative possibilities for designers. Computers are simply a new tool, albeit an especially powerful one, in a designer's toolbox. They allow clothing designers to make garments that are dynamic and interactive. Clothing that can, for example, change color in response to pollution levels, sparkle when a loved one calls you on the phone, or notify you when your blood pressure increases.

How do you involve people in your research?

I engage with people in a few different ways. First, I design hardware and software tools to help people build new and different kinds of technology. The LilyPad is a good example of this kind of work. I hone these designs by teaching workshops to different groups of people, and once a tool is stable, I work hard to disseminate it to users in the real world. The LilyPad has been commercially available since 2007 and it has been fascinating and exciting to see how a group of real-world designers – who are predominantly female – is using it to build things like smart sportswear, plush video game controllers, soft robots, and interactive embroideries.

I also strive to be as open as possible with my own design and engineering explorations. I document and publish as much information as I can about the materials, tools, and processes I use. I apply an open source approach not only to the software and hardware I create but, as much as I can, to the entire creative process. I develop and share tutorials, classroom and workshop curricula, materials references, and engineering techniques.

What excites you most about your work?

I am infatuated with materials. There is nothing more inspiring than a sheet of heavy paper, a length of wool felt, a rough block of wood, or a box of old motors. My thinking about design and technology is largely driven by explorations of materials and their affordances. So materials are always delightful. But the real-world adoption of tools I've designed and the prospect this presents for changing technology culture is perhaps what's most exciting. My most dearly held goal is to expand and diversify technology culture and it's tremendously rewarding to see evidence that my work is starting to do that. ■

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 nth 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

A pilot 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

Show/Hide

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.

A benefit 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. A lot 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. A preset 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.

The focus group hated it. So he showed it to an out-of-focus group.

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

Show/Hide

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.
- A balance 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. A similar 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. A common 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.

A frequently 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.' A brainstorming 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.

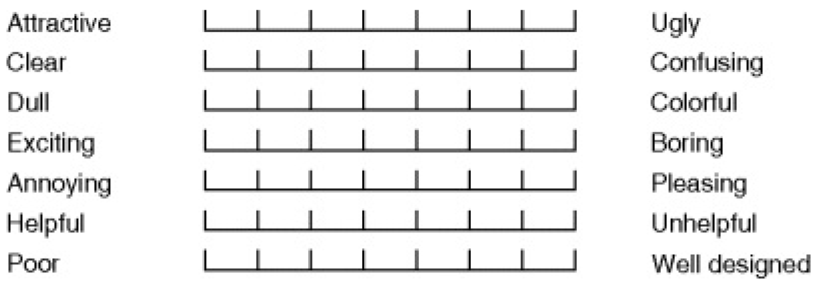


Figure 7.6 An example of a semantic differential scale

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:

☐	☐	☐
yes	don't know	no

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

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

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

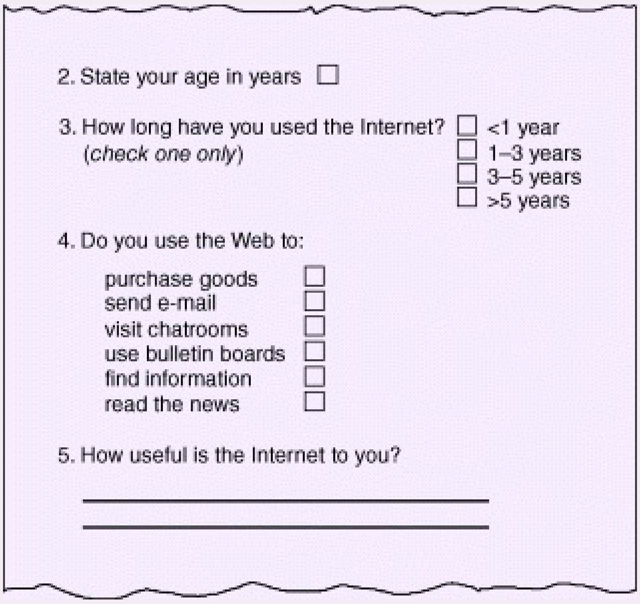
–agree

–slightly disagree

–strongly disagree.■

Activity 7.3

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



2. State your age in years ☐

3. How long have you used the Internet? (*check one only*)
☐ <1 year
☐ 1–3 years
☐ 3–5 years
☐ >5 years

4. Do you use the Web to:
purchase goods ☐
send e-mail ☐
visit chatrooms ☐
use bulletin boards ☐
find information ☐
read the news ☐

5. How useful is the Internet to you?

Figure 7.7 A questionnaire with poorly designed features

Comment

Show/Hide

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).

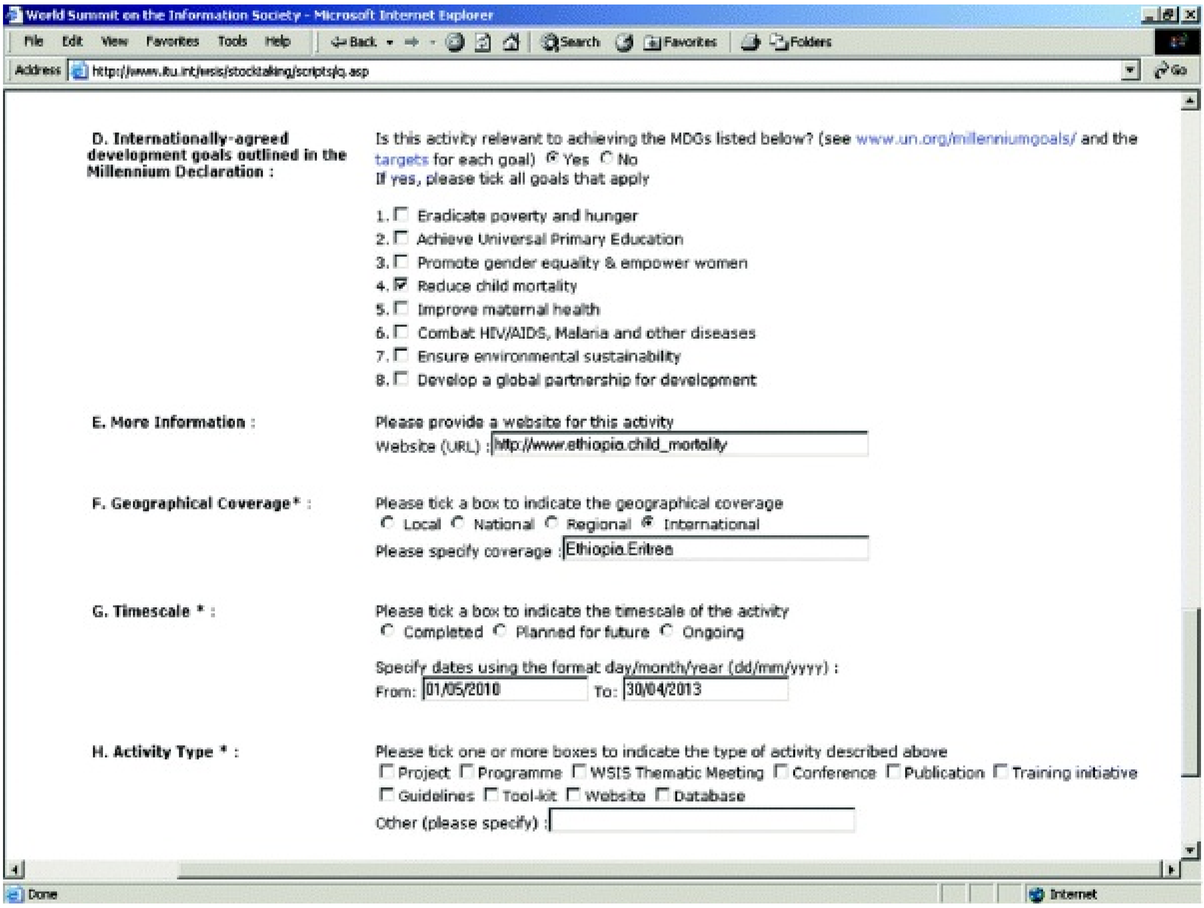


Figure 7.8 An excerpt from a web-based questionnaire showing check boxes, radio buttons, and pull-down menus

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

Show/Hide

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. A usability 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. A usability 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

Show/Hide

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

Show/Hide

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).

(a) (b)

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.

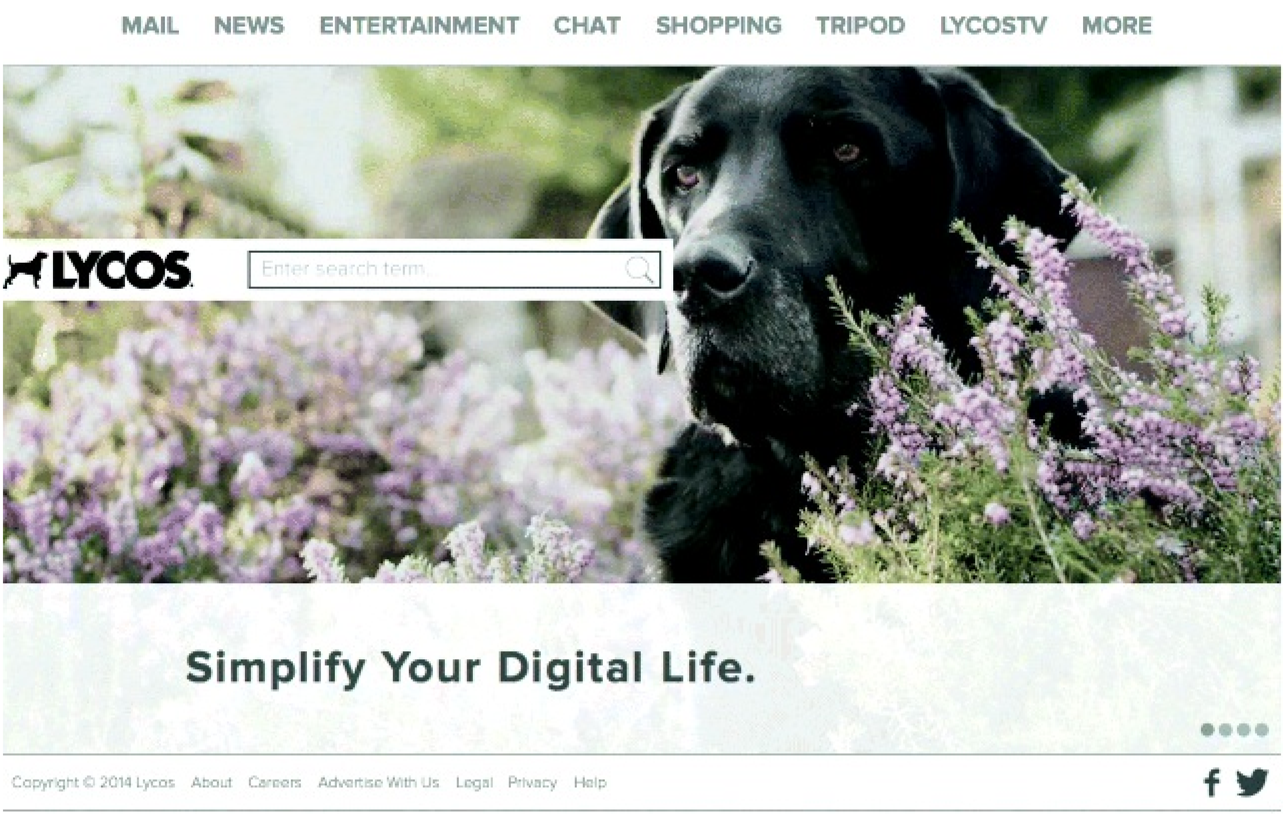


Figure 7.11 Home page of Lycos search engine

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.

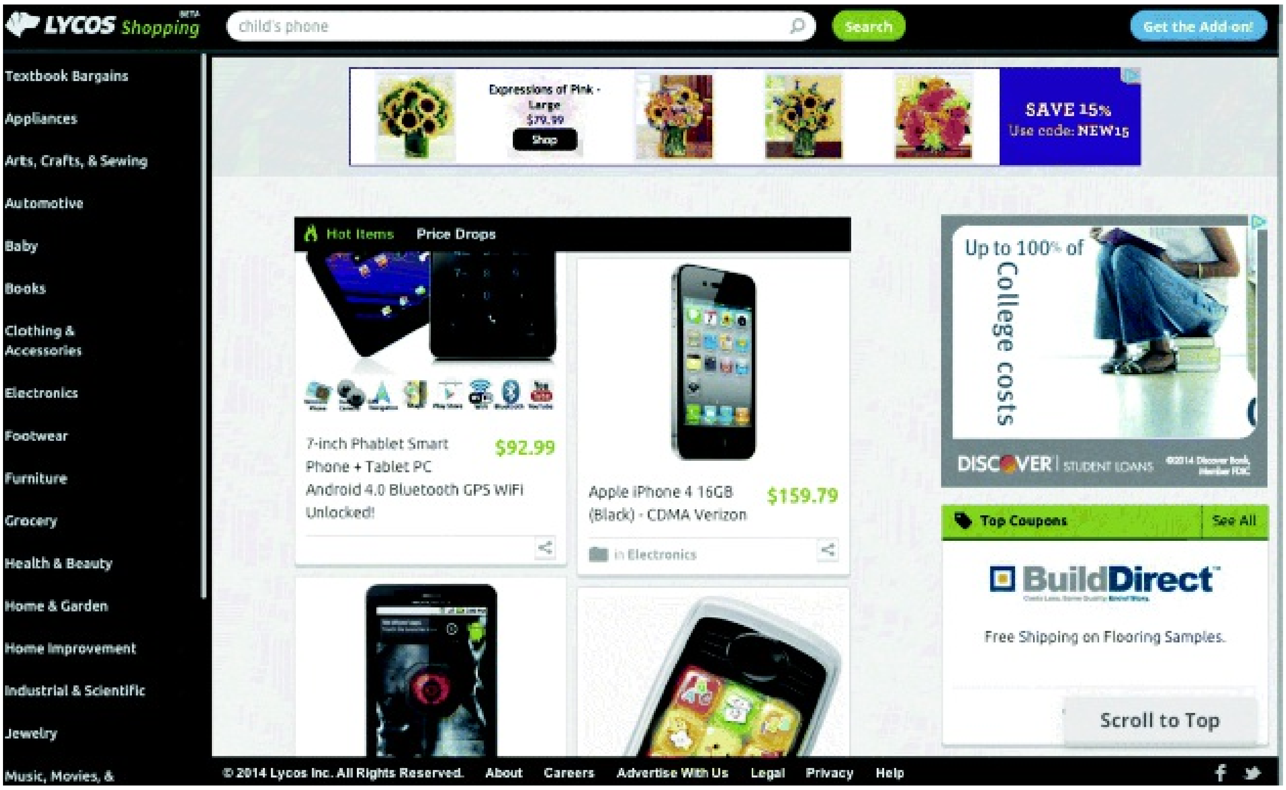


Figure 7.12 The screen that appears in response to choosing ‘child's phone’

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

Show/Hide

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 % of Total: 11.93% (5,768)	74.27% Site Avg: 73.35% (1.25%)	511 % of Total: 12.08% (4,231)	68.46% Site Avg: 56.87% (20.39%)	2.32 Site Avg: 2.94 (-21.17%)	00:01:47 Site Avg: 00:02:32 (-29.56%)
1.	768x1024	173 (25.15%)	60.12%	104 (20.35%)	58.38%	2.66	00:02:17
2.	320x568	137 (19.91%)	84.67%	116 (22.70%)	79.56%	1.53	00:00:45
3.	360x640	40 (5.81%)	87.50%	35 (6.85%)	72.50%	2.00	00:01:50
4.	320x480	34 (4.94%)	82.35%	28 (5.48%)	82.35%	1.35	00:00:07
5.	360x592	19 (2.76%)	94.74%	18 (3.52%)	78.95%	1.68	00:00:13
6.	480x800	19 (2.76%)	52.63%	10 (1.96%)	57.89%	7.05	00:06:24
7.	720x1280	13 (1.89%)	69.23%	9 (1.76%)	69.23%	2.77	00:01:35
8.	1366x768	10 (1.45%)	70.00%	7 (1.37%)	40.00%	4.20	00:05:52
9.	1536x864	10 (1.45%)	50.00%	5 (0.98%)	30.00%	5.20	00:07:35
10.	1280x800	9 (1.31%)	44.44%	4 (0.78%)	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

Show/Hide

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	Understanding context of	Mostly qualitative	Observing gives insights	Very time-consuming.

the field	user activity		that other techniques don't give	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

Show/Hide

Assignment

Part A

The aim of this assignment is for you to practice data gathering. Assume that you have been employed to improve an interactive product such as a smartphone app, an iPod, a DVD recorder, computer software, a photocopying machine, or some other type of technology that interests you. You may either redesign this product, or create a completely new product. To do the assignment you will need to find a group of people or a single individual prepared to be your user group. These could be your family, your friends, or people in your class or local community group.

For this assignment you should:

a. Clarify the basic goal of improving the product by considering what this means in your circumstances.

b. Watch the group (or person) casually to get an understanding of issues that might create challenges for you doing this assignment and get information that might enable you to refine your goals.

c. Explain how you would use each of the three data gathering techniques: interview, questionnaire, and observation in your data gathering program. Explain how your plan takes account of triangulation.

d. Consider your relationship with your user group and decide if an informed consent form is required (Figure 7.1 and Chapter 13 will help you to design your own if needed).

e. Plan your data gathering program in detail:

   - Decide what kind of interview you want to run, and design a set of interview questions for your study. Decide how you will record the data, then acquire and test any equipment needed and run a pilot study.

   - Decide whether you want to include a questionnaire in your data gathering program, and design appropriate questions for it. Run a pilot study to check your questionnaire.

   - Decide whether you want to use direct or indirect observation and where on the outsider–insider spectrum of observers you wish to be. Decide how you will record the data, then acquire and test any equipment needed and run a pilot study.

f. Carry out your study but limit its scope. For example, only interview two or three people or plan only two half-hour observation periods.

g. Reflect on your experience and suggest what you would do differently next time.

   Keep the data you have gathered as this will form the basis of the assignment in Chapter 8.

Part B

This assignment (adapted from Golbeck, 2013) requires you to:

a. Go to https://wordpress.com/ and create a new blog.

b. On your new site, upload an original funny or interesting story, image, or video that you think others would be interested in viewing, making sure that your post does not violate a copyright.

c. Promote your site on social media outlets like Facebook and Twitter and encourage your peers and family to help you.

d. Each day at the same time (e.g. 10am), record the number of people who have visited your blog and the regions visitors came from. Make a chart that shows how these analytics changed over the 7-day period.

e. Analyze your success. How many users were you able to attract? What days attracted the most views? Where did your viewers come from? Based on the analytics data, think about what you could do in the future to attract more visitors.

Take a Quickvote on Chapter 7:
www.id-book.com/quickvotes/chapter7

Summary

This chapter has presented three main data gathering methods that are commonly used in interaction design: interviews, questionnaires, and observation. It has described in detail the planning and execution of each. In addition, five key issues of data gathering were presented, and how to record the data gathered was discussed.

Key points

- All data gathering sessions should have clear goals.
- Depending on the study you plan, you may need to develop an informed consent form and get other permissions to run the study.
- Each planned data gathering session should be tested by running a pilot study.
- Triangulation involves investigating a phenomenon from different perspectives.
- Data may be recorded using handwritten notes, audio or video recording, a camera, or any combination of these.
- There are three styles of interviews: structured, semi-structured, and unstructured.
- Questionnaires may be paper-based, email, or web-based.
- Questions for an interview or questionnaire can be open or closed. Closed questions require the interviewee to select from a limited range of options. Open questions accept a free-range response.
- Observation may be direct or indirect.
- In direct observation, the observer may adopt different levels of participation ranging from insider (participant observer) to outsider (passive observer).
- Choosing appropriate data gathering techniques depends on the focus of the study, the participants involved, the nature of the technique, and the resources available.

Further Reading

Fetterman, D. M. (2010). Ethnography: Step by Step (3rd edn) Applied

Social Research Methods Series, Vol. 17. Sage. This book provides an introduction to the theory and practice of ethnography and is an excellent guide for beginners. It covers both data gathering and data analysis in the ethnographic tradition.

Fulton Suri, J. (2005) Thoughtless Acts? Chronicle Books, San Francisco. This intriguing little book invites you to consider how people react to their environment. It is a good introduction to the art of observation.

Heath, C., hindmarsh, j. and Luff, p. (2010) Video in Qualitative Research: Analyzing social interaction in everyday life. Sage. This is an accessible book which provides practical advice and guidance about how to set up and perform data gathering using video recording. It also covers data analysis, presenting findings and potential implications from video research based on their own experience.

Olson, J. S. and Kellogg, W. A. (eds) (2014) Ways of Knowing in HCI. Springer. This edited collection contains useful chapters on a wide variety of data collection and analysis techniques. Some topics that are particularly relevant to this chapter include: ethnography, experimental design, log data collection and analysis, ethics in research, and more.

Oppenheim, A. N. (1998) Questionnaire Design, Interviewing and Attitude Measurement. Pinter Publishers. This text is now a classic but it is useful for reference. It provides a detailed account of all aspects of questionnaire design, illustrated with many examples. However, care will be needed in applying some of these suggestions to online questionnaires.

Robson, C. (2011) Real World Research (3rd edn). John Wiley & Sons. This book provides comprehensive coverage of data gathering and analysis techniques and how to use them. Early books and related books by Robson also address topics discussed in this chapter.

Sue, V. M. and Ritter, L. A. (2012) Conducting Online Surveys. Sage. This small book describes the process of conducting online surveys including how to set the survey goals, design the questions, implement the questionnaire, identify a suitable group of potential respondents, and analyze the results.

CHAPTER 8
DATA ANALYSIS, INTERPRETATION, AND PRESENTATION

Objectives

The main aims of this chapter are to:

- Discuss the difference between qualitative and quantitative data and analysis.
- Enable you to analyze data gathered from questionnaires.
- Enable you to analyze data gathered from interviews.
- Enable you to analyze data gathered from observation studies.
- Make you aware of software packages that are available to help your analysis.
- Identify some of the common pitfalls in data analysis, interpretation, and presentation.
- Enable you to be able to interpret and present your findings in a meaningful and appropriate manner.

8.1 Introduction

The kind of analysis that can be performed on a set of data will be influenced by the goals identified at the outset, and the data actually gathered. Broadly speaking, you may take a qualitative analysis approach or a quantitative analysis approach, or a combination of qualitative and quantitative. The last of these is very common as it provides a more comprehensive account of the behavior being observed or performance being measured.

Most analysis, whether it is quantitative or qualitative, begins with initial reactions or observations from the data. This might involve identifying patterns or calculating simple numerical values such as ratios, averages, or percentages. This initial analysis is followed by more detailed work using structured frameworks or theories to support the investigation.

Interpretation of the findings often proceeds in parallel with analysis, but there are different ways to interpret results and it is important to make sure that the data supports your conclusions. A common mistake is for the investigator's existing beliefs or biases to influence the interpretation of results. Imagine that through initial analysis of your data you have discovered a pattern of responses to customer care questionnaires which indicates that inquiries from customers that are routed through the Sydney office of an organization take longer to process than those routed through the Moscow office. This result can be interpreted in many different ways. Which do you choose? You may conclude that the customer care operatives in Sydney are less efficient, or you may conclude that the customer care operatives in Sydney provide more detailed responses, or you may conclude that the technology supporting the processing of inquiries needs to be updated in Sydney, or you may conclude that customers reaching the Sydney office demand a higher level of service, and so on. In order to determine which of these potential interpretations is more accurate, it would be appropriate to look at other data such as customer inquiry details, and maybe interviews with staff.

Another common mistake is to make claims that go beyond what the data can support. This is a matter of interpretation and of presentation. The words ‘many’ or ‘often’ or indeed ‘all’ need to be used very carefully when reporting conclusions. An investigator should remain as impartial and objective as possible if the conclusions are to be believed, and showing that your conclusions are supported by your results is an important skill to develop.

Finally, finding the best way to present your findings is equally skilled, and depends on your goals but also on the audience for whom the results were produced. For example, in the requirements activity you might choose to present your findings using a formal notation, while reporting the results of an evaluation to the team of developers might involve a summary of problems found, supported by video clips of users experiencing those problems.

In this chapter we will introduce a variety of methods and describe in more detail how to approach data analysis using some of the common approaches taken in interaction design.

8.2 Qualitative and Quantitative

Quantitative data is data that is in the form of numbers, or that can easily be translated into numbers. For example, the number of years’ experience the interviewees have, the number of projects a department handles at a time, or the number of minutes it takes to perform a task. Qualitative data is not expressed in numerical terms. For example, qualitative data includes descriptions, quotes from interviewees, vignettes of activity, and images. It is possible to express qualitative data in numerical form, but it is not always meaningful to do so – see Box 8.1.

It is sometimes assumed that certain forms of data gathering can only result in quantitative data and others can only result in qualitative data. However, this is a fallacy. All the forms of data gathering discussed in the previous chapter may result in qualitative and quantitative data. For example, on a questionnaire, questions about the participant's age or number of software packages they use a day will result in quantitative data, while any comment fields will result in qualitative data. In an observation, quantitative data you may record includes the number of people involved in a project, or how many hours a participant spends trying to sort out a problem they encounter, while notes about the feelings of frustration, or the nature of interactions between team members, are qualitative data.

Quantitative analysis uses numerical methods to ascertain the magnitude, amount, or size of something; for example, the attributes, behavior, or opinions of the participants. For example, in describing a population, a quantitative analysis might conclude that the average person is 5 feet 11 inches tall, weighs 180 pounds, and is 45 years old. Qualitative analysis focuses on the nature of something and can be represented by themes, patterns, and stories. For example, in describing the same population, a qualitative analysis might conclude that the average person is tall, thin, and middle-aged.

BOX 8.1

Use and Abuse of Numbers

Numbers are infinitely malleable and can make a very convincing argument, but it is important to be clear why you are manipulating quantitative data, and what the implications will be. Before adding a set of numbers together, finding an average, calculating a percentage, or performing any other kind of numerical translation, consider whether the operation is meaningful in your context.

Qualitative data can also be turned into a set of numbers. Translating non-numerical data into a numerical or ordered scale is appropriate at times, and this is a common approach in interaction design. However, you need to be careful that this kind of translation is meaningful in the context of your study. For example, assume that you have collected a set of interviews from sales representatives regarding the use of a new mobile product for reporting sales queries. One way of turning this data into a numerical form would be to count the number of words uttered by each of your interviewees. You might then draw conclusions about how strongly the sales representatives feel about the mobile devices, e.g. the more they had to say about the product, the stronger they feel about it. But do you think this is a wise way to analyze the data? This set of quantitative data is unlikely to be of much use in answering your study questions.

Other, less obvious abuses include translating small population sizes into percentages. For example, saying that 50% of users take longer than 30 minutes to place an order through an e-commerce website carries a different meaning than saying that two out of four users had the same problem. It is better not to use percentages unless the number of data points is at least over 10, and even then it is appropriate to use both percentages and raw numbers, to make sure that your claim is not misunderstood.

It is possible to perform legitimate statistical calculations on a set of data and still to present misleading results by not making the context clear, or by choosing the particular calculation that gives the most favorable result (Huff, 1991). If you are not comfortable dealing with numbers, it is better to ask for help from someone who is, because it is easy to unintentionally misrepresent your data. ■

8.2.1 The First Steps in Analyzing Data

Having performed data gathering sessions, there is some initial processing of the data normally required before data analysis can begin in earnest. There are many different combinations of data, but here we discuss typical data collected through interviews, questionnaires, and observation sessions. This information is summarized in Table 8.1.

Interviews.

Raw interview data is usually in the form of audio recordings and interviewer notes. The notes need to be written up and expanded as soon as possible after the interview has taken place so that the interviewer's memory is clear and fresh. The audio recording may be used to help in this process, or it may be transcribed for more detailed analysis. Transcription takes significant effort, as people talk more quickly than most people can type (or write), and the recording is not always clear. It is therefore worth considering whether or not to transcribe the whole interview, or just sections of it that are relevant to your investigation.

Interviews are sometimes video recorded, especially if the interviewee is given a task to perform or props are used to prompt discussion. The audio channel of the video data may also be transcribed.

Table 8.1 Data gathered and typical initial processing steps for the main data gathering techniques

	Usual raw data	Example qualitative data	Example quantitative data	Initial processing steps
Interviews	Audio recordings. Interviewer notes. Video recordings	Responses to open questions. Video pictures. Respondent's opinions	Age, job role, years of experience. Responses to closed questions	Transcription of recordings. Expansion of notes
Questionnaires	Written responses. Online database	Responses to open questions. Responses in ‘further comments’ fields. Respondent's opinions	Age, job role, years of experience. Responses to closed questions	Clean up data. Filter into different data sets
Observation	Observer's notes. Photographs. Audio and video recordings. Data logs. Think-aloud	Records of behavior. Description of a task as it is undertaken. Copies of informal procedures	Demographics of participants. Time spent on a task. The number of people involved in an activity	Expansion of notes. Transcription of recordings. Synchronization between data recordings

Closed questions are usually treated as quantitative data and analyzed using simple quantitative analysis (see below). For example, a question that asks for the respondent's age range can easily be analyzed to find out the percentage of respondents in each range. More complicated statistical techniques are needed to identify relationships between question responses that can be generalized, such as being able to say that men over the age of 35 all believe that buttons on cell phones are too small. Open questions typically result in qualitative data which might be searched for categories or patterns of response.

Questionnaires.

Raw data from questionnaires consists of the respondents' answers to the questions, and these may be in written format, or for online surveys, the data is likely to be in a database. It may be necessary to clean up the data by removing entries where the respondent has misunderstood a question.

The data can be filtered according to respondent subpopulations, (e.g. everyone under 16) or according to a particular question (e.g. to understand respondents' reactions to color). This allows analyses to be conducted on subsets of the data, and hence to draw detailed conclusions for more specific goals. This is made easier by the use of a simple tool such as a spreadsheet, as discussed below.

As for interviews, closed questions are likely to be analyzed quantitatively and open questions qualitatively.

Observation.

This kind of data gathering can result in a wide variety of raw data including observer's notes, photographs, data logs, think-aloud recordings (often called protocols), video and audio recordings. All this raw data presents a rich picture of the activity under observation, but it can also make it difficult to analyze unless a structured framework is adopted. Initial data processing here would include writing up and expanding notes, and transcribing elements of the audio and video recordings and the think-aloud protocols. For observation in a controlled environment, initial processing might also include synchronizing different data recordings.

Transcriptions and the observer's notes are most likely to be analyzed using qualitative approaches, while photographs provide contextual information. Data logs and some elements of the observer's notes would probably be analyzed quantitatively.

Throughout this initial processing, patterns and themes in the data may present themselves. It is useful to make a note of these initial impressions to use as a basis for further, more detailed analysis, but don't rely on these initial impressions alone as you may be unintentionally biased by them.

8.3 Simple Quantitative Analysis

Explaining statistical analysis requires a whole book on its own. We will not try to explain statistics in any detail, although some basic statistical terms and use of statistics are discussed further in Chapter 14. Here, we introduce some simple quantitative analysis techniques you can use effectively in an interaction design context. The techniques explored here are averages and percentages. Percentages are useful for standardizing the data, particularly if you want to compare two or more large sets of responses.

Averages and percentages are fairly well-known numerical measures. However, there are three different types of average and which one you use changes the meaning of your results. These three are: mean, median, and mode. Mean refers to the commonly understood interpretation of average: i.e. add together all the figures and divide by the number of figures you started with. Median and mode are less well-known but are very useful. The median is the middle value of the data when the numbers are ranked. The mode is the most commonly occurring number. For example, in a set of data (2, 3, 4, 6, 6, 7, 7, 7, 8), the median is 6 and the mode is 7, while the mean is 50/9 = 5.56. In this case, the difference between the different averages is not that great. However, consider the set (2, 2, 2, 2, 450). Now the median is 2, the mode is 2, and the mean is 458/5 = 91.6!

“Looks good. Let me run it past the number-crunchers.”

Use of simple averages can provide useful overview information, but they need to be used with caution. Karapanos et al (2009) go further and suggest that averaging treats diversity among participants as error and proposes the use of a multidimensional scaling approach instead.

Before any analysis can take place, the data needs to be collated into analyzable data sets. Quantitative data can usually be translated into rows and columns, where one row equals one record, e.g. respondent or interviewee. If these are entered into a spreadsheet such as Excel, this makes simple manipulations and data set filtering easier. Before entering data in this way, it is important to decide how you will represent the different possible answers. For example, ‘don't know’ represents a different response from no answer at all and they need to be distinguished, e.g. with separate columns in the spreadsheet. Also, if dealing with options from a closed question, such as job role, there are two different possible approaches which affect the analysis. One approach is to have a column headed Job role and to enter the job role as it is given to you by the respondent or interviewee. The alternative approach is to have a column for each possible answer. The latter approach lends itself more easily to automatic summaries. Note, however, that this option will only be open to you if the original question was designed to collect the appropriate data (see Box 8.2).

BOX 8.2

How Question Design Affects Data Analysis

Activity 7.2 asked you to suggest some interview questions that you might ask a colleague to help evaluate e-readers. We shall use this example here to illustrate how different question designs affect the kinds of analysis that can be performed, and the kind of conclusions that can be drawn.

Assume that you have asked the question: ‘How do you feel about e-readers?’ Responses to this will be varied and may include that they are cool, lightweight, easy to carry, too expensive, difficult to use, technically complex, and so on. There are many possibilities, and the responses would need to be treated qualitatively. This means that analysis of the data must consider each individual response. If you have only 10 or so responses then this may not be too bad, but if you have many more then it becomes harder to process the information, and harder to summarize your findings. This is typical of open-ended questions – answers are not likely to be homogeneous and so will need to be treated individually. In contrast, answers to a closed question, which gives respondents a fixed set of alternatives to choose from, can be treated quantitatively. So, for example, instead of asking ‘How do you feel about e-readers?’ assume that you have asked ‘In your opinion, are e-readers easy to use or tedious to use?’ This clearly reduces the number of options and you would then record the response as ‘easy to use,’ ‘tedious to use,’ or ‘neither.’

When entered in a spreadsheet, or a simple table, initial analysis of this data might look like the following:

Respondent	Easy to use	Tedious to use	Neither
A	1		
B			
C		1	
. . .		1	
Z			
Total			1
	14	5	7

Based on this, we can then say that 14 out of 26 (54%) of our respondents think e-readers are easy to use, 5 out of 26 (19%) think they are tedious to use, and 7 out of 26 (27%) think they are neither easy to use nor tedious. Note also that in the table, respondents' names are replaced by letters so that they are identifiable but anonymous to any onlookers. This strategy is important for protecting participants' privacy, which is usually assured in a consent form (see the example in Figure 7.1).

Another alternative that might be used in a questionnaire is to phrase the question in terms of a Likert scale, such as the one below. This again alters the kind of data and hence the kind of conclusions that can be drawn:

In your opinion, are e-readers easy to use:

strongly agree	agree	neither	disagree	strongly disagree
☐	☐	☐	☐	☐

Then the data could be analyzed using a simple spreadsheet or table:

Respondent	Strongly agree	Agree	Neither	Disagree	Strongly disagree
A		1			
B	1				
C				1	
. . .					
Z					
					1
Total	5	7	10	1	3

In this case we have changed the kind of data we are collecting, and cannot, based on this second set, say anything about whether respondents think e-readers are tedious to use, as we have not asked that question. We can only say that, for example, 4 out of 26 (15%)

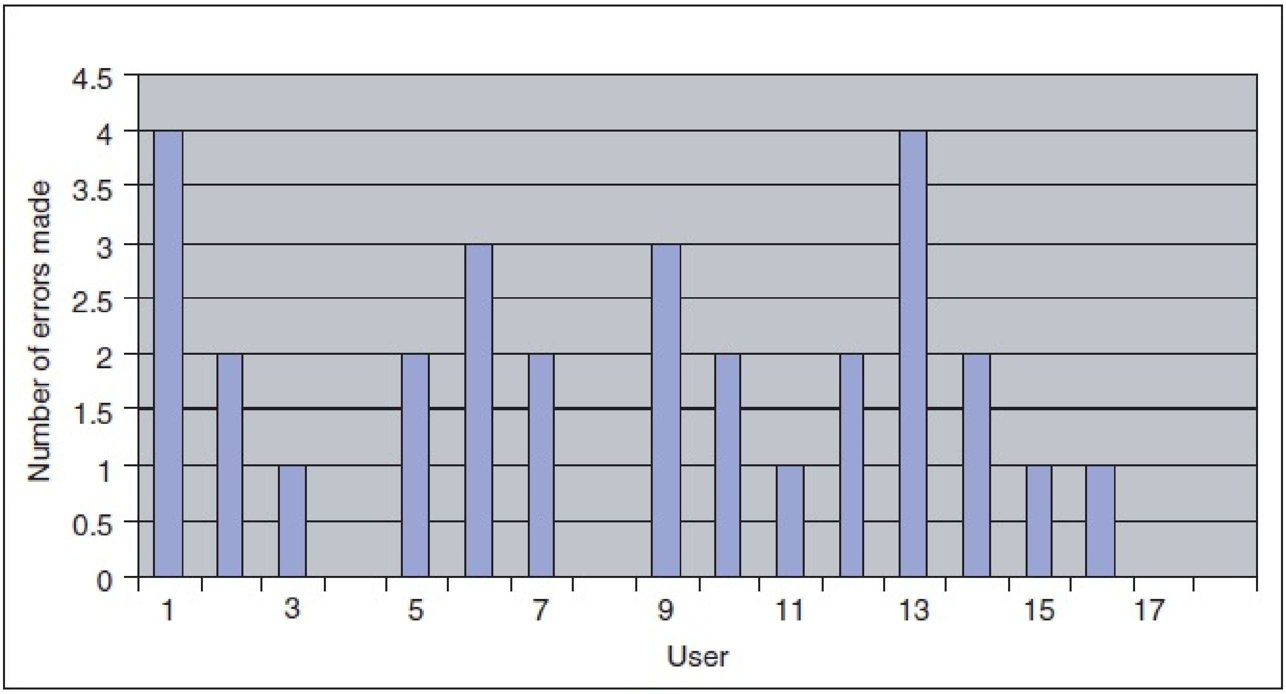
disagreed with the statement that e-readers are easy to use (and of those, 3 (11.5%) strongly disagreed).■

For simple collation and analysis, spreadsheet software such as Excel is often used as it is commonly available, is well understood, and offers a variety of numerical manipulations and graphical representations. Initial analysis might involve finding out averages, and identifying any outliers, i.e. values that are significantly different from the others. Producing a graphical representation of the data helps to get an overall view of the data and any patterns it contains.

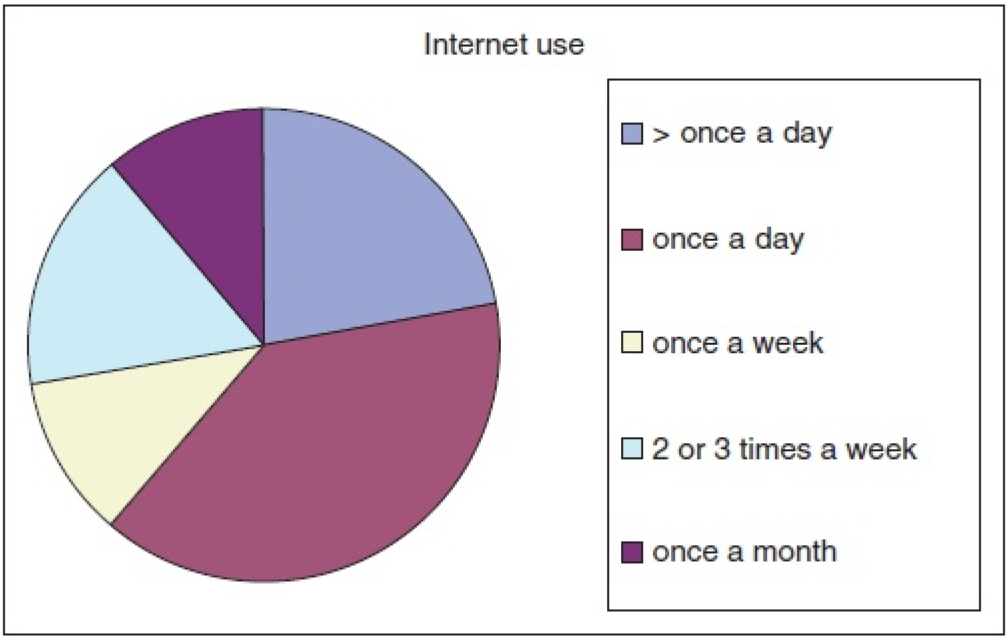
For example, consider the set of data shown in Table 8.2, which was collected during an evaluation study of a document sharing application. This data shows the experience of the users and the number of errors made while trying to complete a controlled task. It was captured automatically and recorded in a spreadsheet; then the totals and averages were calculated. The graphs in Figure 8.1 were generated using the spreadsheet package. They show an overall view of the data set. In particular, we can see that there are no significant outliers in the error rate data. Whether or not you choose to present these graphical representations to your target audience, it is valuable to use them for your own data analysis.

Table 8.2 Data gathered during a study of a document sharing application

Internet use						
User	**More than once a day**	**Once a day**	**Once a week**	**Two or three times a week**	**Once a month**	**Number of errors made**
1		1				4
2	1					2
3			1			1
4	1					0
5				1		2
6		1				3
7	1					2
8		1				0
9						3
10	1				1	2
11				1		1
12			1			2
13		1				4
14		1				2
15						1
16				1		1
17		1				0
18		1			1	0
Totals	**4**	**7**	**2**	**3**	**2**	**30**
					Mean	**1.67** (to 2 decimal places)



(a)



(b)

Figure 8.1 Graphical representations of the data in Table 8.2: (a) the distribution of errors made (take note of the scale used in these graphs, as seemingly large differences may be much smaller in reality), and (b) the spread of Internet experience within the participant group

If we add one more user to Table 8.2 with an error rate of 9, we can see in Figure 8.2 how using a scatter graph helps to identify outliers. Outliers are usually removed from the larger data set because they distort the general patterns. However, they may also be interesting cases to investigate further to see if there are special circumstances surrounding those users and their session.

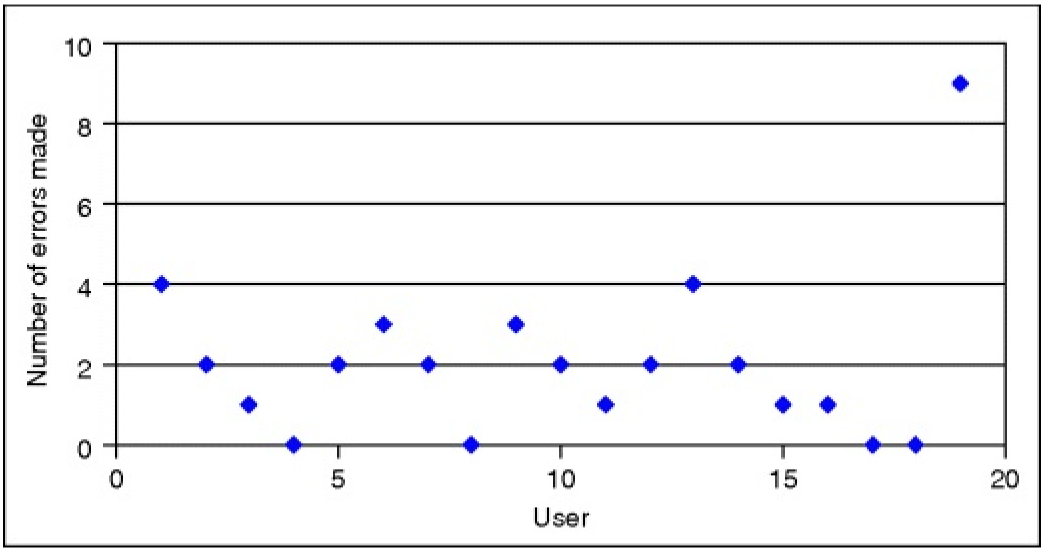


Figure 8.2 Using a scatter diagram helps to identify outliers in your data quite quickly

These initial investigations also help to identify other areas for further investigation. For example, is there something special about the users with error rate 0, or something distinctive about the performance of those who use the Internet only once a month?

Activity 8.1

The data in the table below represents the time taken for a group of users to select and buy an item from an online shopping website.

Using a spreadsheet application to which you have access, generate a bar graph and a scatter diagram to give you an overall view of the data. From this representation, make two initial observations about the data that might form the basis of further investigation.

User	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Time to complete (mins)	15	10	12	10	14	13	11	18	14	17	20	15	18	24	12	16	18	20	26

Comment

Show/Hide

It is fairly straightforward to compare two sets of results, e.g. from the evaluation of two interactive products, using these kinds of graphical representations of the data. Semantic differential data can also be analyzed in this way and used to identify trends, provided the format of the question is appropriate. For example, the following question was asked in a questionnaire to evaluate two different smartphone designs:

For each pair of adjectives, place a cross at the point between them that reflects the extent to which you believe the adjectives describe the smartphone design. Please place only one cross between the marks on each line.

Annoying						*Pleasing*
Easy to use						*Difficult to use*
Value-for-money						*Expensive*
Attractive						*Unattractive*
Secure						*Not secure*
Helpful						*Unhelpful*
Hi-tech						*Lo-tech*
Robust						*Fragile*
Inefficient						*Efficient*
Modern						*Dated*

Tables 8.3 and 8.4 show the tabulated results from 100 respondents who replied to the questionnaire. Note that the responses have been translated into five possible categories, numbered from 1 to 5, based on where the respondent marked the line between each pair of adjectives. It is possible that respondents may have intentionally put a cross closer to one side of the box than the other, but it is acceptable to lose this nuance in the data, provided the original data is not lost, and any further analysis could refer back to it.

Table 8.3 Phone 1

	1	2	3	4	5	
Annoying	35	20	18	15	12	Pleasing
Easy to use	20	28	21	13	18	Difficult to use
Value-for-money	15	30	22	27	6	Expensive
Attractive	37	22	32	6	3	Unattractive
Secure	52	29	12	4	3	Not secure
Helpful	33	21	32	12	2	Unhelpful
Hi-tech	12	24	36	12	16	Lo-tech
Robust	44	13	15	16	12	Fragile
Inefficient	28	23	25	12	12	Efficient
Modern	35	27	20	11	7	Dated

Table 8.4 Phone 2

	1	2	3	4	5	
Annoying						Pleasing
Easy to use	24	23	23	15	15	Difficult to use
Value-for-money	37	29	15	10	9	Expensive
Attractive	26	32	17	13	12	Unattractive
Secure	38	21	29	8	4	Not secure
Helpful	43	22	19	12	4	Unhelpful
Hi-tech	51	19	16	12	2	Lo-tech
Robust	28	12	30	18	12	Fragile
Inefficient	46	23	10	11	10	Efficient
Modern	10	6	37	29	18	Dated
	3	10	45	27	15	

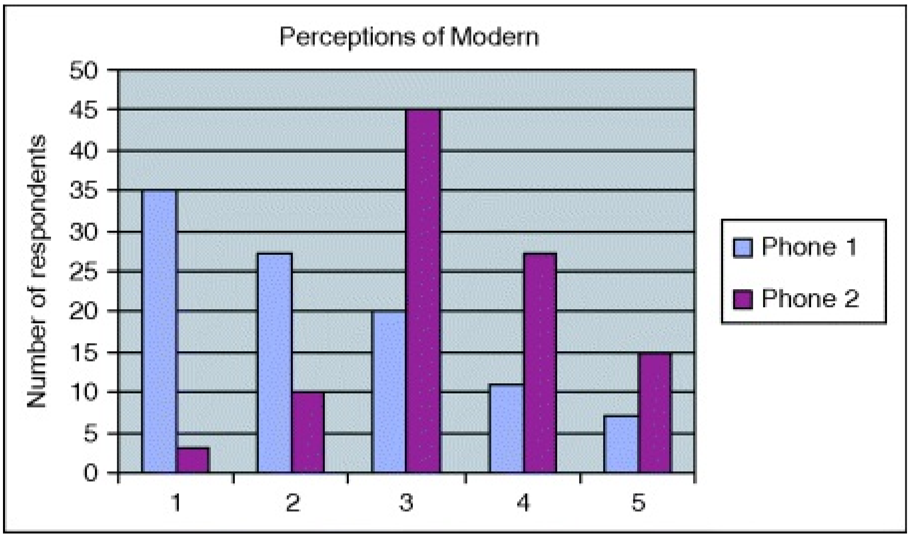


Figure 8.3 A graphical comparison of two smartphone designs according to whether they are perceived as modern or dated

The graph in Figure 8.3 shows how the two smartphone designs varied according to the respondents' perceptions of how modern the design is. This graphical notation shows clearly how the two designs compare.

Data logs that capture users' interactions with a system can also be analyzed and represented graphically, thus helping to identify patterns in behavior. Also, more sophisticated manipulations and graphical images can be used to highlight patterns in collected data. Box 8.3 describes how data logs of an online computer game were used to identify patterns of interactions. This example also shows how observational data can be used to interpret quantitative data.

The examples given in this section have largely focused on data sets which have more than 10 records (i.e. respondents or interviewees). If only a small number of records are collected, then it may be more important to analyze the individual records in more depth than to identify trends. In this case, tabulating the data for ease of presentation may be sufficient.

BOX 8.3

Identifying Interaction Patterns in Log Data from a Massively Multiplayer Online Role-Playing Game

Massively Multiplayer Online Role-Playing Games (MMORPGs) (e.g. World of Warcraft, EverQuest II, Toontown, etc.) involve hundreds of thousands of players interacting on a daily basis within a virtual world, and working towards achieving certain targets. While the challenges offered by these games are often the same as those in a single-player environment, the attraction of MMORPGs is the ability to join a community of gamers, to gain a reputation within that community, and to share gaming experiences directly. Several games have been designed so that players have to collaborate in order to progress.

Ducheneaut and Moore (2004) wanted to investigate how successful MMORPGs are at encouraging interactivity between players. To do this, they analyzed data logs and video recordings of player-to-player interactions in the game Star Wars Galaxies (SWG), complemented by a three-month ethnography of the same environment. The ethnography was achieved by the researchers joining the SWG community as players and using the system regularly over three months. During this time, they identified two locations within the virtual world which were heavily used by other players – the cantina and starport in Coronet City. The cantina is where entertainers can be found and players often go to recover from battle fatigue; players have to go to the starport in order to travel between locations, and shuttles fly about every 9 minutes (see Figure 8.4).

To collect a log of player interactions, they created two characters, placed one in each location for a month, and recorded all public utterances and gestures at these locations. Twenty-six days of data were recorded, with 21 hours a day. This resulted in 100 MB of data and represented a total of 5493 unique players in the two locations. A purpose-built parser was used to identify who was interacting with whom, in what way (gesture or chat), where, when, and what was the content of the interaction. In this context, a gesture may be a smile, greet, clap, cheer, etc. They then analyzed the data for patterns of behavior. One finding was that a small number of players were frequently present in one location while there were many others who visited for only a short time. The median number of days a player was present was 2, while the average was 3.5; only 2% of the total number of players were present more than half the time. Another aspect they investigated was the activity within the cantina over the course of a day. Figure 8.5 shows a summary graph of activity in the cantina for the 26 days. This shows a fairly even distribution of activity throughout the day, with gestures representing about one-third of the events and public utterances representing two-thirds.

Their analysis of the 10 most popular gestures is summarized in the table below:

Figure 8.4 The cantina in SWG's Coronet City

Source: N. Ducheneaut and R.J. Morris (2004): "The social side of gaming: a study of interaction patterns in a massively multiplayer online game" in Proceedings of CSCW 04. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

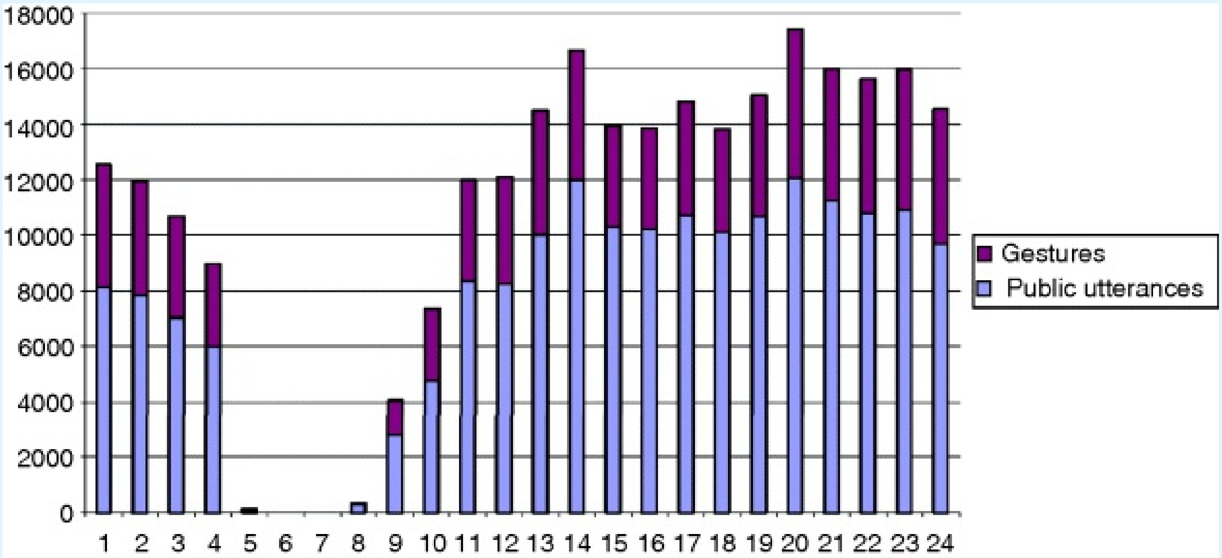


Figure 8.5 Summary of the activity in the cantina over the course of a day. (The gap between 4 a.m. and 7 a.m. is due to a regular server reboot during this time)

Source: N. Ducheneaut and R.J. Morris (2004): "The social side of gaming: a study of interaction patterns in a massively multiplayer online game" in Proceedings of CSCW 04. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

Gesture	% of total (cantina)	Gesture	% of total (starport)
Smile	18.13	Thank	15.95
Cheer	9.57	Bow	12.29
Clap	7.77	Wave	9.81

Wave	6.27	Flail	8.17
Wink	4.22	Smile	7.89
Grin	3.72	Nod	7.03
Nod	3.23	Salute	2.48
Bow	3.22	Pet	1.95
Thank	2.51	Puke	1.89
Greet	2.40	Cheer	1.56

These two kinds of analysis are helpful to get an overview of the different players' interactions but do not indicate the richness of social interaction each player is engaged with. So Ducheneaut and Moore analyzed the interactions on three dimensions for each player: the number of gestures received, the number of gestures made, and the number of public utterances made. Having done so, they concluded that the average player goes into the cantina, makes about one gesture to another player, exchanges about four sentences with him or her, and receives one gesture in return. This conclusion was arrived at from taking averages across the data, but in order to get a clearer view of interactions, they plotted dimensions for each individual set of data on a graph. This is reproduced in Figure 8.6. The x-axis represents the number of gestures received, the y-axis represents the number of gestures made, and the size of the 'dot' is proportional to the number of public utterances made by the player. This graphical representation illustrated an unexpected finding – that the majority of players do not interact very much. Another set of players make a large number of utterances but make or receive no gestures. Yet another (smaller) set of players gesture and talk a lot, but receive few gestures in return.

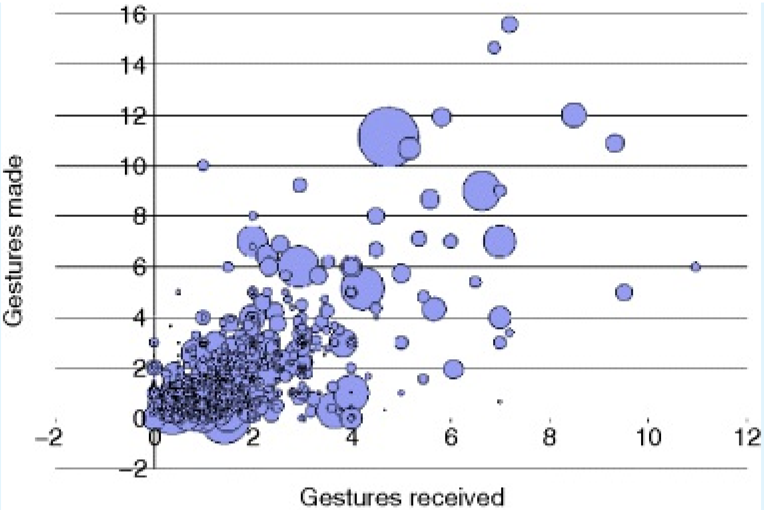


Figure 8.6 Interaction profiles of players in the cantina

Source: N. Ducheneaut and R.J. Morris (2004): "The social side of gaming: a study of interaction patterns in a massively multiplayer online game" in Proceedings of CSCW 04. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

The researchers concluded that these last two kinds of behavior are due to the player programming their avatar to repeat actions even when the player is not logged on. This kind of behavior is not truly interactive as it is designed simply to advance the player within the game (one way of gaining points is to repeatedly perform activities related to the avatar's profession). This behavior then affects the social atmosphere of the cantina because other players are unhappy about the false kind of interaction.

Bubbles to the right of this graph represent players who interact a lot – making and receiving gestures, and chatting. These players are engaging in the kind of social interaction that the designers of SWG want to promote.

A similar analysis was performed for the starport (Figure 8.7), but a different pattern of interactions was found. A large number of players made and received no gestures, but made a lot of public utterances. The ethnographic data helped researchers to interpret this finding too – the starport was a good place to advertise as there were many people gathered waiting for a shuttle. Another set of players at the starport said very little; the researchers believe that these were people looking for trainers to give them a particular skill they needed in order to progress in the game. ■

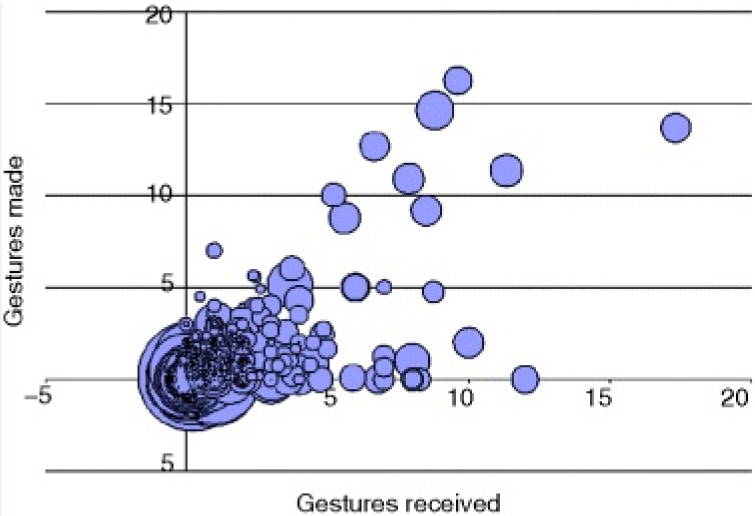


Figure 8.7 Interaction profiles of players in the starport

Source: N. Ducheneaut and R.J. Morris (2004): "The social side of gaming: a study of interaction patterns in a massively multiplayer online game" in Proceedings of CSCW 04. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

8.4 Simple Qualitative Analysis

As with quantitative analysis, the first step in qualitative analysis is to gain an overall impression of the data and to start looking for patterns. Some patterns will have emerged during the data gathering itself, and so you may already have some idea of the kinds of pattern to look for, but it is important to confirm and re-confirm findings to make sure that initial impressions are not biasing analysis. For observation data, the guiding framework will have given some structure to the data. For example, the practitioner's framework for observation introduced in Chapter 7 will have resulted in a focus on who, where, and what, while using Robson's more detailed framework will result in patterns relating to physical objects, people's goals, sequences of events, and so on.

There are three simple types of qualitative analysis that we discuss here: identifying recurring patterns and themes, categorizing data, and analyzing critical incidents. These are not mutually exclusive and can be used in combination.

8.4.1 Identifying Recurring Patterns or Themes

As you become more familiar with the data, possible themes or patterns will emerge. An example might be noticing that people visiting TripAdviser.com look for reviews for a hotel that are rated 'terrible' first. Any initial impressions must be confirmed and refined with more rigorous analysis, seeking both confirming and disconfirming evidence in the data. Sometimes the refined patterns or themes form the primary set of findings for the analysis and sometimes they are just the starting point for different analyses.

The study goals provide an orienting focus for the formulation of themes. For example, consider a survey to evaluate whether the information displayed on a train travel website is appropriate and sufficient. Several of the respondents suggest that the station stops in between the origin and destination stations should be displayed. This theme is relevant to the study goals and would be reported as a main theme. In another part of the survey, under further comments you might notice that several respondents say the company's logo is distracting. Although this too is a theme in the data, it is not directly relevant to the study's goals and may be reported only as a minor theme.

There are different techniques for identifying themes in qualitative data (e.g. Braun and Clarke, 2006). The affinity diagram, which is used in contextual design (Beyer and Holtzblatt, 1998; Holtzblatt, 2001) is one common technique used in qualitative analysis. It aims to organize individual ideas and insights into a hierarchy showing common structures and themes. Notes are grouped together because they are similar in some fashion. The groups are not predefined, but emerge from the data. The process was originally introduced into the software quality community from Japan, where it is regarded as one of the seven quality processes. The affinity diagram is built by a process of induction. One note is put up first, and then the team searches for other notes that are related in some way. For example, De Angeli et al (2004) collected data through field observations and semi-structued interviews to investigate the use of ATMs (automated teller machines) in Mumbai, India. As part of their data analysis they used affinity diagrams to cluster issues into themes (see Figure 8.8).

Figure 8.8 Building the affinity diagram of Indian ATM usage

Source: Figure 1, A. DeAngeli, U. Athavamker, A. Joshi, L. Coventry and G.I. Johnson (2004) "Introducing ATMs in India: a contextual inquiry", Interacting with Computers 16(1), 29–44. Reproduced with permission.

Note that patterns and themes in your data may relate to a variety of aspects: to behavior, to your user group, to places or situations where certain events happen, and so on. Each of these kinds of theme may be relevant to your goals. For example, descriptions of typical users (personas) may be an outcome of data analysis that focuses on patterns of participant characteristics. Although we include thematic analysis under qualitative analysis, patterns and themes may also emerge from quantitative data.

8.4.2 Categorizing Data

Transcripts of meetings, interviews, or think-aloud protocols can be analyzed at a high level of detail, such as identifying stories or themes, or at a fine level of detail in which each word, phrase, utterance, or gesture is analyzed. Either way, elements identified in the data are usually categorized first using a categorization scheme. The categorization scheme may arise from the data itself, if the investigation is exploratory, as it might be in the requirements activity, or it might originate elsewhere in a well-recognized categorization scheme, or a combination of these two approaches may be used. The principle here is that the data is divided up into elements and each element is then categorized.

Which categories to use is largely determined by the goal of the study. One of the most challenging aspects is determining meaningful categories that are orthogonal (i.e. do not overlap). Another is deciding on the appropriate granularity for the categories (e.g. at word, phrase, sentence, or paragraph level); this is also dependent on the goal of the study and the data being analyzed.

The categorization scheme used must be reliable so that the analysis can be replicated. This can be demonstrated by training a second person to use the categories. When training is complete, both people analyze the same data sample. If there is a large discrepancy between the two analyses, either training was inadequate or the categorization is not working and needs to be refined. When a high level of reliability is reached, it can be quantified by calculating the inter-rater reliability. This is the percentage of agreement between the two researchers, defined as the number of items that both people categorized in the same way, expressed as a percentage of the total number of items examined. An alternative measure where two raters have been used is Cohen's kappa (κ), which takes into account the possibility that agreement has occurred due to chance (Cohen, 1960).

To illustrate categorization, we present an example derived from a set of studies looking at the use of different navigation aids in an online educational setting (Ursula Armitage, 2004). These studies involved observing users working through some online educational material (about evaluation methods), using the think-aloud technique. The think-aloud protocol was recorded and then transcribed before being analyzed from various perspectives, one of which was to identify usability problems that the participants were having with the online environment known as Nestor Navigator (Zeiliger et al, 1997). An excerpt from the transcription is shown in Figure 8.9.

I'm thinking that it's just a lot of information to absorb from the screen. I just I don't concentrate very well when I'm looking at the screen. I have a very clear idea of what I've read so far . . . but it's because of the headings I know OK this is another kind of evaluation now and before it was about evaluation which wasn't anyone can test and here it's about experts so it's like it's nice that I'm clicking every now and then coz it just sort of organises the thoughts. But it would still be nice to see it on a piece of paper because it's a lot of text to read.

Am I supposed to, just one question, am supposed to say something about what I'm reading and what I think about it the conditions as well or how I feel reading it from the screen, what is the best thing really?

Observer – What you think about the information that you are reading on the screen . . . you don't need to give me comments . . . if you think this bit fits together.

There's so much reference to all those previously said like I'm like I've already forgotten the name of the other evaluation so it said unlike the other evaluation this one like, there really is not much contrast with the other it just says what it is may be . . . so I think I think of . . .

May be it would be nice to have other evaluations listed to see other evaluations you know here, to have the names of other evaluations other evaluations just to, because now when I click previous I have to click it several times so it would be nice to have this navigation, extra links.

Figure 8.9 Excerpt from a transcript of a think-aloud protocol when using an online educational environment. Note the prompt from the observer about half way through

Source: Excerpts reproduced with permission from Ursula Armitage (2004) Navigation and learning in electronic texts. PhD thesis, Centre for HCI Design, City University London.

This excerpt was analyzed using a categorization scheme derived from a set of negative effects of a system on a user (van Rens, 1997) and was iteratively extended to accommodate the specific kinds of interaction observed in these studies. The categorization scheme is shown in Figure 8.10.

1. Interface Problems

1.1. Verbalizations show evidence of dissatisfaction about an aspect of the interface.

11.2. Verbalizations show evidence of confusion/uncertainty about an aspect of the interface.

11.3. Verbalizations show evidence of confusion/surprise at the outcome of an action.

1.4. Verbalizations show evidence of physical discomfort.

1.5. Verbalizations show evidence of fatigue.

1.6. Verbalizations show evidence of difficulty in seeing particular aspects of the interface.

1.7. Verbalizations show evidence that they are having problems achieving a goal that they have set themselves, or the overall task goal.

1.8. Verbalizations show evidence that the user has made an error.

1.9. The participant is unable to recover from error without external help from the experimenter.

1.10. The participant makes a suggestion for redesign of the interface of the electronic texts.

2. Content Problems

2.1. Verbalizations show evidence of dissatisfaction about aspects of the content of the electronic text.

2.2. Verbalizations show evidence of confusion/uncertainty about aspects of the content of the electronic text.

2.3. Verbalizations show evidence of a misunderstanding of the electronic text content (the user may not have noticed this immediately).

2.4. The participant makes a suggestion for re-writing the electronic text content.

Identified problems should be coded as [UP, << problem no. >>].

Figure 8.10 Criteria for identifying usability problems from verbal protocol transcriptions

Source: Excerpts reproduced with permission from Ursula Armitage (2004) Navigation and learning in electronic texts. PhD thesis, Centre for HCI Design, City University London.

This scheme developed and evolved as the transcripts were analyzed. Figure 8.11 shows the excerpt above coded using this categorization scheme. Note that the transcript is divided up using square brackets to indicate which element is being identified as showing a particular usability problem.

[I'm thinking that it's just a lot of information to absorb from the screen. **UP 1.1]**[I just I don't concentrate very well when I'm looking at the screen **UP 1.1]**. I have a very clear idea of what I've read so far . . . [but it's because of the headings **UP 1.1]** I know OK this is another kind of evaluation now and before it was about evaluation which wasn't anyone can test and here it's about experts so it's like it's nice that I'm clicking every now and then coz it just sort of organises the thoughts. **[**But it would still be nice to see it on a piece of paper **UP 1.10] [**because it's a lot of text to read **UP 1.1]**.

Am I supposed to, just one question, am supposed to say something about what I'm reading and what I think about it the conditions as well or how I feel reading it from the screen, what is the best thing really?

Observer – What you think about the information that you are reading on the screen . . . you don't need to give me comments . . . if you think this bit fits together.

[There's so much reference to all those previously said **UP2.1] [** like I'm like I've already forgotten the name of the other evaluation so it said unlike the other evaluation this one like, there really is not much contrast with the other it just says what it is may be . . . so I think I think of . . . **UP 2.2]**

[May be it would be nice to have other evaluations listed to see other evaluations you know here, to have the names of other evaluations other evaluations **UP 1.10]** just to, **[**because now when I click previous I have to click it several times **UP 1.1, 1.7] [**so it would be nice to have this navigation, extra links **UP 1.10]**.

Figure 8.11 The excerpt in Figure 8.9 coded using the categorization scheme in Figure 8.10

Source: Excerpts reproduced with permission from Ursula Armitage (2004) Navigation and learning in electronic texts. PhD thesis, Centre for HCI Design, City University London.

A rigid categorization scheme means that the data is structured only according to the prespecified categories. However, where a significant set of data cannot be categorized, the scheme can be extended. In this case the categorization scheme and the categorization itself develop in parallel, with the scheme evolving as more analysis is done.

Having categorized the data, the results can be used to answer the study goals. In the example above, the study allowed the researchers to be able to quantify the number of usability problems encountered overall by participants, the mean number of problems per participant for each of the test conditions, and the number of unique problems of each type per participant. This also helped to identify patterns of behavior and recurring problems. Having the think-aloud protocol meant that the overall view of the usability problems could take context into account.

Activity 8.2

The following is another think-aloud extract from the same study. Using the categorization scheme in Figure 8.10, code this extract for usability problems. It is useful to put brackets around the complete element of the extract that you are coding.

> Well, looking at the map, again there's no obvious start point, there should be something highlighted that says 'start here.'
>
> Ok, the next keyword that's highlighted is evaluating, but I'm not sure that's where I want to go straight away, so I'm just going to go back to the introduction.
>
> Yeah, so I probably want to read about usability problems before I start looking at evaluation. So, I, yeah. I would have thought that the links in each one of the pages would take you to the next logical point, but my logic might be different to other people's. Just going to go and have a look at usability problems.
>
> Ok, again I'm going to flip back to the introduction. I'm just thinking if I was going to do this myself I would still have a link back to the introduction, but I would take people through the logical sequence of each one of these bits that fans out, rather than expecting them to go back all the time.
>
> Going back . . . to the introduction. Look at the types. Observation, didn't really want to go there. What's this bit [pointing to Types of UE on map]? Going straight to types of . . .
>
> Ok, right, yeah, I've already been there before. We've already looked at usability problems, yep that's ok, so we'll have a look at these references.
>
> I clicked on the map rather than going back via introduction, to be honest I get fed up going back to introduction all the time.

Comment

Show/Hide

The example above used a form of content analysis (Krippendorff, 2013). Content analysis typically involves categorizing the data and then studying the frequency of category occurrences. The technique may be used for any text, where ‘text’ refers to a range of media including video, newspapers, adverts, and so on. For example, Blythe and Cairns (2009) analyzed 100 videos from a YouTube search by relevance for ‘iPhone 3G’ using content analysis. They categorized the videos into seven categories: review, reportage, ‘unboxing,’ demonstration, satire, advertisement, and vlog commentaries (e.g. complaints about queues).

Another way of analyzing a transcript is to use discourse analysis. Discourse analysis focuses on the dialog, i.e. the meaning of what is said, and how words are used to convey meaning. Discourse analysis is strongly interpretive, pays great attention to context, and views language not only as reflecting psychological and social aspects but also as constructing them (Coyle, 1995). An underlying assumption of discourse analysis is that there is no objective scientific truth. Language is a form of social reality that is open to interpretation from different perspectives. In this sense, the underlying philosophy of discourse analysis is similar to that of ethnography. Language is viewed as a constructive tool and discourse analysis provides a way of focusing on how people use language to construct versions of their worlds (Fiske, 1994).

Small changes in wording can change meaning, as the following excerpts indicate (Coyle, 1995):

> **Discourse analysis is what you do when you are saying that you are doing discourse analysis . . .**
>
> **According to Coyle, discourse analysis is what you do when you are saying that you are doing discourse analysis . . .**

By adding just three words, ‘According to Coyle,’ the sense of authority changes, depending on what the reader knows about Coyle's work and reputation.

Conversation analysis is a very fine-grained form of discourse analysis (Jupp, 2006). In conversation analysis the semantics of the discourse are examined in fine detail, and the focus is on how conversations are conducted. This technique is used in sociological studies and examines how conversations start, how turn-taking is structured, and other rules of conversation. This analysis technique has been used to analyze interactions in a range of settings, and has influenced designers’ understanding about users’ needs in these environments. It can also be used to compare conversations that take place through different media, e.g. face-to-face versus email.

8.4.3 Looking for Critical Incidents

Data gathering sessions can often result in a lot of data. Analyzing all of that data in any detail is very time-consuming, and often not necessary. We have already suggested that themes, patterns, and categories can be used to identify areas where detailed analysis is appropriate. Another approach is to use the critical incident technique (Grison et al, 2013).

The critical incident technique is a set of principles that emerged from work carried out in the United States Army Air Forces where the goal was to identify the critical requirements of good and bad performance by pilots (Flanagan, 1954). It has two basic principles: "(a) reporting facts regarding behavior is preferable to the collection of interpretations, ratings, and opinions based on general impressions; (b) reporting should be limited to those behaviors which, according to competent observers, make a significant contribution to the activity" (Flanagan, 1954, p. 355). In the interaction design context, the use of well-planned observation sessions, as discussed in Chapter 7, satisfies the first principle. The second principle is referring to critical incidents, i.e. incidents that are significant or pivotal to the activity being observed, in either a desirable or an undesirable way.

In interaction design, critical incident analysis has been used in a variety of ways, but the main focus is to identify specific incidents that are significant, and then to focus on these and analyze them in detail, using the rest of the data collected as context to inform interpretation. These may be identified by the users during a retrospective discussion of a recent event, or by an observer either through studying video footage, or in real time. For example, in an evaluation study, a critical incident may be signaled by times when users were obviously stuck – usually marked by a comment, silence, looks of puzzlement, and so on.

In a study by Curzon et al (2002), they identified a set of critical incidents through field trials of an in-car navigation device. One example incident in this context was "On one journey, the system gave directions to turn right when the destination was to the left. Its route was to go round the block to go in the other direction. A car following ignored this turn and went the more obvious way, arriving first." In another study, Grison et al (2013) used the critical incident technique to investigate specific factors that influence travelers' choices of transport mode in Paris in order to adapt new tools and services for mobility, such as dynamic route planners. Participants were asked to report on positive and negative real events they experienced in the context of their route to work or study, and whether they regretted or were satisfied with this choice of transport. Their findings included that contextual factors have a great influence on choice, that people were more likely to choose an alternative route to return home than when setting out, and that emotional state is important when planning a route.

Link to read more on critical incident analysis in HCI, at www.usabilitynet.org

It is common practice to employ more than one complementary data analysis approach. For example, following a critical incident analysis, themes may be identified around the circumstances that caused the incident to occur, and then discourse analysis may be conducted to understand the detail. Analyzing video material frequently employs several methods and can be challenging (see Box 8.4).

Activity 8.3

Set yourself or a friend the task of identifying the next available theater or cinema performance you'd like to attend in your local area. As you perform this task, or watch your friend do it, make a note of critical incidents associated with the activity. Remember that a critical incident may be a positive or a negative event.

Comment

Show/Hide

BOX 8.4

Analyzing Video Material

A good way to start a video analysis is to watch what has been recorded all the way through while writing a high-level narrative of what happens, noting down where in the video there are any potentially interesting events. How you decide which is an interesting event will depend on what is being observed. For example, if you are studying the interruptions that occur in an open plan office, you would include each time a person breaks off from an ongoing activity, e.g. phone rings, someone walks into their cubicle, email arrives. If it is a study of how pairs of students use a collaborative learning tool then activities such as turn-taking, sharing of input device/s, speaking over one another, and fighting over shared objects would be appropriate to record.

Chronological and video times are used to index events. These may not be the same, since recordings can run at different speeds to real time. Labels for certain routine events are also used, e.g. lunchtime, coffee break, staff meeting, doctor's round. Spreadsheets are used to record the classification and description of events, together with annotations and notes of how the events began, how they unfold, and how they end.

Video can be augmented with captured screens or logged data of people's interactions with a computer display. There are various logging and screen capture tools available for this purpose that enable you to play back the interactions as a movie, showing screen objects being opened, moved, selected, and so on. These can then be played in parallel with the video to provide different perspectives on the talk, physical interactions, and the system's responses that occur. Having a combination of data streams can enable more detailed and fine-grained patterns of behavior to be interpreted (Heath et al, 2010).■

8.5 Tools to Support Data Analysis

It would be possible to analyze even large data sets using only manual techniques, however most people would agree that it is quicker, easier, and more accurate to use a software tool of some kind, particularly for ‘big data.’ We introduced the idea of using a simple spreadsheet application in Section 8.3, but there are other more sophisticated tools available – some of which support the organization and manipulation of the data, and some of which are focused on performing statistical tests. For example, Box 8.5 discusses the analysis and presentation of thousands of data points captured in a study on digital library usage.

New tools are developed and existing ones are enhanced on a regular basis, so we do not attempt to provide a comprehensive survey of this area. Instead, we discuss the kind of support available, and describe briefly some of the more popular tools used in interaction design.

Tools to support the organization and manipulation of data include facilities for categorization, theme-based analysis, and quantitative analysis. These typically provide facilities to associate labels (categories, themes, and so on) with sections of data, search the data for key words or phrases, investigate the relationships between different themes or categories, and help to develop the coding scheme further; some tools can also generate graphical representations. In addition, some provide help with techniques such as content analysis and sometimes very sophisticated mechanisms to show the occurrence and co-occurrence of words or phrases. In addition, searching, coding, project management, writing and annotating, and report generation facilities are common.

More detail regarding software tools to support the analysis of qualitative data can be found through the CAQDAS Networking Project, based at the University of Surrey (http://caqdas.soc.surrey.ac.uk/).

Two well-known tools that support some of these data analysis activities are Nvivo and Atlas.ti. Nvivo, for example, supports the annotation and coding of data including PDF documents, photos, and video and audio files. Using Nvivo, field notes can be searched for key words or phrases to support coding or content analysis; codes and data can be explored, merged, and manipulated in several ways. The information can also be printed in a variety of forms such as a list of every occasion a word or phrase is used in the data, and a tree structure showing the relationships between codes. Like all software packages, Nvivo has advantages and disadvantages, but it is particularly powerful for handling very large sets of data and it can generate output for statistical packages such as SPSS.

SPSS (Statistical Package for the Social Sciences) is one of the more popular quantitative analysis packages that supports the use of statistical tests. It is a sophisticated package offering a wide range of statistical tests such as frequency distributions, rank correlations (to determine statistical significance), regression analysis, and cluster analysis. SPSS assumes that the user knows and understands statistical analysis.

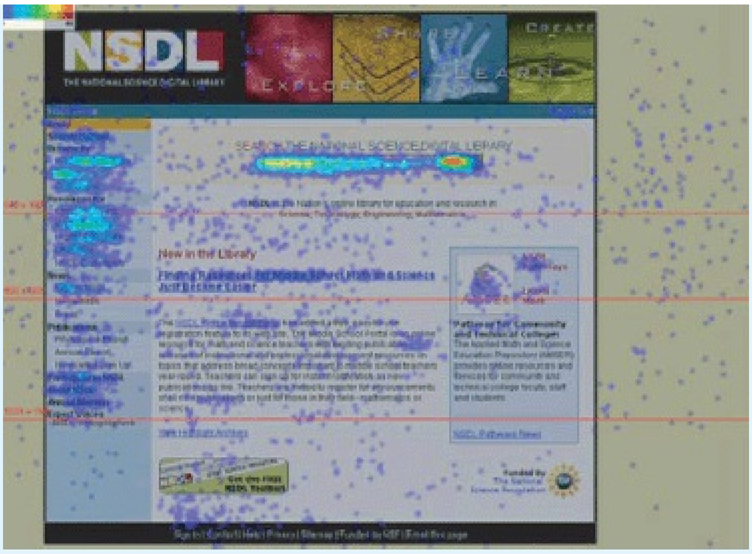
BOX 8.5

Web Analytics for Digital Libraries – Presenting Session Length Data

Khoo et al (2008) discuss four case studies focused on different digital libraries, each of which investigated the use of session length as a useful metric in their context. Associated with session length are various terms including a visit (the sequential viewing of one or more pages from the same IP address), page views (number of times a page is accessed during a visit), and hit (request for a page or page element). One of the problems faced by users of web analytic software is that these elements are measured differently by different tools – hence the need to investigate how best to make use of session length data.

The four digital libraries presented in their article are the Library of Congress, Teachers' Domain of the National Science Digital Library (NSDL), the NSDL itself, and The Instructional Architect. Each of these had different technical and organizational concerns but several common themes emerged from the investigations including the need to identify suitable tools and reconcile differences in metric definitions, and the need to triangulate web metrics with other research.

Of particular interest are the different visualizations used to present the web metrics data for interpretation and discussion (in the article) and other reports mentioned as influencing the conclusions. Data from the NSDL was presented as a heat map (Figure 8.12a) and as timeframes (Figure 8.12b); data from the Library of Congress was presented in a statistical table (Figure 8.12c) and data from the Teachers Domain was presented as a comparison between several months' figures, on one graph (Figure 8.12d). The Instructional Architect investigations used Google Analytics software. The researchers used a range of reports including users' geographical location, where users were before arriving at the library site, the ratio of student login paths to teacher login paths, and session time (as time on site and pages viewed per visit).■



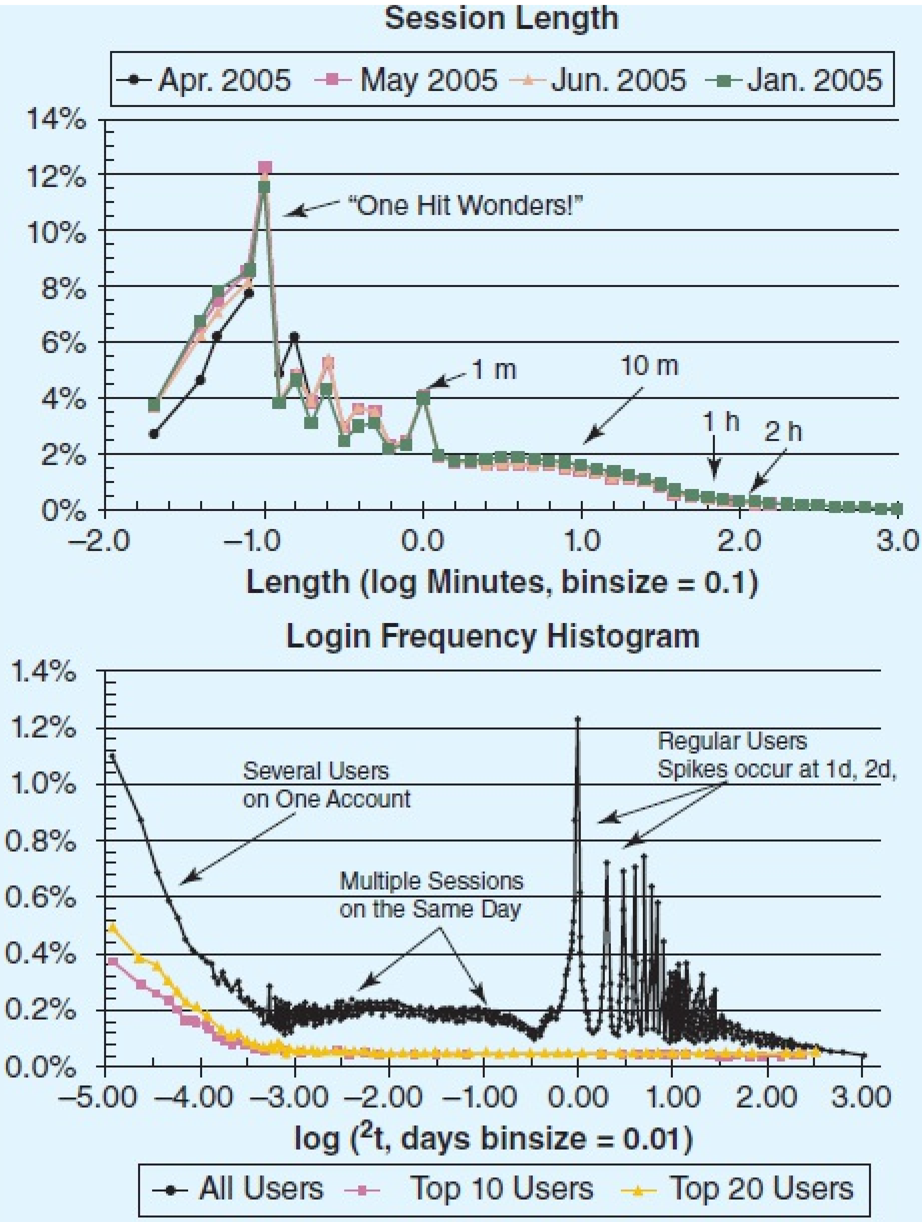
(a)

(b)

Filecode: 18ME44 Approximate words: 900 Sum of views: 13931

Minutes	.25 - .5	.5 - 1	1 - 3	3 - 5	5 - 10	10 - 15	15 - 20
Views	1354	1303	1580	792	2442	3646	2814
Std. Dev.		147.7					
Hours		77.4					
Std. Dev.			399.7				
Hours			121.8				
Std. Dev.				825.3			
Hours				410.7			
Std. Dev.					1432.8		
Hours					1117.6		
Std. Dev.						616.5	
Hours						1885.6	

(c)



(d)

Figure 8.12 Different visualizations used to present web analytical data from three different digital library case studies: (a) heat map from NSDL; (b) weekly traffic timeline from NSDL; (c) standard deviation and time spent on a page data from the Library of Congress; and (d) session length data of four different months from Teachers' Domain (NSDL)

Source: Khoo, M., Pagano, J., Washington, A. L., Recker, M., Palmer, B., and Donahue, R. A. (2008) Using web metrics to analyze digital libraries. Proceedings of Joint Conference on Digital Libraries, Pittsburgh, June 16–20. ©2008 Association for Computing Machinery, Inc. Reprinted by permission.

8.6 Using Theoretical Frameworks

Structuring the analysis of qualitative data around a theoretical framework can lead to additional insights that go beyond the results found from the simple techniques introduced earlier. However, these frameworks are quite sophisticated and using them requires investment to make sure that the framework is understood and applied appropriately. This section discusses three frameworks that are commonly used in interaction design to structure the analysis of data gathered in the field, such as observational and interview data: grounded theory, distributed cognition, and activity theory.

8.6.1 Grounded Theory

Grounded theory is an approach to qualitative data analysis that aims to develop theory from the systematic analysis and interpretation of empirical data, i.e. the theory derived is grounded in the data. In this respect it is a bottom-up development of theory since the data is needed to develop the theory. This contrasts with some types of analysis in which the theory (or previous published research) provides the categories used for the analysis of the data. The approach was originally developed by Glaser and Strauss (1967) and since has been adopted by several researchers, with some adaptations to different circumstances. In particular, Glaser and Strauss have individually (and with others) developed grounded theory in slightly different ways, but the aim of this approach remains the same. Glaser (1992) provides further information about the differences and areas of controversy.

The aim of grounded theory is to develop a theory that fits a set of collected data. In this context, theory is: "a set of well-developed concepts related through statements of relationship, which together constitute an integrated framework that can be used to explain or predict phenomena" (Strauss and Corbin, 1998). Development of a 'grounded' theory progresses through alternating data collection and data analysis: first data is collected and analyzed to identify categories, then that analysis may lead to further data collection and analysis to extend and refine the categories and so on; during this cycle, parts of the data may be reanalyzed in more detail. Data gathering and subsequent analysis are hence driven by the emerging theory. This approach continues until no new insights emerge and the theory is well-developed. During this process, the researcher needs to maintain a balance between objectivity and sensitivity. Objectivity is needed to maintain accurate and impartial interpretation of events; sensitivity is required to notice the subtleties in the data and identify relationships between concepts.

The thrust of the analysis undertaken is to identify and define the properties and dimensions of relevant categories and then to use these as the basis for constructing a theory. Category identification and definition is achieved by coding the data, i.e. marking it up according to the emerging categories. According to Corbin and Strauss (2014), this coding has three aspects, which are iteratively performed through the cycle of data collection and analysis:

1. Open coding is the process through which categories, their properties, and dimensions are discovered in the data. This process is similar to our discussion of categorization above, including the question of granularity of coding (at the word, line, sentence, conversation level, and so on).
2. Axial coding is the process of systematically fleshing out categories and relating them to their subcategories.
3. Selective coding is the process of refining and integrating categories to form a larger theoretical scheme. The categories are organized around one central category that forms the backbone of the theory. Initially, the theory will contain only an outline of the categories but as more data is collected, they are refined and developed further.

Early books on grounded theory say little about what data collection techniques should be used, but focus instead on the analysis. Some later books place more emphasis on data collection. For example, Charmaz (2014) discusses interviewing techniques, and collection and analysis of documents for grounded theory analysis. When analyzing data, Corbin and Strauss (2014) encourage the use of written records of analysis and diagrammatic representations of categories (which they call memos and diagrams). These memos and diagrams evolve as data analysis progresses. Some researchers also look to digital tools such as spreadsheets and diagramming tools, but many like to develop their own physical code books such as the one Rotman et al (2014) constructed in a study to understand the motivations of citizens to contribute to citizen science projects. The data that she analyzed was from in-depth semi-structured interviews of 33 citizen scientists and 11 scientists from the USA, India, and Costa Rica (see Figure 8.13).

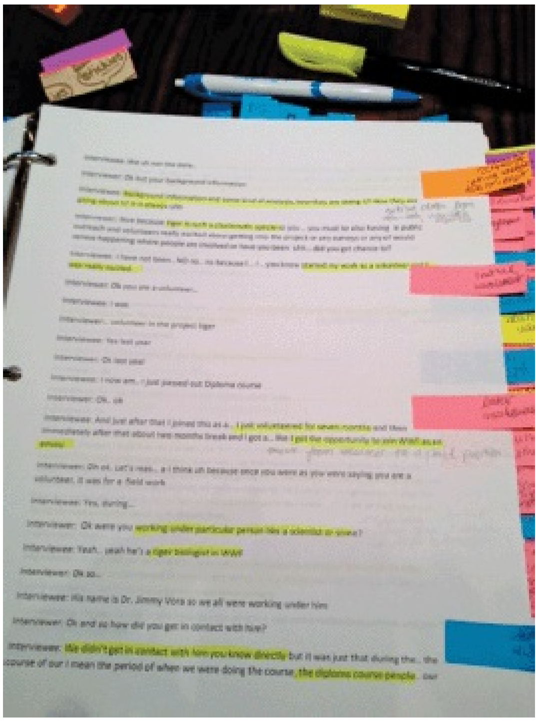


Figure 8.13 Code book used in a grounded theory analysis of citizens’ motivations to contribute to citizen science

Source: Rotman, D. et al (2014). Does motivation in citizen science change with time and culture? In Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing (CSCW Companion ‘14). ACM, New York, NY, USA, 229–232. ©2014 Association for Computing Machinery, Inc. Reprinted by permission.

The following analytic tools are used to help stimulate the analyst's thinking and identify and characterize relevant categories:

- The use of questioning (not questioning your participants, but questioning the data): questions can help an analyst to generate ideas or consider different ways of looking at the data. It can be useful to ask questions when analysis appears to be in a rut.
- Analysis of a word, phrase, or sentence: considering in detail the meaning of an utterance can also help to trigger different perspectives on the data.
- Further analysis through comparisons: comparisons may be made between objects or between abstract categories. In either case, comparing one with the other brings alternative interpretations. Sharp et al (2005) take this idea further and use metaphor as an analysis technique with qualitative data.

One thing that distinguishes a grounded theory approach to data gathering and analysis from ethnographic approaches is that researchers are encouraged to draw on their own theoretical backgrounds to help inform the study, provided that they are alert to the possibility of unintentional bias.

Performing a Grounded Theory Analysis

Sarker et al (2001) used the grounded theory approach to develop a model of collaboration in virtual teams. The virtual teams used in the study were made up of students from two universities – one in Canada and one in the United States of America. Each team consisted of four to five members from each university. Each team was given the task of studying a business systems problem, producing a systems design for it, and developing a working prototype. The projects themselves lasted about 14 weeks and a total of 12 teams participated in the study. The team members could communicate directly with each other using various technologies such as email, videoconferencing, telephone, and fax. The main communication channel, however, was Webboard, a collaborative message board tool supporting threaded discussions, email discussions, chat rooms, instant messaging, calendar, whiteboard, blogging, and so on. Using Webboard meant that communication was more public, and could be recorded more easily.

All communication data through Webboard was automatically saved, minutes of any videoconferences, telephone calls, or Internet chat sessions undertaken were posted on Webboard, and the team members were also asked to provide any additional emails they received or sent to other team members. In addition to this data, the team's final project reports, individual team members' reflections on the lessons learned through the project, feedback on fellow team members' performance, and comments on the virtual project itself were all collected and used as data for the study.

As soon as the teams were formed, informal data analysis began and two of the researchers became participant observers in the project teams, developing sensitivity to the project and its goals. They also began to reflect on their own backgrounds to see what theoretical frameworks they could draw on.

Open coding.

This was done initially on a line-by-line basis, but later coding was done at the message level, while other documents such as reports and reflections were coded at document level. Over 200 categories were generated, and as these were refined, some informal axial coding was also done. Table 8.5 shows two messages posted at the beginning of a project, and illustrates how these messages were coded during the open coding process.

Table 8.5 An illustration of open coding

Message	Post date, week #, time	Sample codes generated (underlined) and notes
Hi there in UB, I'm Henry. I just wanted to say hello and provide you with the rest of our group members' email address. [Names and email addresses] Well, I guess we'll see each other on Saturday at the videoconference.	1/22/98, week 1, 1:41:52 PM	1. Leadership – initiative to represent. 2. Establishing team's co-presence on the Internet. 3. Implying preference for communication technology (email). 4. Implying technology (VC) can bridge the time and space gap.
Hello UB. Just letting you know that you are free to email us anytime. I might be getting an ICQ account going so that if any of you are into real-time chat and wish to communicate that way, it might be something to try . . .	1/26/98, week 1, 2:56:37 PM	1. UB members' identity viewed at an aggregate level (as in msg. #1). 2. Collapsing/bridging across time boundaries. 3. Invitation. 4. Implying preference for communication technology. 5. Properties of communication technology/medium (real-time, synchronous?).

		6. Novelty of technology, recognizing the need to try/explore.

Sarker et al note that codes emerged and continued to be refined over the life of the project. Also, a significant number of the codes that were ultimately used in the theory building were recurrent; for example, preference for technology and time gaps/boundaries. Finally, some of the key categories were identified when considering messages as one unit and looking at comparable strips in other data segments.

Through constant comparison of data across categories, the names of categories were refined, merged, and changed over time.

Axial coding.

Sarker et al found the suggestions in Corbin and Strauss about how to relate subcategories too constraining. They instead used a two-step process for axial coding:

1. The major categories, e.g. technology, norms, social practices, stages of team development, and frames of reference, were hierarchically related to subcategories. For example, the category technology was linked to the subcategories purpose of technology, nature of ownership, accessibility (by time, location, cost, awareness), future potential, degree of novelty, and interconnectedness. At the next level, purpose of technology was linked to information sharing, triggering, and so on (see Figure 8.14). During this process, the researchers returned to open coding and refined categories further.
2. For each major category, researchers created a description (called a memo) that attempted to integrate as many of the categories and subcategories as possible. These memos also evolved as analysis progressed. Figure 8.15 contains an excerpt from an early draft memo for the technology category.

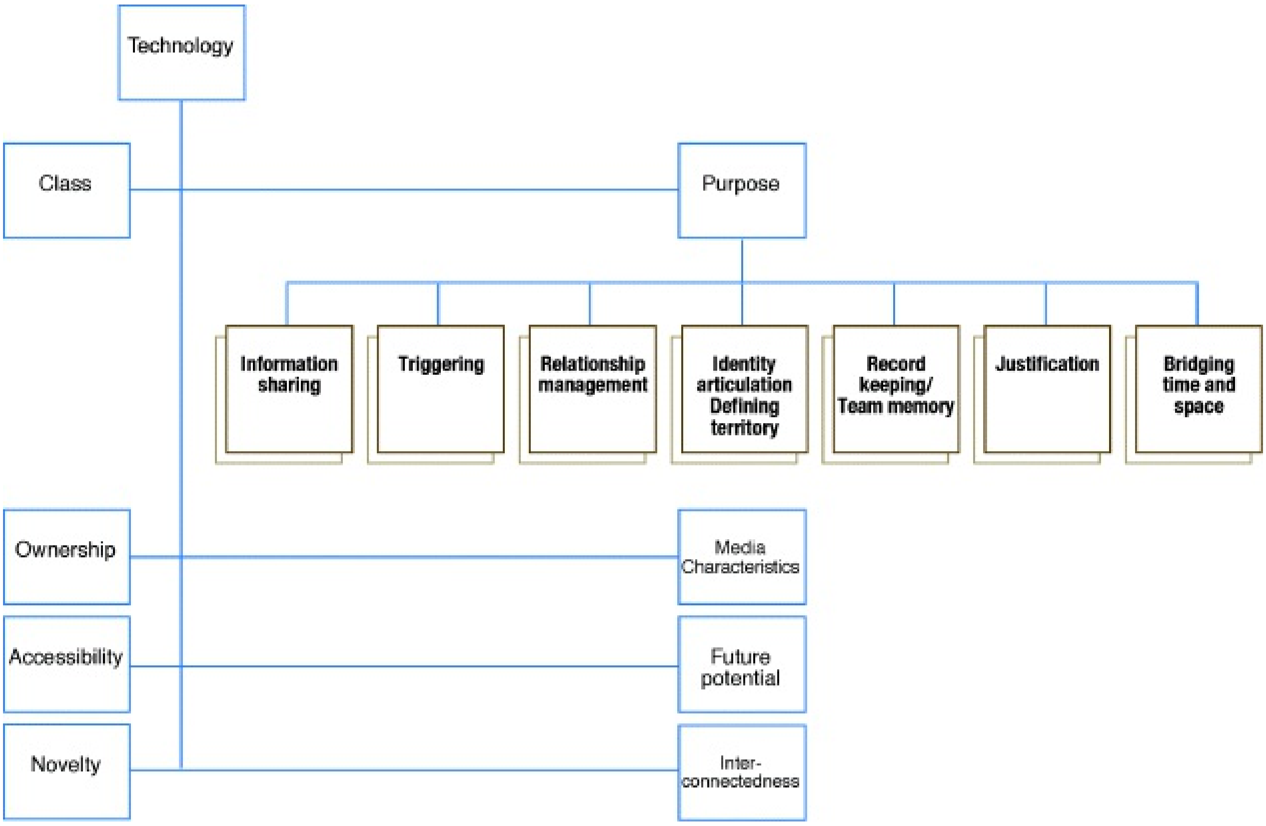


Figure 8.14 Axial coding for the technology category

Source: S. Sarker, F. Lau and S. Sahay (2001): "Using an adapted grounded theory approach for inductivetheory building about virtual team development". The Data Base for Advances in Information Systems, 32(1), pp. 38–56 ©2001 Association for Computing Machinery, Inc. Reprinted by permission.

Collaboration across time and space requires mediation by technology for both symbolic and substantive purposes. Substantive purposes include sharing information, record-keeping, managing relationships, pacing and triggering of activities in collaboration. Some symbolic uses of technology involve the articulation of the self and group identity and legitimizing different courses of action by appealing to the use of technology.

Different classes of technology provide different capabilities, some of them different to the features of technology as defined from the designers' or the implementers' point of view. For example, we wanted Webboard to be a public record . . . students have extended this use by creating a local enclave for information exchange with local members in a domain traditionally thought of as being public. The Webboard has also become a project archive, conserving team memory through the documentation of agendas, minutes, project steps, and deliverables.

Figure 8.15 An excerpt from an early draft of an integrative memo for the technology category

Source: Excerpt from S. Sarker, F. Lau and S. Sahay (2001): "Using an adapted grounded theory approach for inductive theory building about virtual team development". The Data Base for Advances in Information Systems, 32(1), pp. 38–56.

Selective coding.

This stage of coding involves linking the categories and subcategories into a theory, and as theory building is quite complex, we only present an overview of the process here.

Sarker et al wanted to develop a theory of virtual teams, and so they used two approaches from their background to help them. One of these approaches (Couch, 1996) emphasizes the concepts that a theory of human conduct must use. The other focuses on social structure (Giddens, 1984). Using these two approaches, the category 'stages of team development' was chosen as the core category for the grounded theory to be built from this data set, and other categories were linked around it. This theory was further elaborated upon through discussions and reading of theory, and evolved into a theory of how virtual teams develop over time. More details can be found in Sarker and Sahay (2003).

8.6.2 Distributed Cognition

We introduced the distributed cognition approach in Chapter 3, as a theoretical account of the distributed nature of cognitive phenomena across individuals, artifacts, and internal and external representations (Hutchins, 1995).

Typically, a distributed cognition analysis results in an event-driven description which emphasizes information and its propagation through the cognitive system under study. The cognitive system under study might be one person's use of a computational tool, such as a calculator, two people's joint activities when designing the layout for the front page of a newspaper, using a shared authoring tool, or, more widely, a large team of software developers, examining how they coordinate their work with one another, using a variety of mediating artifacts, such as schedules, clocks, to-do lists, and shared files.

The granularity of analysis varies depending on the activities and cognitive system being observed and the research or design questions being asked. For example, if the goal is to examine how a team of pilots fly a plane – with a view to improving communication between them – then the focus will be on the interactions and communications that take place between them and their instruments, at a fine level of granularity. If the goal is to understand how pilots learn how to fly – with a view to developing new training materials – then the focus will be at a coarser grain of analysis, taking into account the cultural, historical, and learning aspects involved in becoming a pilot.

The description produced may cover a period of a day, an hour, or only minutes, depending on the study's focus. For the longer periods, verbal descriptions are primarily used. For the shorter periods, micro-level analyses of the cognitive processes are meticulously plotted using diagrammatic forms and other graphical representations. The rationale for performing the finer levels of analysis is to reveal practices and discrepancies that would go unnoticed using coarser grain analysis, but which reveal themselves as critical to the work activity.

Ed Hutchins (1995) emphasizes that an important part of doing a distributed cognition analysis is to have a deep understanding of the work domain that is being studied. He recommends, where possible, that the investigators learn the trade under study. This can take a team of researchers several months and even years to accomplish and in most cases this is impractical for a research or design team to do.

Alternatively, it is possible to spend a few weeks immersed in the culture and setting of a specific team to learn enough about the organization and its work practices to conduct a focused analysis of a particular cognitive system. For example, I spent six weeks with an engineering team, where I was able to learn enough about their work practice to gain a good understanding of how they worked together on projects, how they coordinated their work with each other, and how the technologies that were used mediated their work activities. I was then able to document and analyze a number of problems they were experiencing through the introduction of new networking technology. Using the distributed cognition framework, I described how seemingly simple communication problems led to large delays and recommended how the situation could be improved (Rogers, 1993, 1994).

More recently, distributed cognition has been applied to studying medical teams. For example, Rajkomar and Blandford (2012) examined how healthcare technologies are used; specifically they examined the use of infusion pumps by nurses in an intensive care unit (ICU). They gathered data through ethnographic observations and interviews, which they analysed by constructing representational models that focused on information flows, physical layouts, social structures, and artifacts. They note that “the findings showed that there was significant distribution of cognition in the ICU: socially, among nurses; physically, through the material environment; and through technological artefacts.” Based on the results of this study, they were able to suggest changes that would improve the safety and efficiency of the nurses’ interactions with the infusion technology.

Performing a Distributed Cognition Analysis

It should be stressed that there is not one single way of doing a distributed cognition analysis, nor is there an off-the-shelf manual that can be followed. A good way to begin analyzing and interpreting the data collected is to describe the official work practices, in terms of the routines and procedures followed, and the work-arounds that teams develop when coping with the various demands placed upon them at different times during their work. In so doing, any breakdowns, incidents, or unusual happenings should be highlighted, especially where it was discovered that excessive time was being spent doing something, errors were made using a system, or a piece of information was passed on incorrectly to someone else or misheard. While writing these observations down it is good to start posing specific research questions related to them (e.g. ‘Why did X not let Y know the printer was broken when he came back from his break?') and to contemplate further (e.g. ‘Was it a communication failure, a problem with being overloaded at the time, or a technology problem?').

It is at this point that knowledge of the theory of distributed cognition can help in interpreting and representing the observations of a work setting (see Chapter 3 and Hutchins, 1995). It provides an analytic framework and a set of concepts to describe what is happening at a higher level of abstraction.

Problems can be described in terms of the communication pathways that are being hindered or the breakdowns arising due to information not propagating effectively from one representational state to another (see Box 8.6). The framework can reveal where information is being distorted, resulting in poor communication or inefficiency. Conversely, it can show when different technologies and the representations displayed via them are effective at mediating certain work activities and how well they are coordinated.

Performing a detailed distributed cognition analysis enables researchers and designers to explore the trade-offs and likely outcomes of potential solutions and in so doing suggest a more grounded set of cognitive requirements, e.g. types of information resources that are considered suitable for specific kinds of activities, and those that could be dealt with by an automated system. Clearly, such a painstaking level of analysis and the expertise required in the interpretation are very costly. In the commercial world, where deadlines and budgets are always looming, it is unlikely to be practical. However, in large-scale and safety-critical projects, where more time and resources are available, it can be a valuable analytic tool to use.

BOX 8.6

Distributed Cognition Concepts

A distributed cognition analysis involves producing a detailed description of the domain area at varying levels of granularity. At the micro-level, a small set of cognitive terms are used to depict the representations employed in a cognitive activity and the processes acting upon them. The terms are intended to steer the analysis towards conceptualizing problems in terms of distributed information and representations. This level of description can also directly lead to recommendations, suggesting how to change or redesign an aspect of the cognitive system, such as a display or a socially mediated practice. The main terms used are:

- The cognitive system – the interactions among people, the artifacts they use, and the environment they are working in.
- The communicative pathways – the channels by which information is passed between people (e.g. phone, email, physical gesture).
- Propagation of representational states – how information is transformed across different media. Media refers to external artifacts (e.g. instruments, maps, paper notes) and internal representations (e.g. human memory). These can be socially mediated (e.g. passing on a message verbally), technologically mediated (e.g. pressing a key on a computer), or mentally mediated (e.g. reading the time on a clock). ■

Furniss and Blandford (2006) applied distributed cognition to an emergency medical dispatch setting (ambulance control). They identified 22 principles underlying the literature on distributed cognition, and used diagrams from Contextual Design (Beyer and Holtzblatt, 1998) to capture relevant aspects of activity they observed. The analysis resulted in suggestions for improving the dispatch room operation. The resulting method for applying distributed cognition, called DiCOT (distributed cognition for teamwork, Blandford and Furniss (2006)), has subsequently been used to understand software team interactions (Sharp and Robinson, 2008), mobile healthcare settings (McKnight and Doherty, 2008), and the use of infusion pumps by nurses (Rajkomar and Blandford, 2012).

8.6.3 Activity Theory

Activity theory (AT) is a product of Soviet psychology that explains human behavior in terms of our practical activity with the world. It originated as part of the attempt to produce a Marxist psychology, an enterprise usually associated with Vygotsky (e.g. 1926/1962) and later Leontiev (e.g. 1978, 1989). In the last 20–30 years, versions of AT have become popular elsewhere, particularly in Scandinavia and Germany, and interest is now growing in the USA and UK. The newer versions of AT have been popular in research investigating applied problems, particularly those to do with work, technology, and education.

Activity theory provides a framework that focuses analysis around the concept of an activity and helps to identify tensions between the different elements of the system. For example, in what is now viewed as a classic study of the use of AT in HCI, Mackay et al (2000) analyzed a 4-minute excerpt from a video of users working with a new software tool. They identified 19 shifts in attention between different parts of the tool interface and the task at hand. In fact, some users spent so much time engaged in these shifts that they lost track of their original task. Using the theory helped evaluators to focus on relevant incidents.

AT outlines two key framings: the individual model that constitutes an activity and one that models the mediating role of artifacts.

The Individual Model

AT models activities in a hierarchical way. At the bottom level are operations, routinized behaviors that require little conscious attention (e.g. rapid typing). At an intermediate level are actions, behavior that is characterized by conscious planning, e.g. producing a glossary. The top level is the activity, and that provides a minimum meaningful context for understanding the individual actions, e.g. writing an essay (see Figure 8.16). There may be many different operations capable of fulfilling an action, and many actions capable of serving the same activity.

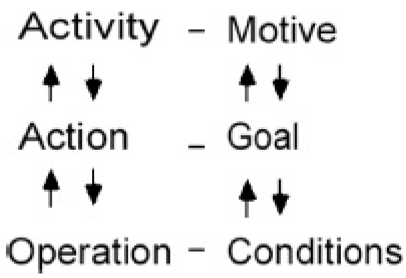


Figure 8.16 The original activity theory model

Activities can be identified on the basis of the motives that elicit them, actions on the basis of conscious goals that guide them, and operations by the conditions necessary to attain the goals. However, there is an intimate and fluid link between levels. Actions can become operations as they become more automatic and operations can become actions when an operation encounters an obstacle, thus requiring conscious planning. Similarly there is no strict demarcation between action and activity. If the motive changes, then an activity can become an action. It is also important to realize that activities are not self-contained. Activities relate to others while actions may be part of different activities, and so on.

The Role of Artifacts

Artifacts can be physical, such as a book or a stone, or they can be abstract, such as a system of symbols or a set of rules. Physical artifacts have physical properties that cause humans to respond to them as direct objects to be acted upon. They also embody a set of social practices, their design reflecting a history of particular use. Leontiev describes the process of learning what these inherent properties are as one of appropriation, signifying the active nature of the learning that is needed. The kind of learning involved is one of identifying and participating in the activity appropriate to the artifact. Consider an infant learning to feed with a spoon. Leontiev (1981) observed that, at first, the infant carries the spoon to its mouth as though it were handling any other object, not considering the need to hold it horizontal. Over time, with adult guidance, the spoon is shaped in the way it is because of the social practice – the activity – of feeding. In turn, the infant's task is to learn that relationship – to discover what practice(s) the object embodies. By contrast a spoon dropped into the cage of a mouse, say, will only ever have the status of just another physical object – no different from that of a stone.

The idea of abstract artifacts follows from the idea of mediation, i.e. a fundamental characteristic of human development is the change from a direct mode of acting on the world to one that is mediated by something else. In AT, the artifacts involved in an activity mediate between the elements of it.

AT also emphasizes the social context of an activity. Even when apparently working alone, an individual is still engaged in activities that are given meaning by a wider set of practices.

The classic view of an activity has a subject (who performs the activity) and an object (on which the activity is performed). Engeström (e.g. 1999) and Nardi and Kaptelinin (2012) have widened the focus from the individual triangle of a single activity (subject, activity, and object) to include supra-individual concepts – tools, rules, community, and division of labor. By tool is meant the artifacts, signs, and means that mediate the subject and object; by community is meant those who share the same object; by rules is meant a set of agreed conventions and policies covering what it means to be a member of that community (set by laws, parents, managers, boards, and so forth); and by division of labor is meant the primary means of classifying the labor in a workplace (e.g. manager, engineer, receptionist). The extended versions allow consideration of networks of interrelated activities – forming an activity system (see Figure 8.17).

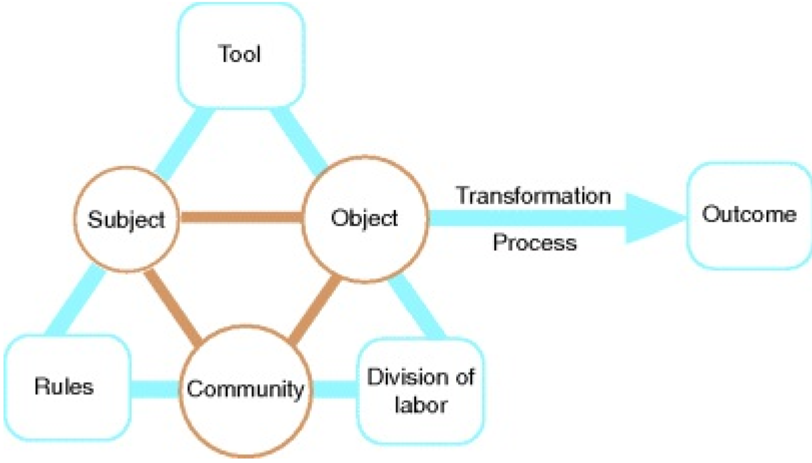


Figure 8.17 Engeström's (1999) activity system model. The tool element is sometimes referred to as the mediating artifact

Source: Reproduced from Engeström, Y. (1999) Perspectives on Activity Theory, CUP.

Performing an Analysis Driven by Activity Theory

AT does not present a clear methodological prescription for the description or analysis of behavior as a set of procedures to be followed. The model shown in Figure 8.16 is the main framework that is used to describe levels within an activity. This means that identifying elements will be highly dependent on individual interpretation. Christiansen (1996, p. 177) summarizes: “Methodologically . . . the ideal data for an application of AT consist of longitudinal ethnographic observation, interviews and discussion in real-life settings, supplemented by experiments.” She continues that you “cannot interview people directly through rote questions but must interpret their actions and opinions after some careful reflection,” which is a difficult process. Nevertheless, the original and later versions of the AT framework have become popular amongst researchers and some practitioners as a way of identifying needs for new tools and to analyze problems that are present in a work or other setting. For example, Saguna and Chakraborty (2013) point out that AT has much to offer in the development of mobile applications where analyzing users’ activities and particularly changes in activities can inform design. However, they also point out that there is a lack of a unifying theoretical framework which fully addresses all aspects of the activity and activity domain.

One of the biggest problems with doing an AT analysis is working out when something should be described as a top-level activity and when something is better described as a lower-level action.

Performing an AT analysis enables researchers and designers to identify the tensions in a workplace leading to specific needs for new technological tools; it is also used in evaluation studies. It can be difficult, however, getting to grips with the concepts and being able to determine how to label the points raised in the interviews. Expertise and a good background in the Soviet version of activity theory are recommended to become competent in AT. Similar to the distributed cognition approach in the commercial world, where deadlines and budgets are always looming, it is unlikely to be practical. Where more time and resources are available, it can be a valuable analytic tool.

Activity 8.4

How does activity theory (AT) analysis differ from and how is it similar to distributed cognition (DC) analysis?

Comment

Show/Hide

For those interested in exploring activity theory and distributed cognition further, Baumer and Thomlinson (2011) provide a comparison of them, including examples of the use of these theories in video analysis.

8.7 Presenting the Findings

The best way to present findings depends on the audience, and the original goals of the study. However, it also is dependent on the data gathering and analysis techniques used.

In the previous sections of this chapter, you met many different ways of presenting findings – as numbers, through various graphical devices, in tables, in textual descriptions, as a set of themes or categories, and so on. These representations may be used directly to report your findings, provided they are appropriate for your audience and your purpose, or they may be used as background evidence for a different form of representation.

Broadly speaking, data gathering and analysis in interaction design are carried out for one of two purposes: to derive requirements for an interactive product, or to evaluate an interactive product under development. These two purposes have their own needs in terms of the notations to use and the information to be highlighted, but they also have similarities in terms of the choices to be made for presentation. For example, they usually involve reporting findings to a technical design team who will act on the findings.

In this section, we discuss three kinds of presentation style that we have not focused on as yet: using rigorous notations, using stories, and summarizing. There are other ways of presenting findings, but these are representative of the main options; they are not mutually exclusive and are often used in combination.

8.7.1 Rigorous Notations

A number of rigorous notations have been developed to analyze, capture, and present information for interaction design. The term rigorous is not intended to imply formality or rigidity, but simply to say that the notations have clear syntax and semantics. For example, the work models promoted in contextual design (Beyer and Holtzblatt, 1998) use simple but clear conventions for representing flows, breakdowns, individual roles, and so on. The modeling language UML (Unified Modeling Language) has stricter and more precise syntax to be followed and is more often used to specify internal software designs (although it has also been used and extended in the context of user interface design, e.g. Van den Bergh and Coninx, 2005).

Advantages of using a rigorous notation are that it gives you clear guidance on what to look for in the findings and what to highlight, and that it forces you to be precise. Disadvantages include that by highlighting specific elements, it inevitably also downplays or ignores other aspects, and that the precision expressed by the notation may be lost on an audience if they don't know the notation well. Producing diagrams in these notations inevitably requires further analysis of the findings in order to identify the specific characteristics and properties that the notation highlights.

To overcome their disadvantages, rigorous notations are usually used in combination with stories or other more easily accessible formats.

8.7.2 Using Stories

Storytelling is an easy and intuitive approach for people to communicate ideas and experiences. It is not surprising then that stories (also called narratives) are used extensively in interaction design, both to communicate findings of investigative studies, and as the basis for further development such as product design or system enhancements.

Storytelling may be employed in three different ways. First, participants (i.e. interviewees, questionnaire respondents, and those you have observed) may have told stories of their own during data gathering. These stories can be extracted, compared, and may be used to communicate findings to others (e.g. as anecdotes to bring a summary report to life).

Second, stories about participants may be employed. For example, in reporting her findings about European culture, Bell (2001) presents general themes and overall findings, but then illustrates these with specific stories about participants she observed. For example, one of the themes from her observations was that food shopping is important in European cultures. To illustrate the importance of the local market in France, she tells the following story:

> **Patrice lives with her husband Frederico and their two small children in a tiny village in Brittany. In her early 30s, Patrice has retired from working in the company she and her husband started. She now runs her household and looks after her kids and husband. Like many other European women we interviewed, she expressed serious reservations about catalog shopping, saying "I like to touch and see things before I buy them." And although she lives just a five-minute drive from a town with a large supermarket, Patrice chooses to shop in the local produce markets, which cycle through southern Brittany. It is important to buy what is locally seasonally available: "It just tastes better," Patrice says.**

Including these kinds of specific stories gives credibility and adds weight to the summary. Making a multimedia presentation of the story by adding video or audio excerpts and photographs will illustrate the story further. This kind of approach can be particularly effective if presenting data from an evaluation study that involves observation, as it is hard to contest well-chosen video excerpts of users interacting with technology or anecdotes from interview transcripts.

Third, stories may be constructed from smaller anecdotes or repeated patterns that are found in the data. In this case, stories provide a way of rationalizing and collating data to form a representative account of a product's use or a certain type of event.

Any stories collected through data gathering may be used as the basis for constructing scenarios. Scenarios are hypothesized stories about people and their daily life. They are a powerful technique for interaction design and can be used throughout the lifecycle. See Chapters 10 and 11 for more information on scenarios.

8.7.3 Summarizing the Findings

Clearly written reports with an overview at the beginning and a detailed content list make for easy reading and a good reference document. Including anecdotes, quotations, pictures, and video clips helps to bring the study to life, stimulate interest, and make the written description more meaningful. Some teams emphasize quantitative data, but its value depends on the type of study. Often both qualitative and quantitative data analysis are used because they provide alternative perspectives.

Some audiences are likely to be more interested in the headline findings than in the details of stories or precise specifications. These may be high-level managers, the set of users who acted as participants in studies, or product designers who want to use the results to develop a new product. Whoever they are, being able to present a summary of the findings is important. This is where numbers and statistical values (if you have any) can be really powerful. However, in these summaries it is important not to overstate your findings – if 8 out of 10 users preferred design A over design B, this does not mean that design A is 80% more attractive than design B. If you found 800 out of 1000 users preferred design A then you have more evidence to suggest that design A is better, but there are still other factors to take into account. In general, be wary of using words such as ‘most,’ ‘all,’ ‘majority,’ ‘none,’ and be careful when writing justifications to ensure that they reflect the data.

Activity 8.5

Consider each of the findings below and the associated summary statement about it. For each one, comment on whether the finding supports the statement.

1. Finding: Two out of four people who filled in the questionnaire ticked the box that said they prefer not to use the ring-back facility on their cell phone.

   Statement: Half of the users don't use the ring-back facility.

2. Finding: Joan who works in the design department was observed one day walking for 10 minutes to collect printout from the high-quality colour printer.

   Statement: Significant time is wasted by designers who have to walk a long distance to collect printout.

3. Finding: A data log of 1000 hours of interaction with a website recorded during January, February, and March records 8 hours spent looking at the helpfiles.

   Statement: The website's helpfiles were used less than 1% of the time during the first quarter of the year.

Comment

Show/Hide

Assignment

The aim of this assignment is for you to practice data analysis and presentation. Assume that you are to present the findings of your data gathering assignment from Chapter 7 to a group of peers, e.g. through a seminar.

a. Review the data you have gathered and identify any qualitative data and any quantitative data in your data set.

b. Is there any qualitative data that could sensibly and helpfully be translated into quantitative measures? If so, do the translation and add this data to your quantitative set.

c. Consider your quantitative data.

   - Decide how best to enter your quantitative data into your spreadsheet software. For example, you need to consider how to handle answers to closed questions. Then enter the data and generate some graphical representations. As you are likely to have very few records, i.e. respondents or interviewees, in your data set, you will have to think carefully about what, if any, graphical representations will provide meaningful summaries of your findings.
   - Is there any data for which simple measures such as percentages or averages will be helpful? If so, calculate the three different types of average.

d. Consider your qualitative data.

   - Based on your refinement of the study question 'improving the product,' identify some themes from your qualitative data, e.g. what features of the product cause people difficulties, did any of your participants suggest alternative designs or solutions? Refine your themes and collate extracts of data which support the theme.
   - Identify any critical incidents in your data. These may arise from interview or questionnaire responses, or from observation. Describe these incidents carefully and choose one or two to analyze in more depth, focusing on the context in which they occurred.

e. Collate your findings as a presentation and deliver them to a group of your peers.

f. Review your presentation and any questions you received from your peers and consider where your analysis and presentation could be improved.

Take a Quickvote on Chapter 8:
www.id-book.com/quickvotes/chapter8

Summary

This chapter has described in detail the difference between qualitative and quantitative data and between qualitative and quantitative analysis.

Quantitative and qualitative data can be analyzed for patterns and trends using simple techniques and graphical representations. Qualitative data may be analyzed using a variety of approaches including identifying categories or themes, and using theoretical frameworks, such as grounded theory, distributed cognition, and activity theory.

It was noted that presenting the results is just as important as analyzing the data, hence it is important to make sure that any summary or claim arising from the analysis is carefully contextualized, and can be justified by the data.

Key points

- The kind of data analysis that can be done depends on the data gathering techniques used.
- Qualitative and quantitative data may be collected from any of the main data gathering techniques: interviews, questionnaires, and observation.
- Quantitative data analysis for interaction design usually involves calculating percentages and averages.
- There are three different kinds of average: mean, mode, and median.
- Graphical representations of quantitative data help in identifying patterns, outliers, and the overall view of the data.
- Qualitative data analysis may be framed by theories. Three such theories are grounded theory, activity theory, and distributed cognition.

Further Reading

CHARMAZ, K. (2014) Constructing Grounded Theory (2nd edn). Sage Publications. This popular book also provides a useful account of how to actually do grounded theory.

Corbin, J. M. and STRAUSS, A. (2014) Basics of Qualitative Research: Techniques and procedures for developing grounded theory. Sage. This presents a readable and practical account of applying the grounded theory approach. It is not tailored specifically to interaction design and therefore requires some interpretation. It is a good discussion of the basic approach.

Huff, D. (1991) How to Lie with Statistics. Penguin. This wonderful little book illustrates the many ways in which numbers can be misrepresented. Unlike some (many) books on statistics, the text is easy to read and amusing.

kuutti, K. (1996) Activity theory as a potential framework for human–computer interaction. In B. A. Nardi (ed.) Context and Consciousness. MIT Press, pp. 17–44. This provides a digestible description of activity theory and how it can be applied in HCI.

LAZAR, J., FENG, J. H. and HOCHHEISER, H. (2010) Research Methods in Human–Computer Interaction. John Wiley & Sons Ltd. This is a good resource for a more detailed discussion of data gathering and analysis in interaction design.

Rogers, Y. (2006) Distributed cognition and communication. In K. Brown (ed.) The Encyclopedia of Language and Linguistics (2nd edn). Elsevier, pp. 731–733. (A version can also be downloaded from Yvonne Rogers' website.) This chapter provides a readable introduction to the background and application of distributed cognition.

CHAPTER 9
THE PROCESS OF INTERACTION DESIGN

Objectives

The main aims of this chapter are to:

- Consider what doing interaction design involves.
- Explain some advantages of involving users in development.
- Explain the main principles of a user-centered approach.
- Present a simple lifecycle model of interaction design.
- Ask and provide answers for some important questions about the interaction design process.
- Consider how interaction design activities can be integrated into the wider product development lifecycle.

9.1 Introduction

Design is a practical and creative activity with the aim of developing a product that helps its users achieve their goals. In previous chapters, we looked at different kinds of interactive products, issues that need to be taken into account when doing interaction design, some of the theoretical bases for the field, and techniques for gathering and analyzing data to understand users' goals. In this chapter we start to explore how we can design and build interactive products.

Chapter 1 defined interaction design as being concerned with 'designing interactive products to support the way people communicate and interact in their everyday and working lives.' But how do you go about doing this? Developing a product must begin with gaining some understanding of what is required of it, but where do these requirements come from? Whom do you ask about them? Underlying good interaction design is the philosophy of user-centered design, i.e. involving users throughout development, but who are the users? Will they know what they want or need even if we can find them to ask? For an innovative product, users are unlikely to be able to envision what is possible, so where do these ideas come from?

In this chapter, we raise and answer these kinds of questions, discuss user-centered design, and revisit the four basic activities of the interaction design process that were introduced in Chapter 1. We also introduce a lifecycle model of interaction design that captures these activities.

9.2 What Is Involved in Interaction Design?

The previous chapters have introduced you to many exciting ideas, approaches, theories, and opinions about interaction design, but what does it mean to actually do interaction design? The following activity is intended to start you thinking about this by asking you to produce an initial design for an interactive product.

Activity 9.1

Imagine that you want to design a travel planner for yourself. You might use this system to plan your route, check visa requirements, book flights or train tickets, investigate the facilities at your destination, and so on. Assume that the system is destined to run on a tablet for the purposes of this activity.

1. Make a list of the user experience and usability goals for the system.
2. Outline the initial screen or two for this system, showing its main functionality and its general look and feel. Spend about 10 minutes on this.
3. Having produced an outline, spend 5 minutes reflecting on how you went about tackling this activity. What did you do first? Did you have any particular artifacts or experience to base your design upon? What process did you go through?

Comment

Show/Hide

BOX 9.1

Four Approaches to Interaction Design

Saffer (2010) suggests four main approaches to interaction design, each of which is based on a distinct underlying philosophy: user-centered design, activity-centered design, systems design, and genius design. He acknowledges that the purest form of any of these is unlikely to be realized, and takes an extreme view of each in order to distinguish between them. In user-centered design, the user knows best and is the only guide to the designer; the designer's role is to translate the users' needs and goals into a design solution. Activity-centered design focuses on the behavior surrounding particular tasks. Users still play a significant role but it is their behavior rather than their goals and needs that are important. Systems design is a structured, rigorous, and holistic design approach that focuses on context and is particularly appropriate for complex problems. In systems design it is the system (i.e. the people, computers, objects, devices, and so on) that are the center of attention while the users' role is to set the goals of the system.

Finally, genius design is different from the other three approaches because it relies solely on the experience and creative flair of a designer. Jim Leftwich, an experienced interaction designer interviewed in Saffer (2010, pp. 44–5) prefers the term 'rapid expert design.' In this approach the users' role is to validate ideas generated by the designer, and users are not involved during the design process itself. Saffer points out that this is not necessarily by choice, but may be due to limited or no resources for user involvement. Apple, for example, does very little user research or testing, yet the Apple iPod is acknowledged as a significant design achievement.

Different design problems lend themselves more easily to different approaches, and different designers will tend to gravitate towards using the approach that suits them best. Although an individual designer may prefer a particular approach, it is important that the approach for any one design problem is chosen with that design problem in mind. ■

There are many fields of design; for example, graphic design, architectural design, industrial and software design, and although each discipline has its own interpretation of 'designing,' there are three fundamental activities that are recognized in all design: understanding the requirements, producing a design that satisfies those requirements, and evaluating the design.

Interaction design also involves these activities, and in addition we focus attention very clearly on users and their goals. For example, we investigate the artifact's use and target domain by taking a user-centered approach to development, we seek users' opinions and reactions to early designs, and we involve users appropriately in the development process itself. This means that users' concerns direct the development rather than just technical concerns, and for interaction design, the three fundamental activities of design listed above are extended to include an activity of prototyping so that users can interact with the design.

So design involves work on requirements, designing a solution, producing an interactive version of the solution, and evaluating it. But design is also about trade-offs, about balancing conflicting requirements. One common form of trade-off when developing a system to offer advice is deciding how much choice will be given to the user and how much direction the system should offer. Often the division will depend on the purpose of the system, e.g. for business travel or for vacations. Getting the balance right requires experience, but it also requires the development and evaluation of alternative solutions.

Generating alternatives is a key principle in most design disciplines, and one that should be encouraged in interaction design. Linus Pauling, twice a Nobel Prize winner, once said: “The best way to get a good idea, is to get lots of ideas.” This is not necessarily easy, however, and unlike many design disciplines, interaction designers are not generally trained to generate alternative designs. The good news is that the ability to brainstorm and contribute alternative ideas can be learned. For example, Kelley (2008) describes seven secrets for better brainstorms, including sharpen the focus (have a well-honed problem statement), playful rules (to encourage ideas), and get physical (use visual props). He also discusses six ways to kill a brainstorm, including do it off-site and write everything down.

Involving users and others in the design process means that the designs and potential solutions will need to be communicated to people other than the original designer. This requires the design to be captured and expressed in some suitable form that allows review, revision, and improvement. There are many ways of doing this, one of the simplest being to produce a series of sketches. Other common approaches are to write a description in natural language, to draw a series of diagrams, and to build prototypes. A combination of these techniques is likely to be the most effective. When users are involved, capturing and expressing a design in a suitable format is especially important since they are unlikely to understand jargon or specialist notations. In fact, a form that users can interact with is most effective, and building prototypes is an extremely powerful approach (see Box 9.2 and Chapter 11 for more on prototyping).

In the rest of this section, we explore in more depth the significance and practicality of involving users in design, i.e. using a user-centered approach, and consider again the four activities of interaction design that were introduced in Chapter 1.

BOX 9.2

The Value of Prototyping

I learned the value of prototyping through a very effective role-playing exercise. I was on a course designed to introduce new graduates to different possible careers in industry. One of the themes was production and manufacturing and the aim of one group exercise was to produce a book. Each group was told that it had 30 minutes to deliver 10 books to the person in charge. Groups were given various pieces of paper, scissors, sticky tape, staples, etc., and told to organize ourselves as best we could. So in my group we set to work organizing ourselves into a production line, with one of us cutting up the paper, another stapling the pages together, another sealing the binding with the sticky tape, and so on. One person was even in charge of quality assurance. It took us less than 10 minutes to produce the 10 books, and we rushed off with our delivery. When we showed the person in charge, he replied, ‘That's not what I wanted; I need it bigger than that.’ Of course, the size of the notebook wasn't specified in the description of the task, so we found out how big he wanted it, got some more materials, and scooted back to produce 10 more books. Again, we set up our production line and produced 10 books to the correct size. On delivery, we were again told that it was not what was required: he wanted the binding down the other edge. This time we got as many of the requirements as we could and went back, developed one book, and took that back for further feedback and refinement before producing the 10 required.

If we had used prototyping as a way of exploring our ideas and checking requirements in the first place, we could have saved so much effort and resource! ■

9.2.1 The Importance of Involving Users

The description above emphasizes the need to involve users in interaction design, but why is it important? Before the impact that user involvement can have on project success was recognized, it was common for developers to talk to managers or to proxy users, i.e. people who role-played as users, when eliciting requirements, or even to use their own judgment without reference to anyone else. While a proxy user can provide useful information, they will not have the same perspective as someone who performs the task every day, or who will use the intended product on a regular basis. For example, several years ago, I was involved with a system to process and record financial transactions from the foreign exchange (forex) dealers in a large international bank. The users of this system took the handwritten transaction records completed by the forex dealers and entered the details into the system. The system then validated the transaction and communicated a confirmation to the relevant parties. When the requirements for this system were developed, no one from the development team spoke to the end-users of the system; the requirements were identified by higher level managers. Although the system was successfully used to support the task, the end-users had developed several work-arounds and crib sheets. For example, each desk had a sheet of paper with lists of buyer codes and names. The system required both buyer code and buyer name to be entered, but it was quite common for the dealers to write only the buyer name on the transaction record. The list of names and codes was used to identify the codes manually.

The best way to ensure that development continues to take users' activities into account is to involve real users throughout development. In this way, developers can gain a better understanding of users' goals, leading to a more appropriate, more usable product. However, two other aspects that have nothing to do with functionality are equally as important if the product is to be usable and used: expectation management and ownership.

Expectation management is the process of making sure that the users' expectations of the new product are realistic. The purpose of expectation management is to ensure that there are no surprises for users when the product arrives. If users feel they have been cheated by promises that have not been fulfilled, then this will cause resistance and even rejection. Marketing of the new arrival must be careful not to misrepresent the product, although it may be particularly difficult to achieve with a large and complex system (Nevo and Wade, 2007). How many times have you seen an advert for something you thought would be really good to have, but when you see one, discover that the marketing hype was a little exaggerated? I expect you felt quite disappointed and let down. This is the kind of feeling that expectation management tries to avoid.

It is better to exceed users' expectations than to fall below them. This does not mean adding more features, but that the product supports the users' goals more effectively than they expect. Involving users throughout development helps with expectation management because they can see from an early stage what the product's capabilities are. They will also understand better how it will affect their jobs and lives, and why the features are designed that way. Adequate and timely training is another technique for managing expectations. If users have the chance to work with the product before it is released, through training or hands-on demonstrations of a pre-release version, then they will understand better what to expect when the final product is available.

A second reason for user involvement is ownership. Users who are involved and feel that they have contributed to a product's development are more likely to feel a sense of ownership towards it and support its use.

9.2.2 Degrees of User Involvement

Different degrees of user involvement may be implemented in order to manage expectations and to create a feeling of ownership. At one end of the spectrum, users may be co-opted to the design team so that they are major contributors. For any one user, this may be on a full-time basis or a part-time basis, and it may be for the duration of the project or for a limited time only. There are advantages and disadvantages to each situation. If a user is co-opted full-time for the whole project, their input will be consistent and they will become very familiar with the product and its rationale. However, if the project takes many years, they may lose touch with the rest of the user group, making their input less valuable. If a user is co-opted part-time for the whole project, she will offer consistent input to development while remaining in touch with other users. Depending on the situation, this will need careful management as the user will be trying to learn new jargon and handle unfamiliar material as a member of the design team, yet concurrently trying to fulfill the demands of her original job. This can become very stressful for the individuals. If a number of users from each user group are co-opted part-time for a limited period, input is not necessarily consistent across the whole project, but careful coordination between users can alleviate this problem. In this case, one user may be part of the design team for 6 months, then another takes over for the next 6 months, and so on.

At the other end of the spectrum, users may simply be kept informed through regular newsletters or other channels of communication. Provided they are able to influence the development process through workshops or similar events, this can be an effective approach to expectation management and ownership. In a situation with hundreds or even thousands of users, it would not be feasible to involve them all as members of the team, and so this might be the only viable option. In this case a compromise situation is probably the best. Representatives from each user group may be co-opted onto the team on a full-time basis, while other users are involved through design workshops, evaluation sessions, and other data-gathering activities.

The individual circumstances of the project affect what is realistic and appropriate. If your end-user groups are identifiable, e.g. you are developing a product for a particular company, then it is easier to involve them. If, however, you are developing a product for the open market, it is unlikely that you will be able to co-opt a user to your design team, and so alternative approaches are needed. Box 9.3 outlines an alternative way to obtain user input.

How actively users should be involved is a matter for debate. Some studies have shown that too much user involvement can lead to problems. This issue is discussed in the next Dilemma box.

BOX 9.3

Ongoing User Involvement after a Product is Released

Once a product has been released and the focus of development moves to future versions, a different kind of user involvement is possible – one that captures data about the real use of the product. This may be obtained in a number of ways, e.g. through interaction between users and customer service agents, or through automated error reporting systems. For example, Microsoft has millions of customers around the world, about 30% of whom call their customer support lines with problems and frustrations resulting from poor features or software errors. This data about customer behavior and their problems with the products is fed back into product development and improvement (Cusumano and Selby, 1995).

Error reporting systems (ERS, also called online crashing analysis) automatically collect information from users (with their permission), which is used to improve applications in the longer term. Figure 9.2 shows two typical dialog boxes for the Windows error reporting system that is built into Microsoft operating systems (Glerum et al., 2009). ■

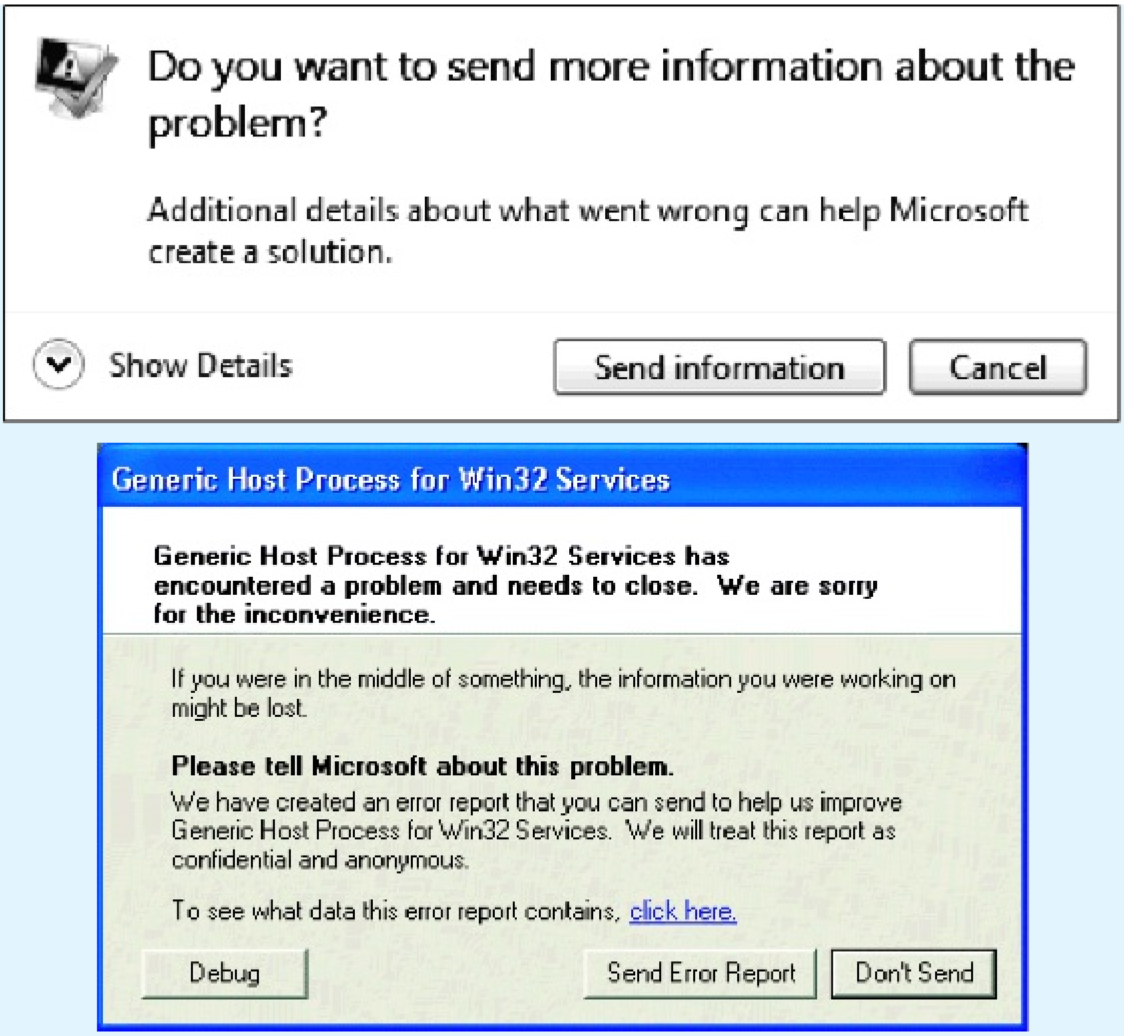


Figure 9.2 Two typical dialog boxes from the Windows error reporting system. This kind of reporting can have a significant effect on the quality of applications: for example, 29% of the errors fixed by the Windows XP (Service Pack 1) team were based on information collected through their ERS

Dilemma

Too Much of a Good Thing?

Involving users in development is a good thing. Or is it? And how much should they become involved? In what role(s)? Are users qualified to lead a technical development project? Or is it more appropriate for their role to be confined to evaluating prototypes? Involving users to any degree incurs costs, so what evidence is there that user involvement is productive, or that it is worth putting the required level of resources into it?

Research by Keil and Carmel (1995) indicates that the more successful projects have direct links to users and customers, while Kujala and Mäntylä (2000) concluded that user studies conducted early in development produce benefits that outweigh the costs of conducting them. Subrayaman et al (2010) found a mixed result. They investigated the impact of user participation on the satisfaction with the product by both developers and users. They found that for new products, developer satisfaction increased as user participation increased, while user satisfaction was higher where participation was low, but satisfaction dropped as participation increased. For maintenance projects, both developers and users were most satisfied with a moderate level of participation (approximately 20% of overall project development time). If we focus just on the user satisfaction as an indication of project success, then it seems that low user participation is most beneficial.

According to some research, a low level of user involvement is beneficial whereas too much can lead to problems. For example, Heinbokel et al (1996) found that high user involvement projects tended to run less smoothly. Subrayaman et al (2010) also identified that high levels of user involvement can generate unnecessary conflicts and increased reworking.

The kind of product being developed, the kind of user involvement possible, and the application domain all influence the impact that users can have on a project. Scaife et al (1997) suggest that involving different kinds of user at different stages of design yields positive results, while Wagner and Piccoli (2007) provide evidence that user involvement is most effective when the product becomes salient to their daily lives, which is often close to or even after deployment.

Abelein et al (2013) performed a detailed review of the literature in this area and concluded that, overall, the evidence indicates user involvement has a positive effect on user satisfaction and system use. However, they also found that even though the data clearly indicates this positive effect, some links have a large variation, suggesting that there is still no clear conceptual model to measure the effects consistently. In addition, they found that most studies with negative correlations involving users and system success were published over a decade ago.

While user involvement is widely acclaimed as beneficial, there are still many dimensions that need to be considered and balanced. ■

9.2.3 What Is a User-Centered Approach?

Throughout this book, we emphasize the need for a user-centered approach to development. By this we mean that the real users and their goals, not just technology, are the driving force behind product development. As a consequence, a well-designed system will make the most of human skill and judgment, will be directly relevant to the activity in hand, and will support rather than constrain the user. This is less of a technique and more of a philosophy.

In 1985, Gould and Lewis (1985) laid down three principles they believed would lead to a 'useful and easy to use computer system':

1. Early focus on users and tasks. This means first understanding who the users will be by directly studying their cognitive, behavioral, anthropomorphic, and attitudinal characteristics. This requires observing users doing their normal tasks, studying the nature of those tasks, and then involving users in the design process.
2. Empirical measurement. Early in development, the reactions and performance of intended users to printed scenarios, manuals, etc., are observed and measured. Later on, users interact with simulations and prototypes and their performance and reactions are observed, recorded, and analyzed.
3. Iterative design. When problems are found in user testing, they are fixed and then more tests and observations are carried out to see the effects of the fixes. This means that design and development is iterative, with cycles of design–test–measure–redesign being repeated as often as necessary.

These three principles are now accepted as the basis for a user-centered approach (e.g. see Mao et al, 2005) but when Gould and Lewis wrote their paper, they were not accepted by most developers. In fact, they remark in their paper that when they started recommending these to designers, the designers' reactions implied that these principles were obvious. However, when they asked designers at a human factors symposium for the major steps in software design, most of them did not cite many of the principles – in fact, only 2% mentioned all of them. So maybe they had obvious merit, but they were not so easy to put into practice. The Olympic Messaging System (OMS) (Gould et al, 1987) was the first reported large computer-based system to be developed using these three principles. Here a combination of techniques was used to elicit users' reactions to designs, from the earliest prototypes through to the final product. In this case, users were mainly involved in evaluating designs. Below, we discuss these principles in more detail.

Early Focus on Users and Tasks

This principle can be expanded and clarified through the following five further principles:

1. Users' tasks and goals are the driving force behind the development.

   In a user-centered approach to design, while technology will inform design options and choices, it should not be the driving force. Instead of saying 'Where can we deploy this new technology?' say 'What technologies are available to provide better support for users' goals?'

2. Users' behavior and context of use are studied and the system is designed to support them.

   This is about more than just capturing the tasks and the users' goals. How people perform their tasks is also significant. Understanding behavior highlights priorities, preferences, and implicit intentions. One argument against studying current behavior is that we are looking to improve work, not to capture bad habits in automation. The implication is that exposing designers to users is likely to stifle innovation and creativity, but experience tells us that the opposite is true (Beyer and Holtzblatt, 1998). In addition, if something is designed to support an activity with little understanding of the real work involved, it is likely to be incompatible with current practice, and users don't like to deviate from their learned habits if operating a new device with similar properties (Norman, 2013).

3. Users' characteristics are captured and designed for.

   When things go wrong with technology, we often say that it is our fault. But as humans, we are prone to making errors and we have certain limitations, both cognitive and physical. Products designed to support humans should take these limitations into account and should limit the mistakes we make. Cognitive aspects such as attention, memory, and perception issues were introduced in Chapter 3. Physical aspects include height, mobility, and strength. Some characteristics are general, such as about one man in 12 has some form of color blindness, but some characteristics may be associated more with the job or particular task at hand. So, as well as general characteristics, we need to capture those specific to the intended user group.

4. Users are consulted throughout development from earliest phases to the latest and their input is seriously taken into account. As discussed above, there are different levels of user involvement and there are different ways in which to consult users. However involvement is organized, it is important that users are respected by designers.

5. All design decisions are taken within the context of the users, their work, and their environment. This does not necessarily mean that users are actively involved in design decisions.

   As long as designers remain aware of the users while making their decisions, then this principle will be upheld. Keeping this context in mind can be difficult, but using personas is one way to achieve this (see Chapter 10); an easily accessible collection of gathered data is another. Some design teams set up a specific design room for the project where data and informal records of brainstorming sessions are pinned on the walls or left on the table.

Activity 9.2

Assume that you are involved in developing a novel online experience for buying garden plants. Although several websites exist for buying plants online, you want to produce a distinct experience to increase the organization's market share. Suggest ways of applying the above principles in this task.

Comment

Show/Hide

Empirical Measurement

Specific usability and user experience goals should be identified, clearly documented, and agreed upon at the beginning of the project. They can help designers to choose between alternative designs and to check on progress as the product is developed. Identifying specific goals up front means that the product can be empirically evaluated at regular stages as it is developed.

Iterative Design

Iteration allows designs to be refined based on feedback. As users and designers engage with the domain and start to discuss requirements, needs, hopes, and aspirations, then different insights into what is needed, what will help, and what is feasible will emerge. This leads to a need for iteration, for the activities to inform each other and to be repeated. However good the designers are and however clear the users may think their vision is of the required artifact, it will be necessary to revise ideas in light of feedback, several times. This is particularly true when trying to innovate. Innovation rarely emerges whole and ready to go. It takes time, evolution, trial and error, and a great deal of patience. Iteration is inevitable because designers never get the solution right the first time (Gould and Lewis, 1985).

9.2.4 Four Basic Activities of Interaction Design

Four basic activities for interaction design were introduced in Chapter 1, some of which you will have engaged in when doing Activity 9.1. These are: establishing requirements for the user experience, designing alternatives that meet those requirements, prototyping the alternative designs so that they can be communicated and assessed, and evaluating what is being built throughout the process and the user experience it offers. They are fairly generic activities and can be found in other design disciplines too, such as graphic design, architectural design, and product design.

Establishing Requirements

In order to design something to support people, we must know who our target users are and what kind of support an interactive product could usefully provide. These needs form the basis of the product's requirements and underpin subsequent design and development. This activity is fundamental to a user-centered approach, and is very important in interaction design. Understanding these needs is gleaned through data gathering and analysis, which are discussed in Chapters 7 and 8. The requirements activity is discussed further in Chapter 10.

Designing Alternatives

This is the core activity of designing: actually suggesting ideas for meeting the requirements. This activity can be viewed as two sub-activities: conceptual design and concrete design. Conceptual design involves producing the conceptual model for the product, and a conceptual model describes an abstraction outlining what people can do with a product and what concepts are needed to understand how to interact with it. Concrete design considers the detail of the product including the colors, sounds, and images to use, menu design, and icon design. Alternatives are considered at every point. You met some of the ideas for conceptual design in Chapter 2, and some more design issues for specific interface types in Chapter 6; we go into more detail about how to design an interactive product in Chapter 11.

Prototyping

Interaction design involves designing interactive products. The most sensible way for users to evaluate such designs is to interact with them, and this can be achieved through prototyping. This does not necessarily mean a piece of software is required. There are different prototyping techniques, not all of which require a working piece of software. For example, paper-based prototypes are very quick and cheap to build and are very effective for identifying problems in the early stages of design, and through role-playing users can get a real sense of what it will be like to interact with the product. Prototyping is also covered in Chapter 11.

Evaluating

Evaluation is the process of determining the usability and acceptability of the product or design that is measured in terms of a variety of usability and user experience criteria. Interaction design requires a high level of user involvement throughout development, and this enhances the chances of an acceptable product being delivered. Evaluation does not replace the activities concerned with quality assurance and testing to make sure that the final product is fit for purpose, but it complements and enhances them. We devote Chapters 13 to 15 to evaluation.

The activities of establishing requirements, designing alternatives, building prototypes, and evaluating them are intertwined: alternatives are evaluated through the prototypes and the results are fed back into further design or might identify missing requirements. This iteration is one of the key characteristics of a user-centered approach.

9.2.5 A Simple Lifecycle Model for Interaction Design

Understanding what activities are involved in interaction design is the first step to being able to do it, but it is also important to consider how the activities are related to one another so that the full development process can be seen. The term lifecycle model is used to represent a model that captures a set of activities and how they are related. Software engineering has spawned many lifecycle models including the waterfall, spiral, and RAD (rapid applications development) models (Sommerville, 2010). One lifecycle model that was not widely used, but which was intended to capture the iterative nature of software development, was called the fountain lifecycle model (Henderson-Sellers and Edwards, 1993). The field of HCI has also been associated with several lifecycle models such as the Star (Hartson and Hix, 1989) and an international standard model ISO 9241-210.

Existing models have varying levels of sophistication and complexity. For projects involving only a few experienced developers, a simple process would probably be adequate. However, for larger systems involving tens or hundreds of developers with hundreds or thousands of users, a simple process just isn't enough to provide the management structure and discipline necessary to engineer a usable product. So something is needed that will provide more formality and more discipline (described in Box 9.1 as the systems design approach to interaction design). Note that this does not necessarily mean that innovation is lost or that creativity is stifled, just that a structured process is used to provide a more stable framework for creativity.

However simple or complex it appears, any lifecycle model is a simplified version of reality. It is intended as an abstraction and, as with any good abstraction, only the amount of detail required for the task at hand will be included. Any organization wishing to put a lifecycle model into practice will need to add detail specific to its particular circumstances and culture.

The activities of interaction design are related as shown in Figure 9.3. This model incorporates the four activities of interaction design and the three principles of user-centered design discussed above. Depending on the kind of product being developed, it may not be possible or appropriate to follow this model for every aspect, and more detail would be required to put the lifecycle into practice in a real project. We have not specified outputs from each activity, although measurable usability and user experience criteria would be specified early on and referred to across all the activities.

The model is not intended to be prescriptive; that is, we are not suggesting that this is how all interactive products are or should be developed. It is based on our observations of interaction design and on information we have gleaned in the research for this book. It has its roots in the software engineering and HCI lifecycle models mentioned above and it represents what we believe is practiced in the field.

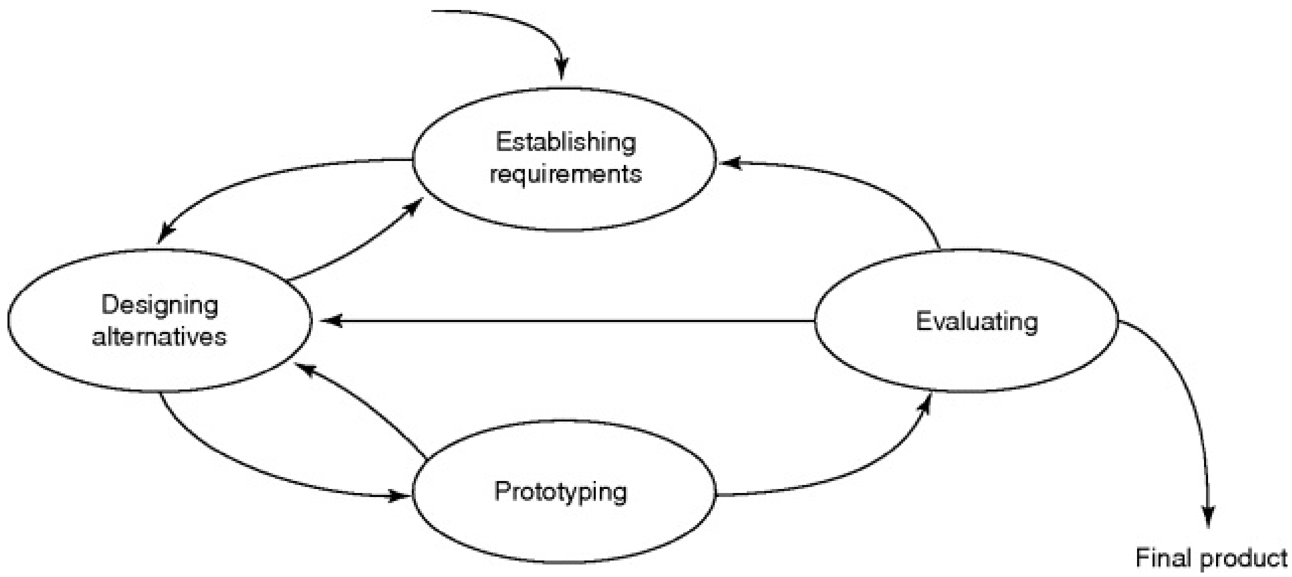


Figure 9.3 A simple interaction design lifecyle model

Most projects start by establishing requirements. The project may have arisen because of some evaluation that has been done, but the lifecycle of the new (or modified) product can be thought of as starting at this point. From this activity, some alternative designs are generated in an attempt to meet the requirements that have been identified. Then prototype versions of the designs are developed and evaluated. Based on the feedback from the evaluations, the team may need to return to identify more needs or refine requirements, or it may go straight into redesigning. It may be that more than one alternative design follows this iterative cycle in parallel with others, or it may be that one alternative at a time is considered. Implicit in this cycle is that the final product will emerge in an evolutionary fashion from a rough initial idea through to the finished product. Exactly how this evolution happens may vary from project to project, and we return to this issue in Chapter 11. The only factor limiting the number of times through the cycle is the resources available, but whatever the number is, development ends with an evaluation activity that ensures the final product meets the prescribed user experience and usability criteria.

9.3 Some Practical Issues

The discussion so far has highlighted some issues about the practical application of user-centered design and the simple lifecycle of interaction design that we have introduced. These issues must be addressed in order to be able to do interaction design in practice. We capture these issues in the following questions:

- Who are the users?

- What do we mean by needs?
- How do you generate alternative designs?
- How do you choose among alternatives?
- How do you integrate interaction design activities with other lifecycle models?

9.3.1 Who Are the Users?

With all this emphasis on users and user involvement in the interaction design process, a fairly basic question to ask is 'Who are the users?'.

Identifying the users may seem like a straightforward activity, but in fact there are many interpretations of the term 'user,' and involving the right users is crucial to successful user-centered design. The most obvious definition is those people who interact directly with the product to achieve a task. Most people would agree with this definition; however, there are others who can also be thought of as users. For example, Holtzblatt and Jones (1993) include in their definition of users those who manage direct users, those who receive products from the system, those who test the system, those who make the purchasing decision, and those who use competitive products. Eason (1987) identifies three categories of user: primary, secondary, and tertiary. Primary users are those likely to be frequent hands-on users of the system; secondary users are occasional users or those who use the system through an intermediary; and tertiary users are those affected by the introduction of the system or who will influence its purchase.

The trouble is that there is a surprisingly wide collection of people who all have a stake in the development of a successful product. These people are called stakeholders. Stakeholders are "people or organizations who will be affected by the system and who have a direct or indirect influence on the system requirements" (Kotonya and Sommerville, 1998). Dix et al (2004) make an observation that is very pertinent to a user-centered view of development: "It will frequently be the case that the formal 'client' who orders the system falls very low on the list of those affected. Be very wary of changes which take power, influence or control from some stakeholders without returning something tangible in its place."

The group of stakeholders for a particular product will be larger than the group of people normally thought of as users, although it will of course include users. Using the definition above, the group of stakeholders includes the development team itself as well as its managers, the direct users and their managers, recipients of the product's output, people who may lose their jobs because of the introduction of the new product, and so on (Sharp et al, 1999).

For example, consider again the travel planner in Activity 9.1. According to the description we gave you, the user group for the system has just one member: you. However, the stakeholders for the system would also include the people you are going to visit, the airlines you book flights with, staff in the hotels you might stay at, a wide selection of companies and staff members who have an interest to make sure that any information you are given is correct, and even the restaurants on the route chosen for your journey, since the route suggested by the system will determine whether or not you drive past those restaurants.

Identifying the stakeholders for your project means that you can make an informed decision about who should be involved and to what degree, but how to make sure that you include stakeholders who are relevant is more complex. Alexander and Robertson (2004) suggest using an onion diagram to model stakeholders and their involvement. This diagram shows concentric circles of stakeholder zones with the product being developed sitting in the middle. Lim et al (2010) developed a process and supporting tool called StakeNet which relies on recommendations through social networking to identify and prioritize relevant stakeholders. Ellen Gottesdiener expands on the role of stakeholders in her interview at the end of this chapter.

Activity 9.3

Who are the stakeholders for an automated check-out system of a large supermarket?

Comment

Show/Hide

9.3.2 What Do We Mean by ‘Needs’?

If you had asked someone in the street in the late 1990s what she needed, I doubt that the answer would have included interactive television, or a ski jacket with integrated MP3 player, or a robot pet. If you presented the same person with these possibilities and asked whether she would buy them if they were available, then the answer may have been more positive. When we talk about identifying needs, it is not simply a question of asking people, ‘What do you need?’ and then supplying it, because people don't necessarily know what is possible. Robertson and Robertson (2013) refer to ‘un-dreamed-of’ needs, which are those that users are unaware they could have. Instead, we have to approach it by understanding the characteristics and capabilities of the users, what they are trying to achieve, how they achieve it currently, and whether they would achieve their goals more effectively and have a more enjoyable experience if they were supported differently.

There are many dimensions along which a user's characteristics and capabilities may vary, and that will have an impact on the product's design. You have met some of the cognitive ones in Chapter 3. A person's physical characteristics may also affect the design: size of hands may affect the size and positioning of input buttons, and motor abilities may affect the suitability of certain input and output devices; height is relevant in designing a physical kiosk, for example; and strength in designing a child's toy – a toy should not require too much strength to operate, but may require strength greater than expected for the target age group to change batteries or perform other operations suitable only for an adult. Cultural diversity and experience may affect the terminology the intended user group is used to, or how nervous about technology a set of users may be, or how a facility is used (we discuss user requirements in more detail in Chapter 10).

If a product is a new invention, then it can be difficult to identify the users and representative tasks for them, e.g. before in-car navigation systems were first developed, there were no users to consult about requirements and there were no representative tasks to identify. Those developing the system had to imagine who might want to use it and what they might want to do with it.

It may be tempting for designers simply to design what they would like to use themselves, but their ideas would not necessarily coincide with those of the target user group, because they have different experiences and expectations. It is imperative that representative users from the real target group be consulted. For example, Netpliance developed a new product that would seamlessly integrate all the services necessary for the user to achieve a specific task on the Internet (Isensee et al, 2000). They took a user-centered approach and employed focus group studies and surveys to understand their customers’ needs, but developers observed the focus groups to learn more about their intended user group. Isensee et al (p. 60) comment that, “It is always tempting for developers to create products they would want to use or similar to what they have done before.”

Whether the product is a new invention or not, it is always useful to start by understanding similar behavior that is already established. Introducing something new into people's lives, especially a new everyday item, requires a culture change in the target user population, and it takes a long time to effect a culture change.

Focusing on people's goals and on usability and user experience goals is a more promising approach to interaction design than focusing on people's needs and expecting them to be able to tell us the requirements for a product. Techniques for data gathering to investigate these goals and to establish requirements are discussed more in Chapters 7 and 10.

9.3.3 How Do You Generate Alternative Designs?

A common human tendency is to stick with something that we know works. We recognize that a better solution may exist out there somewhere, but it is very easy to accept this one because we know it works – it is ‘good enough.’ Settling for a solution that is good enough is not necessarily bad, but it may be undesirable because good alternatives may never be considered, and considering alternative solutions is a crucial step in the process of design. But where do these alternative ideas come from?

One answer to this question is that they come from the individual designer's flair and creativity (the genius design described in Box 9.1). Although it is certainly true that some people are able to produce wonderfully inspired designs while others struggle to come up with any ideas at all, very little in this world is completely new. For example, if you think of something commonly believed to be an invention, such as the steam engine, this was in fact inspired by the observation that the steam from a kettle boiling on the stove lifted the lid. Clearly there was an amount of creativity and engineering involved in making the jump from a boiling kettle to a steam engine, but the kettle provided the inspiration to translate experience gained in one context into a set of principles that could be applied in another. Innovations usually arise through cross-fertilization of ideas from different perspectives, individuals, and contexts; the evolution of an existing product through use and observation; or straightforward copying of other, similar products.

Cross-fertilization may result by discussing ideas with other designers, while Buxton (2007) reports that different perspectives from users generated original ideas about alternative designs. As an example of evolution, consider the word processor. The capabilities of suites of office software have gradually increased from the time they first appeared. Initially, a word processor was just an electronic version of a typewriter, but gradually other capabilities, including the spell-checker, thesaurus, style sheets, graphical capabilities, and so on, were added.

Although creativity and invention are often wrapped in mystique, we do understand something of the process and of how creativity can be enhanced or inspired. We know, for instance, that browsing a collection of designs will inspire designers to consider alternative perspectives, and hence alternative solutions. The field of case-based reasoning (Maher and Pu, 1997) emerged from the observation that designers solve new problems by drawing on knowledge gained from solving previous similar problems. As Schank (1982, p. 22) puts it, "An expert is someone who gets reminded of just the right prior experience to help him in processing his current experiences." And while those experiences may be the designer's own, they can equally well be others'.

Another approach to creativity has been taken by Maiden et al (2007a). They have been running creativity workshops to generate innovative requirements in an air traffic management (ATM) application domain. Their idea is to introduce experts in different fields into the workshop, and then invite stakeholders to identify analogies between their own field and this new one. For example, they have invited an Indian textile expert, a musician, a TV program scheduler, and a museum exhibit designer. Although not all obviously analogical domains, they sparked creative ideas for the air traffic management application. For example, participants reported that one textile design was elegant, i.e. simple, beautiful, and symmetrical. They then transferred these properties to a key area of the ATM domain – that of aircraft conflict resolution. They explored the meaning of elegance within this context, and realized that elegance is perceived differently by different controllers. From this they generated the requirement that the system should be able to accommodate different air traffic controller styles during conflict resolution.

A more pragmatic answer to this question, then, is that alternatives come from seeking different perspectives and looking at other designs. The process of inspiration and creativity can be enhanced by prompting a designer's own experience and studying others' ideas and suggestions. Deliberately seeking out suitable sources of inspiration is a valuable step in any design process. These sources may be very close to the intended new product, such as competitors' products, or they may be earlier versions of similar systems, or something completely different.

Having said this, under some circumstances the scope to consider alternative designs may be limited. Design is a process of balancing constraints and constantly trading off one set of requirements with another, and the constraints may be such that there are very few viable alternatives available. For example, if you are designing a software system to run under the

Windows operating system, then elements of the design will be prescribed because you must conform to the Windows look and feel, and to other constraints intended to make Windows programs consistent for the user. If you are producing an upgrade to an existing system, then you may want to keep the familiar elements of it and retain basically the same user experience.

Activity 9.4

Consider again the travel planner introduced at the beginning of the chapter. Reflecting on the process again, what do you think inspired your outline design? See if you can identify any elements within it that you believe are truly innovative.

Comment

Show/Hide

BOX 9.4

A Box Full of Ideas

The innovative product design company IDEO was introduced in Chapter 1. Underlying some of their creative flair is a collection of weird and wonderful engineering housed in a large flatbed filing cabinet called the TechBox (see Figure 9.4). The TechBox holds hundreds of gizmos and interesting materials, divided into categories such as: Amazing Materials, Cool Mechanisms, Interesting Manufacturing Processes, Electronic Technologies, and Thermal and Optical. Each item has been placed in the box because it represents a neat idea or a new process. Staff at IDEO take along a selection of items from the TechBox to brainstorming meetings. The items may be chosen because they provide useful visual props or possible solutions to a particular issue, or simply to provide some light relief.

Each item is clearly labeled with its name and category, but further information can be found by accessing the TechBox's online catalog. Each item has its own page detailing what the item is, why it is interesting, where it came from, and who has used it or knows more about it. Items in the box include an example of metal-coated wood, and materials with and without holes that stretch, bend, and change shape or color at different temperatures.

Each of IDEO's offices has a TechBox and each TechBox has its own curator who is responsible for maintaining and cataloging the items and for promoting its use within the office. Anyone can submit a new item for consideration, and as items become commonplace they are removed from the TechBox to make way for the next generation of fascinating curios. ■

Figure 9.4 The TechBox at IDEO

Source: Reproduced by permission of IDEO. Photo by Jorge Davies.

Dilemma

Copying for Inspiration: Is it Legal?

Designers draw on their experience of design when approaching a new project. This includes the use of previous designs that they know work, both designs they have created themselves and those that others have created. Others’ creations often spark inspiration that also leads to new ideas and innovation. This is well known and understood. However, the expression of an idea is protected by copyright, and people who infringe that copyright can be taken to court and prosecuted. Note that copyright covers the expression of an idea and not the idea itself. This means, for example, that while there are numerous MP3 players all with similar functionality, this does not represent an infringement of copyright as the idea has been expressed in different ways, and it is the expression that has been copyrighted. Copyright is free and is automatically invested in the author of something, e.g. the writer of a book or a programmer who develops a program, unless he signs the copyright over to someone else. People who produce something through their employment, such as programs or products, may have in their employment contract a statement saying that the copyright relating to anything produced in the course of that employment is automatically assigned to the employer and does not remain with the employee.

Patenting is an alternative to copyright that does protect the idea rather than the expression. There are various forms of patenting, each of which is designed to allow the inventor the chance to capitalize on an idea. It is unusual for software to be patented, since it is a long, slow, and expensive process, although there have been some examples of patenting business processes. For example, Amazon has patented its one-click purchasing process, which allows regular users simply to choose a purchase and buy it with one mouse click (US Patent No. 5960411, September 29, 1999). This is possible because the system stores its customers’ details and recognizes them when they access the Amazon site again.

In recent years, the creative commons community (creativecommons.org) has suggested more flexible licensing arrangements that allow others to reuse and extend a piece of created work, thereby supporting collaboration. In the Open Source software development movement, for example, software code is freely distributed and can be modified, incorporated into other software, and redistributed under the same open source conditions. No royalty fees are payable on any use of open source code. These movements do not replace copyright or patent law, but they provide an alternative route for the dissemination of ideas.

So the dilemma comes in knowing when it is OK to use someone else's work as a source of inspiration and when you are infringing copyright or patent law. The issues around this question are complex and detailed, and well beyond the scope of this book, but up-to-date information and examples of law cases that have been brought successfully and unsuccessfully can be found in Bainbridge (2014). ■

9.3.4 How Do You Choose Among Alternative Designs?

Choosing among alternatives is about making design decisions: Will the device use keyboard entry or a touch screen? Will the product provide an automatic memory function or not? These decisions will be informed by the information gathered about users and their tasks, and by the technical feasibility of an idea. Broadly speaking, though, the decisions fall into two categories: those that are about externally visible and measurable features, and those that are about characteristics internal to the system that cannot be observed or measured without dissecting it. For example, in a photocopier, externally visible and measurable factors include the physical size of the machine, the speed and quality of copying, the different sizes of paper it can use, and so on. Underlying each of these factors are other considerations that cannot be observed or studied without dissecting the machine. For example, the choice of materials used in a photocopier may depend on its friction rating and how much it deforms under certain conditions.

In an interactive product there are similar factors that are externally visible and measurable and those that are hidden from the users' view. For example, exactly why it takes 30 seconds for a web page to load, or why it takes an hour for a cell phone text message to arrive, will be influenced by technical decisions made when the web page or cell phone software was constructed. From the users' viewpoint the important observation is the fact that it does take 30 seconds to load or an hour to arrive.

In interaction design, the way in which the users interact with the product is considered the driving force behind the design and so we concentrate on the externally visible and measurable behavior. Detailed internal workings are still important to the extent that they affect external behavior or features.

One answer to the question posed above is that we choose between alternative designs by letting users and stakeholders interact with them and by discussing their experiences, preferences, and suggestions for improvement. This is fundamental to a user-centered approach to development. This in turn means that the designs must be available in a form that can be reasonably evaluated with users, not in technical jargon or notation that seems impenetrable to them.

One form traditionally used for communicating a design is documentation, e.g. a description of how something will work or a diagram showing its components. The trouble is that a static description cannot easily capture the dynamics of behavior, and for an interactive product we need to communicate to the users what it will be like to actually operate it.

In many design disciplines, prototyping is used to overcome potential client misunderstandings and to test the technical feasibility of a suggested design and its production. Prototyping involves producing a limited version of the product with the purpose of answering specific questions about the design's feasibility or appropriateness. Prototypes give a better impression of the user experience than simple descriptions, and there are different kinds of prototyping that are suitable for different stages of development and for eliciting different kinds of information. Prototyping is discussed in detail in Chapter 11.

Another basis on which to choose between alternatives is quality, but this requires a clear understanding of what quality means. People's views of what is a quality product vary. Whenever we use anything we have some notion of the level of quality we are expecting, wanting, or needing. Whether this level of quality is expressed formally or informally does not matter. The point is that it exists and we use it consciously or subconsciously to evaluate alternative items. For example, if you have to wait too long to download a web page, then you are likely to give up and try a different site – you are applying a certain measure of quality associated with the time taken to download the web page. If one smartphone makes it easy to access your favorite music channel while another involves several complicated key sequences, then you are likely to buy the former rather than the latter. Here, you are applying a quality criterion concerned with efficiency.

If you are the only user of a product, then you don't necessarily have to express your definition of quality since you don't have to communicate it to anyone else. However, as we have seen, most projects involve many different stakeholder groups, and you will find that each of them has a different definition of quality and different acceptable limits for it. For example, although all stakeholders may agree on targets such as ‘response time will be fast’ or ‘the menu structure will be easy to use,’ exactly what each of them means by this is likely to vary. Disputes are inevitable when, later in development, it transpires that ‘fast’ to one set of stakeholders meant ‘under a second’ while to another it meant ‘between 2 and 3 seconds.’ Capturing these different views in clear unambiguous language early in development takes you halfway to producing a product that will be well-regarded by all your stakeholders. It helps to clarify expectations, provides a benchmark against which products of the development process can be measured, and gives you a basis on which to choose among alternatives.

The process of writing down formal, verifiable – and hence measurable – usability criteria is a key characteristic of an approach to interaction design called usability engineering. This has emerged over many years and with various proponents (Whiteside et al, 1988; Nielsen, 1993). Usability engineering involves specifying quantifiable measures of product performance, documenting them in a usability specification, and assessing the product against them. One way in which this approach is used is to make changes to subsequent versions of a system based on feedback from carefully documented results of usability tests for the earlier version.

Activity 9.5

Consider the travel planner that you designed in Activity 9.1. Suggest some usability criteria that you could use to determine the planner's quality. Use the usability goals introduced in Chapter 1: effectiveness, efficiency, safety, utility, learnability, and memorability. Be as specific as possible. Check your criteria by considering exactly what you would measure and how you would measure its performance.

Then try to do the same thing for some of the user experience goals introduced in Chapter 1 (these relate to whether a system is satisfying, enjoyable, motivating, rewarding, and so on).

Comment

Show/Hide

9.3.5 How Do You Integrate Interaction Design Activities with Other Lifecycle Models?

There are several lifecycle models associated with other disciplines that contribute to interaction design (see Figure 1.4). Prominent among these lifecycle models are those associated with software engineering. Discussion about how best to integrate user-centered design and software engineering, and how to raise awareness of user-centered techniques with software engineers has been ongoing for several years, e.g. see Seffah et al (2005).

The latest, and some would argue the most promising, attempts at integration focus on a relatively recent trend in software engineering, called agile software development. Agile methods began to emerge in the late 1990s. The most well known of these are eXtreme Programming (Beck and Andres, 2005), Crystal (Cockburn, 2005), Scrum (Schwaber and Beedle, 2002), and Adaptive Software Development (ASD) (Highsmith, 2000). Dynamic Systems Development Method (DSDM) (DSDM, 2014), although established before the current agile movement, also belongs to the agile family as it adheres to the agile manifesto (reproduced below). These methods differ, but they all stress the importance of iteration, early and repeated user feedback, being able to handle emergent requirements, and striking a good balance between flexibility and structure. They also all emphasize collaboration, face-to-face communication, streamlined processes to avoid unnecessary activities, and the importance of practice over process, i.e. of getting work done.

The opening statement for the Manifesto for Agile Software Development (www.agilemanifesto.org/) is:

> **We are uncovering better ways of developing software by doing it and helping others do it. Through this work we have come to value:**
>
> **Individuals and interactions over processes and tools**
>
> **Working software over comprehensive documentation**
>
> **Customer collaboration over contract negotiation**
>
> **Responding to change over following a plan**

This manifesto is underpinned by a series of principles, which range from communication with the business through to excellence of coding and maximizing the amount of work done. The agile approach to development is particularly interesting from the point of view of interaction design because it incorporates tight iterations and feedback, and collaboration with the customer (e.g. Armitage, 2004; Sharp et al, 2006). For example, in eXtreme[1] Programming (XP), each iteration is between one and three weeks, with a product of value being delivered at the end of each iteration. Also, XP stipulates that the customer should be on site with developers. In practice, the customer role is usually taken by a team rather than one person (Martin et al, 2009), and integration is far from straightforward (Ferreira et al, 2012). Several companies have integrated agile methods with interaction design practices to produce better quality products. Chapter 12 discusses this aspect in more detail.

One of the main proponents for integrating user-centered design and agile development in practice, Jeff Patton, has articulated several patterns of common behaviors for successful agile product development, including the importance of designers being part of the team, using parallel tracks with user research, doing just enough user research, modeling and design up front, buying design time when developers are working on complex engineering stories, and cultivating a user group for continuous user validation. Although they first appeared a while ago, these still represent good and practical advice.

> Link to see Jeff Patton's common behaviours for successful agile product development at http://tinyurl.com/3rfnm2

These recommendations are echoed in a report from the Nielsen Norman Group (Nielsen Norman Group, 2014) who conducted a study of agile projects, investigating case studies in depth from 16 organizations who care about user experience and who have also embraced an agile approach to development. They make two main recommendations for success: have development and design running in separate tracks (as described in Figure 12.2 in Chapter 12) and maintain a coherent vision of the interface architecture (also see Kollmann et al, 2009).

Assignment

Nowadays, timepieces (such as clocks, wristwatches, etc.) have a variety of functions. They not only tell the time and date but they can speak to you, remind you when it's time to do something, and provide a light in the dark, among other things. Mostly, the interface for these devices, however, shows the time in one of two basic ways: as a digital number such as 23:40 or through an analog display with two or three hands – one to represent the hour, one for the minutes, and one for the seconds.

In this assignment, we want you to design an innovative timepiece for your own use. This could be in the form of a wristwatch, a mantelpiece clock, an electronic clock, or any other kind of timepiece you fancy. Your goal is to be inventive and exploratory. We have broken this assignment down into the following steps to make it clearer:

a. Think about the interactive product you are designing: what do you want it to do for you? Find three to five potential users and ask them what they would want. Write a list of requirements for the clock, together with some usability criteria and user experience criteria based on the definitions in Chapter 1.

b. Look around for similar devices and seek out other sources of inspiration that you might find helpful. Make a note of any findings that are interesting, useful, or insightful.

c. Sketch out some initial designs for the timepiece. Try to develop at least two distinct alternatives that both meet your set of requirements.

d. Evaluate the two designs, using your usability criteria and by role playing an interaction with your sketches. Involve potential users in the evaluation, if possible. Does it do what you want? Is the time or other information being displayed always clear? Design is iterative, so you may want to return to earlier elements of the process before you choose one of your alternatives.

Take a Quickvote on Chapter 9:
www.id-book.com/quickvotes/chapter9

Summary

In this chapter, we have looked at user-centered design and the process of interaction design, i.e. what is user-centered design, what activities are required in order to design an interactive product, and how are these activities related. A simple interaction design lifecycle model consisting of four activities was introduced and issues surrounding the involvement and identification of users, generating alternative designs, evaluating designs, and integrating user-centered concerns with other lifecycles were discussed.

Key points

- The interaction design process consists of four basic activities: establishing requirements, designing alternatives that meet those requirements, prototyping the designs so that they can be communicated and assessed, and evaluating them.
- User-centered design rests on three principles: early focus on users and tasks, empirical measurement, and iterative design. These principles are also key for interaction design.
- Involving users in the design process helps with expectation management and feelings of ownership, but how and when to involve users is a matter of dispute.
- Before you can begin to establish requirements, you must understand who the users are and what their goals are in using the product.
- Looking at others' designs and involving other people in design provides useful inspiration and encourages designers to consider alternative design solutions, which is key to effective design.
- Usability criteria, technical feasibility, and users' feedback on prototypes can all be used to choose among alternatives.
- Prototyping is a useful technique for facilitating user feedback on designs at all stages.
- Integrating interaction design activities with other lifecycle models requires careful planning.

Further Reading

Greenbaum, J. and Kyng, M. (eds) (1991) Design at Work: Co-operative

design of computer systems. Lawrence Erlbaum. This classic book is a good collection of papers about the co-design of software systems: both why it is worthwhile and early experiences of how to do it.

Highsmith, J. (2002) Agile Software Development Ecosystems. Addison-Wesley. This book introduces the main agile methods and their proponents. Highsmith explains the motivation behind the agile approach to development and extracts some common themes. The book includes some case studies, and how you the reader can go about developing your own agile method that suits your own particular environment.

Kelley, T., with Littman, J. (2004) The Art of Innovation. Profile Books. Tom Kelley is general manager of IDEO. In this book, Kelley explains some of the innovative techniques used at IDEO, but more importantly he talks about the culture and philosophy underlying IDEO's success. There are some useful practical hints in here as well as an informative story about building and maintaining a successful design company.

Nielsen, J. (1993) Usability Engineering. Morgan Kaufmann. This is a seminal book on usability engineering. If you want to find out more about the philosophy, intent, history, or pragmatics of usability engineering, then this is the main text.

Seffah, A., Gulliksen, J. and Desmarais, M. C. (2005) Human-Centered Software Engineering. Springer. This book is an edited collection of papers focusing on issues relating to the integration of user-centered design and software engineering. It is split into five parts: introduction; requirements, scenarios, and use cases; principles, myths, and challenges; user-centered design, unified, and agile processes; and user-centered design knowledge and UI design patterns. If you want to pursue the thorny issue of how to bring these two communities together, then this is a good starting point.

Sommerville, I. (2010) Software Engineering (9th edn). Addison-Wesley. If you are interested in pursuing the software engineering aspects of the lifecycle models section, then this book provides a useful overview of the main models and their purpose.

Interview with Ellen Gottesdiener

Ellen Gottesdiener, EBG Consulting CEO/Principal, helps organizations discover and deliver the right product at the right time. Ellen is an internationally recognized leader in the collaborative convergence of requirements management + product management + project management. In addition to working with clients, Ellen presents, writes, and trains globally. She is the author of three acclaimed books: Discover to Deliver: Product Planning and Analysis – co-authored with Mary Gorman (2012), **Requirements by Collaboration** (Addison-Wesley, 2002), and **The Software Requirements Memory Jogger** (Goal/QPC, 2005). Visit these websites for resources and more: www.ebgconsulting.com and www.DiscoverToDeliver.com

What are requirements?

Product requirements are needs that must be satisfied to achieve a goal, solve a problem, or take advantage of an opportunity. The word 'requirement' literally means something that is absolutely, positively, without question, necessary. Product requirements need to be defined in sufficient detail for planning and development. But before going to that effort and expense, are you sure they are not only must-haves but also the right and relevant requirements?

To arrive at this level of certainty, stakeholders ideally start by exploring the product's options. An option represents a potential characteristic, facet, or quality of the product. Stakeholders, who I like to refer to as product partners, use expansive thinking to surface a range of options that could fulfill the vision. Then they collaboratively analyze the options and collectively select options, based on value.

Every product has multiple dimensions, seven in fact. Discovering options for each of the 7 Product Dimensions yields a comprehensive, realistic view of the product. They are as follows.

Product	Description

dimension	
User	Users interact with the product
Interface	The product connects users, systems, and devices
Action	The product provides capabilities to users
Data	The product includes a repository of data and useful information
Control	The product enforces constraints
Environment	The product conforms to physical properties and technical platforms
Quality attribute	The product has certain properties that qualify its operation and development

(You can view and download an image here: http://tinyurl.com/m59lo9x

You want to engage diverse stakeholders across the full product lifecycle, from birth to retirement and demise.

So how do you know who the stakeholders are?

Successful teams work hand in hand with their stakeholders as product partners, defining value and then actively discovering – and delivering – high-value solutions. This goes beyond feature requests and requirements documents – beyond user stories and product backlogs – beyond the push-pull of competing interests. It's a partnership where the ideas, perspectives and experiences of three different stakeholder groups converge. The result? Product partners who collaborate to discover and deliver value.

A product partnership includes people from three realms: customer, business, and technology. Each offers a unique perspective and has its own ideas of what is valuable.

The customer partners represent users, buyers, and advisers – people or systems that interface with the product, choose to buy it, or influence others to buy it. They tend to value improved productivity, heightened efficiency, greater speed, entertainment, and similar benefits.

Business partners represent the people in your organization who authorize, champion, or support the product or who provide subject matter expertise. They find value in improving market position, complying with regulations, achieving a business case, reducing overhead costs, enhancing internal performance, and so on.

Technology partners (your delivery team, internal or third parties) design, deliver, test, and support the product or advise those who do. They may value building a high-quality product, offering smooth, continual delivery, adopting a stable architecture, and the like.

This mix of partners and perspectives is essential, no matter what kind of delivery method you adopt (agile, traditional, hybrid, or another approach). For the partnership to work, these three disparate groups must collaborate to reach their shared goal: discover and deliver value.

How do you go about identifying requirements?

Requirements discovery is highly proactive, interactive, and, well, sometimes hyperactive! You are engaged in eliciting, analyzing, specifying, prototyping, and testing. And in the best practices we've been involved in, you are constantly discovering (a.k.a. identifying) product needs. It's not a ‘one-and-done’ activity.

Elicitation includes interviews, existing documentation study, exploratory prototypes, facilitated workshops, focus groups, observation (including apprenticing, contextual inquiry, and ethnography), surveys (and other research-based techniques), and user task analysis (including storyboarding and scenario analysis). There are a number of specific techniques within each of these general categories, and some techniques overlap. Analyzing involves using lightweight models, often combined with specifications, which are often in the form of acceptance tests or prototypes or both.

It's not enough to get the right people together and ask the right questions. To communicate efficiently and effectively about how to deliver, product partners need a focused way to communicate and make decisions together.

What we've found in our work is that the most efficient and effective discovery mechanism is a collaborative approach called the ‘structured conversation.’ In a structured conversation, the product partners first explore possible requirements (options) for their next increment. They do this within and across the 7 Product Dimensions. This enables product partners to collaboratively and creatively explore a range of possibilities. This expansive thinking opens up product innovation, experimentation, and mutual learning.

They then evaluate these many options in terms of value. This means having shared understanding of what value really means at that point in time. Once they have narrowed the list of options through the evaluation process, they confirm how they will verify and validate these candidate solutions with unambiguous acceptance criteria. The validation includes how to test that they delivered the right requirements, and that they achieved the anticipated value from each delivery.

How do you know when you have collected enough requirements to go on to the next step?

I often get asked by clients how I know when I have a complete set of requirements. I think it's more important to ask whether you are going after the right requirements.

I characterize a ‘right requirement’ as one that is:

1. **Just in time, just enough.** It is essential for achieving the business objectives, in this time period.
2. **Realistic.** It is capable of being delivered with the available resources.
3. **Clearly and unambiguously defined.** Acceptance criteria exist that all partners understand and will use to verify and validate the product.
4. **Valuable.** It is indispensable for achieving the anticipated outcomes for the next delivery cycle.

What's the hardest thing about establishing requirements?

People.

Seriously.

We humans are non-linear creatures. We are unpredictable, fickle, and (as adults) often inflexible. As requirements seekers, we swim in a stew of complex, ever-evolving human systems that interoperate as we do our requirements work.

To top that off, most products’ requirements are fraught with complexity and interdependency; there are truly wicked problems, whereby the problem space overlaps with solution space. As Frederick Brooks said [in his essay No Silver Bullet], “the hardest single part of building a software system is deciding precisely what to build.”

You can't make those decisions without trust. And trust is not an easy thing to build.

Do you have any other tips for establishing requirements?

Employ small, tightly wound cycles of requirements-build-release. Use interactive and incremental (a.k.a. agile) practices to get feedback early and often on the smallest viable releases.

For successful requirements discovery, you need to keep the focus on value – the why behind the product and the value considerations of the product partners. During discovery work, some people view a specific option as a 'requirement' for the next delivery cycle, whereas others consider it a 'wish list' item for a future release.

Such was the case in a recent release planning workshop. The team wrestled with a particular option, questioning if it could deliver enough value to justify the cost to develop it. The product champion explained why the option was a requirement – without it the organization was at risk for regulatory noncompliance. Once the others understood this, they all agreed it would be included in the release.

In the end, requirements work is human-centric, and central to successful product delivery. At the same time, the subject matter and content of product requirements is complex. Thus, requirements work is the hardest part of software and will always be.

To be successful with requirements, engineer collaboration into requirements work.

Personally, I'm excited and grateful for the growing recognition of the value of collaboration and the explosion in interest in collaborative practices in the product and software development community – because: collaboration works! ■

Note

1. The method is called 'extreme' because it pushes a key set of good practices to the limit, i.e. it is good practice to test often, so in XP the development is test-driven and a complete set of tests is executed many times a day; it is good practice to talk to people about their requirements, so rather than having weighty documentation, XP reduces documentation to a minimum, thus forcing communication, and so on.

CHAPTER 11
DESIGN, PROTOTYPING, AND CONSTRUCTION

Objectives

The main aims of this chapter are to:

- Describe prototyping and different types of prototyping activities.
- Enable you to produce simple prototypes from the models developed during the requirements activity.
- Enable you to produce a conceptual model for a product and justify your choices.
- Explain the use of scenarios and prototypes in design.
- Introduce physical computing kits and software development kits, and their role in construction.

11.1 Introduction

Design activities begin once some requirements have been established. The design emerges iteratively, through repeated design–evaluation–redesign cycles involving users. Broadly speaking, there are two types of design: conceptual and concrete. The former is concerned with developing a conceptual model that captures what the product will do and how it will behave, while the latter is concerned with details of the design such as menu structures, haptic feedback, physical widgets, and graphics. As design cycles become shorter, the distinction between these two becomes blurred, but they are worth distinguishing because each emphasizes a different set of design concerns.

For users to evaluate the design of an interactive product effectively, designers must prototype their ideas. In the early stages of development, these prototypes may be made of paper and cardboard, or ready-made components pulled together to allow evaluation, while as design progresses, they become more polished, compact, and robust so that they resemble the final product.

Broadly speaking, the design process may start from two distinct situations: when starting from scratch or when modifying an existing product. Much of design comes from the latter, and it is tempting to think that additional features can be added, or existing ones tweaked, without extensive investigation, prototyping, or evaluation. Although prototyping and evaluation activities can be reduced if changes are not significant, they are still valuable and should not be skipped.

In Chapter 10, we discussed some ways to identify user needs and establish requirements. In this chapter, we look at the activities involved in progressing a set of requirements through the cycles of prototyping to construction. We begin by explaining the role and techniques of prototyping and then explain how prototypes may be used in the design process. We end with an exploration of physical computing and software development kits (SDKs) that provide a basis for construction.

11.2 Prototyping

It is often said that users can't tell you what they want, but when they see something and get to use it, they soon know what they don't want. Having established some requirements, the next step is to try out design ideas through prototyping and evaluation cycles.

11.2.1 What Is a Prototype?

A prototype is one manifestation of a design that allows stakeholders to interact with it and to explore its suitability; it is limited in that a prototype will usually emphasize one set of product characteristics and de-emphasize others. When you hear the term prototype, you may imagine a scale model of a building or a bridge, or a piece of software that crashes every few minutes. A prototype can also be a paper-based outline of a display, a collection of wires and ready-made components, an electronic picture, a video simulation, a complex piece of software and hardware, or a three-dimensional mockup of a workstation.

In fact, a prototype can be anything from a paper-based storyboard through to a complex piece of software, and from a cardboard mockup to a molded or pressed piece of metal. For example, when the idea for the PalmPilot was being developed, Jeff Hawkin (founder of the company) carved up a piece of wood about the size and shape of the device he had imagined. He used to carry this piece of wood around with him and pretend to enter information into it, just to see what it would be like to own such a device (Bergman and Haitani, 2000). This is an example of a very simple (some might even say bizarre) prototype, but it served its purpose of simulating scenarios of use. Advances in 3D printer technologies, coupled with reducing prices, have increased their use in design. It is now possible to take a 3D model from a software package and print a prototype. Even soft toys and chocolate may be ‘printed’ in this way (see Figure 11.1).

(a)

(b)

(c)

Figure 11.1 (a) Color output from a 3D printer: all the gears and rods in this model were 'printed' in one pass from bottom to top, and when one gear is turned, the others turn too. (b) James Bond's Aston Martin in Skyfall was in fact a 3D-printed model (http://www.telegraph.co.uk/technology/news/9712435/The-names-Printing-3D-Printing.html). (c) A teddy bear 'printed' from a wireframe design http://www.disneyresearch.com/project/printed-teddy-bears/

Source: (a) The Computer Language Company, Inc., courtesy of Alan Freedman (b) Courtesy of voxeljet and Propshop Modelmakers Ltd (c) Courtesy of Scott Hudson, Human–Computer Interaction Institute, Carnegie Mellon University.

Video showing a teddy bear being 'printed' is available at http://www.disneyresearch.com/project/printed-teddy-bears/

11.2.2 Why Prototype?

Prototypes are useful when discussing or evaluating ideas with stakeholders; they are a communication device among team members, and an effective way for designers to explore design ideas. The activity of building prototypes encourages reflection in design, as described by Schön (1983) and is recognized by designers from many disciplines as an important aspect of design.

Prototypes answer questions and support designers in choosing between alternatives. Hence, they serve a variety of purposes: for example, to test out the technical feasibility of an idea, to clarify some vague requirements, to do some user testing and evaluation, or to check that a certain design direction is compatible with the rest of product development. The purpose of your prototype will influence the kind of prototype you build. So, for example, if you want to clarify how users might perform a set of tasks and whether your proposed design would support them in this, you might produce a paper-based mockup. Figure 11.2 shows a paper-based prototype of a handheld device to help an autistic child communicate. This prototype shows the intended functions and buttons, their positioning and labeling, and the overall shape of the device, but none of the buttons actually work. This kind of prototype is sufficient to investigate scenarios of use and to decide, for example, whether the button images and labels are appropriate and the functions sufficient, but not to test whether the speech is loud enough or the response fast enough. In the development of SITU, a smart food nutrition scale and tablet application, a range of prototypes and representations were used from initial idea to final product. These included screen sketches, paper and cardboard mockups, wireframes, and many post-its.

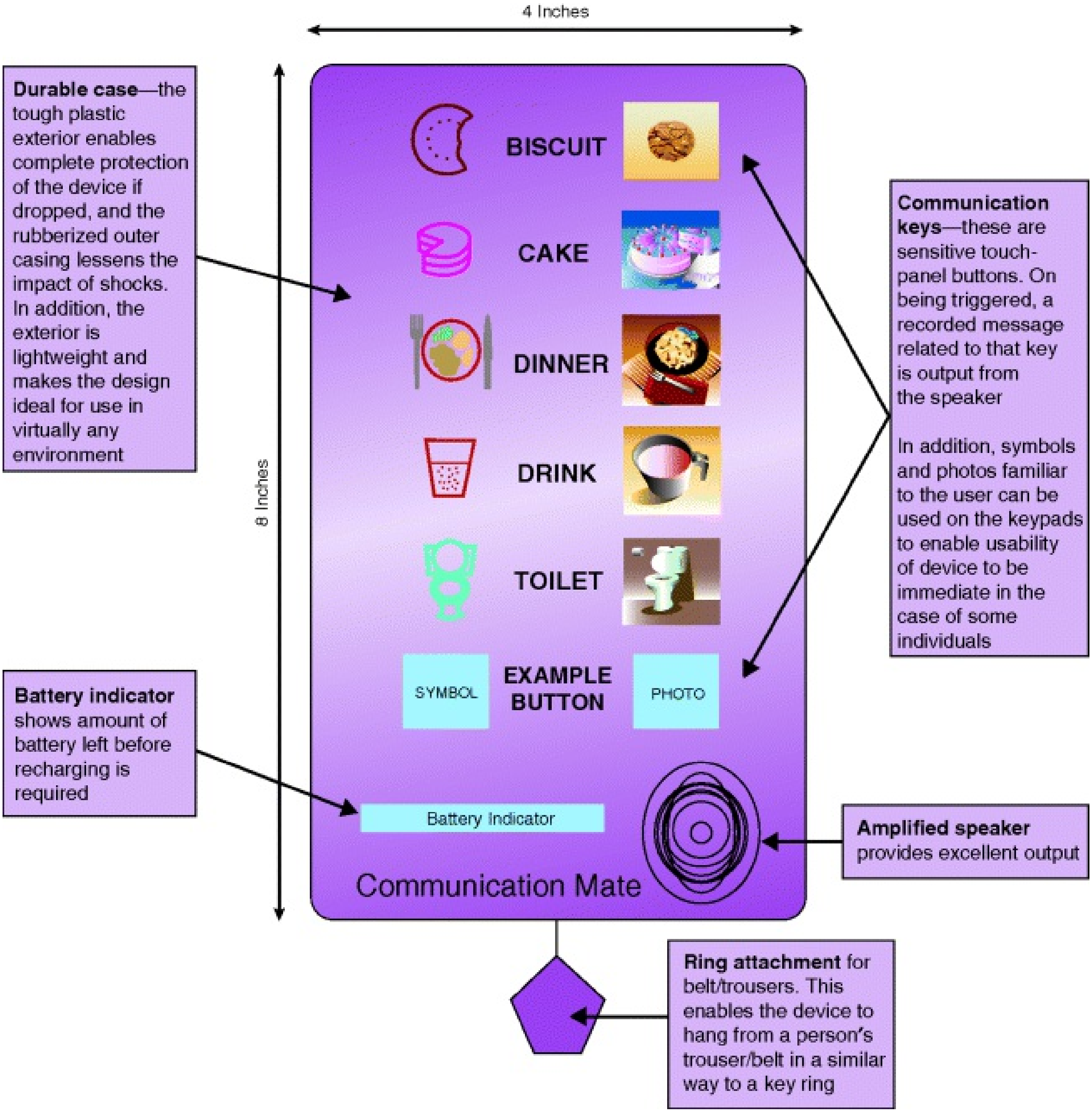


Figure 11.2 A paper-based prototype of a handheld device to support an autistic child

Source: Reprinted by permission of Sigil Khwaja.

Link to see the full story of SITU at
https://www.kickstarter.com/projects/situ/situ-smart-food-nutrition-scale

Saffer (2010) distinguishes between a product prototype and a service prototype, where the latter involves role playing and people as an integral part of the prototype as well as the product itself. Service prototypes are sometimes captured as video scenarios and used in a similar way to the scenarios introduced in Chapter 10.

11.2.3 Low-Fidelity Prototyping

A low-fidelity prototype does not look very much like the final product and does not provide the same functionality. For example, it may use very different materials, such as paper and cardboard rather than electronic screens and metal, it may perform only a limited set of functions, or it may only represent the functions and not perform any of them. The lump of wood used to prototype the PalmPilot described above is a low-fidelity prototype.

Low-fidelity prototypes are useful because they tend to be simple, cheap, and quick to produce. This also means that they are simple, cheap, and quick to modify so they support the exploration of alternative designs and ideas. This is particularly important in early stages of development, during conceptual design for example, because prototypes that are used for exploring ideas should be flexible and encourage exploration and modification. Low-fidelity prototypes are not meant to be kept and integrated into the final product. They are for exploration only.

Storyboarding.

Storyboarding is one example of low-fidelity prototyping that is often used in conjunction with scenarios, as described in Chapter 10. A storyboard consists of a series of sketches showing how a user might progress through a task using the product under development. It can be a series of screen sketches or a series of scenes showing how a user can perform a task using an interactive device. When used in conjunction with a scenario, the storyboard provides more detail and offers stakeholders a chance to role-play with a prototype, interacting with it by stepping through the scenario. The example storyboard shown in Figure 11.3 depicts a person (Christina) using a new mobile device for exploring historical sites. This example shows the context of use for this device and how it might support Christina in her quest for information about the pottery trade at The Acropolis in Ancient

Greece.

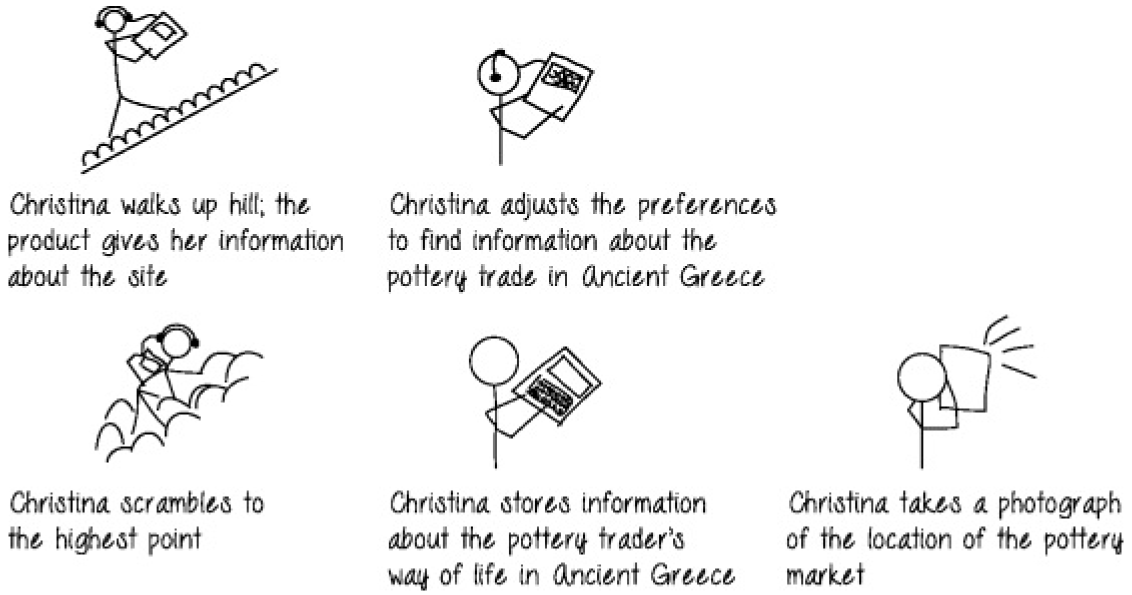


Figure 11.3 An example storyboard for a mobile device to explore ancient sites such as The Acropolis

Sketching.

Low-fidelity prototyping often relies on hand-drawn sketches, and many people find it difficult to engage in this activity because they are inhibited about the quality of their drawing, but as Greenberg et al (2012) put it, “Sketching is not about drawing. Rather, it is about design” (p. 7). You can get over any inhibition by devising your own symbols and icons and practicing them – referred to by Greenberg et al as a ‘sketching vocabulary’ (p. 85). They don't have to be anything more than simple boxes, stick figures, and stars. Elements you might require in a storyboard sketch, for example, include digital devices, people, emotions, tables, books, etc., and actions such as give, find, transfer, and write. If you are sketching an interface design, then you might need to draw various icons, dialog boxes, and so on. Some simple examples are shown in Figure 11.4. The next activity requires other sketching symbols, but they can still be drawn quite simply. Baskinger (2008) provides further tips for those new to sketching.

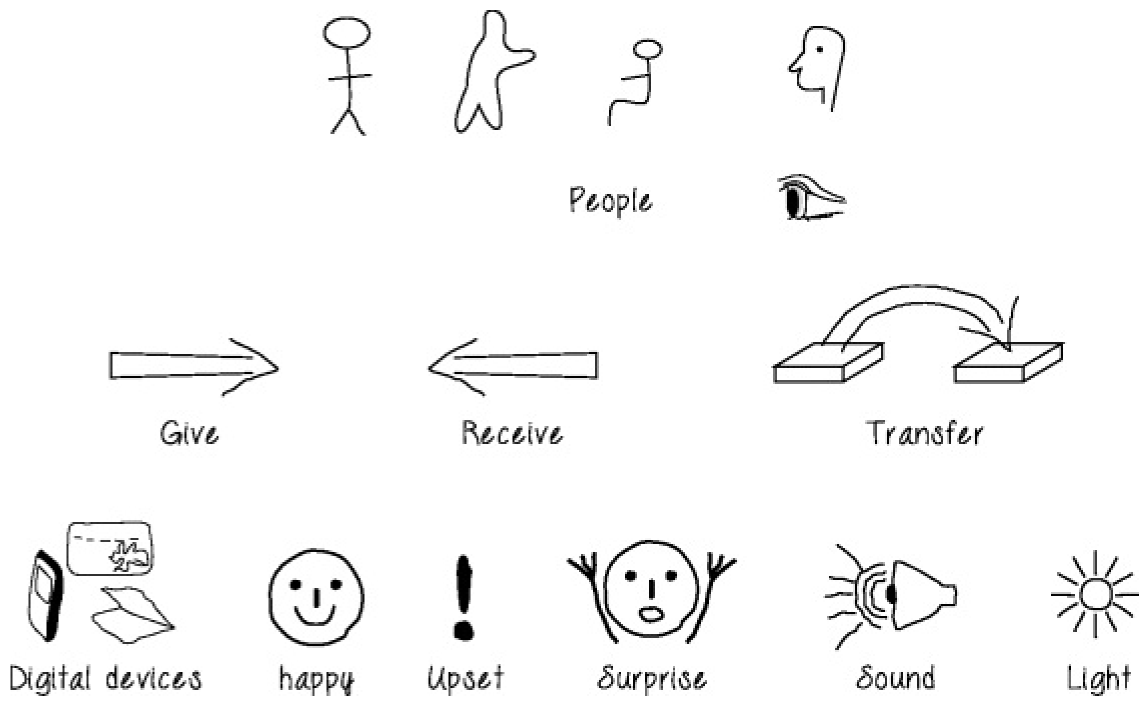


Figure 11.4 Some simple sketches for low-fidelity prototyping

Prototyping with index cards.

Using index cards (small pieces of cardboard about 3 × 5 inches) is a successful and simple way to prototype an interaction, and is used often when developing websites. Each card represents a screen or one element of the interaction. In user evaluations, the user can step through the cards, pretending to perform the task while interacting with the cards. A more detailed example of this kind of prototyping is given in Section 11.6.2.

Activity 11.1

Produce a storyboard that depicts how to fill a car with gas (petrol).

Comment

Show/Hide

Wizard of Oz.

Another low-fidelity prototyping method called Wizard of Oz assumes that you have a software-based prototype. In this technique, the user interacts with the software as though interacting with the product. In fact, however, a human operator simulates the software's response to the user. The method takes its name from the classic story of the little girl who is swept away in a storm and finds herself in the Land of Oz (Baum and Denslow, 1900). The Wizard of Oz is a small shy man who operates a large artificial image of himself from behind a screen where no one can see him. The Wizard of Oz style of prototyping has been used successfully for various applications, including PinTrace, a robotic system that helps surgeons to position orthopedic pins accurately during the surgery of hip fractures (Molin, 2004), and to identify gestures for full body interaction with digital games (Norton et al, 2010).

11.2.4 High-Fidelity Prototyping

A high-fidelity prototype looks like the final product and/or provides more functionality than a low-fidelity prototype. For example, a prototype of a software system developed in Visual Basic is higher fidelity than a paper-based mockup; a molded piece of plastic with a dummy keyboard is a higher-fidelity prototype of the PalmPilot than the lump of wood. High-fidelity prototyping is useful for selling ideas to people and for testing out technical issues.

High-fidelity prototypes can be developed by modifying and integrating existing components – both hardware and software. In robotics this approach has been called tinkering (Hendriks-Jansen, 1996) while in software development it has been referred to as Opportunistic System Development (Ncube et al, 2008). Banzi (2009) comments that: “Reusing existing technology is one of the best ways of tinkering. Getting cheap toys or old discarded equipment and hacking them to make them do something new is one of the best ways to get great results.” Bird et al (2009) describe how they used this approach to develop a tactile vision sensory substitution system, i.e. a system that translates a camera image of the user's environment into tactile stimulation on their body.

11.2.5 Compromises in Prototyping

By their very nature, prototypes involve compromises: the intention is to produce something quickly to test an aspect of the product. Lim et al (2008) suggest an anatomy of prototyping that structures the different aspects of a prototype and what it aims to achieve. Their ideas are expanded in Box 11.1. The kind of questions that any one prototype can answer is limited, and the prototype must be built with the key issues in mind. In low-fidelity prototyping, it is fairly clear that compromises have been made. For example, with a paper-based prototype an obvious compromise is that the device doesn't actually work! For software-based prototyping, some of the compromises will still be fairly clear; for example, the response speed may be slow, or the look and feel may not be finalised, or only a limited amount of functionality may be available.

BOX 11.1

The Anatomy of Prototyping: Filters and Manifestations

Lim et al (2008) propose a view of prototypes which focuses on their role as filters, i.e. to emphasize specific aspects of a product being explored by the prototype, and as manifestations of designs, i.e. as tools to help designers develop their design ideas through external representations.

They suggest three key principles in their view of the anatomy of prototypes:

1. Fundamental prototyping principle: Prototyping is an activity with the purpose of creating a manifestation that, in its simplest form, filters the qualities in which designers are interested, without distorting the understanding of the whole.
2. Economic principle of prototyping: The best prototype is one that, in the simplest and the most efficient way, makes the possibilities and limitations of a design idea visible and measurable.
3. Anatomy of prototypes: Prototypes are filters that traverse a design space and are manifestations of design ideas that concretize and externalize conceptual ideas.

Lim et al identify several dimensions of filtering and of manifestation that may be considered when developing a prototype, although they point out that these dimensions are not complete but provide a useful starting point for consideration of prototype development. These are shown in Tables 11.1 and 11.2. □

Table 11.1 Example variables of each filtering dimension

Filtering dimension	Example variables
Appearance	size; color; shape; margin; form; weight; texture; proportion; hardness; transparency; gradation; haptic; sound
Data	data size; data type (e.g. number; string; media); data use; privacy type; hierarchy; organization
Functionality	system function; users' functionality need
Interactivity	input behavior; output behavior; feedback behavior; information behavior
Spatial structure	arrangement of interface or information elements; relationship among interface or information elements – which can be either two- or three-dimensional, intangible or tangible, or mixed

Table 11.2 The definition and variables of each manifestation dimension

Manifestation dimension	Definition	Example variables
Material	Medium (either visible or invisible) used to form a prototype	Physical media, e.g. paper, wood, and plastic; tools for manipulating physical matters, e.g. knife, scissors, pen, and sandpaper; computational prototyping tools, e.g. Macromedia Flash and Visual Basic; physical computing tools, e.g. Phidgets and Basic Stamps; available existing artifacts, e.g. a beeper to simulate a heart attack
Resolution	Level of detail or sophistication of what is manifested (corresponding to fidelity)	Accuracy of performance, e.g. feedback time responding to an input by a user (giving user feedback in a paper prototype is slower than in a computer-based one); appearance details; interactivity details; realistic versus faked data
Scope	Range of what is covered to be manifested	Level of contextualization, e.g. website color scheme testing with only color scheme charts or color schemes placed in a website layout structure; book search navigation usability testing with only the book search related interface or the whole navigation interface

Two common compromises that often must be traded against each other are breadth of functionality provided versus depth. These two kinds of prototyping are called horizontal prototyping (providing a wide range of functions but with little detail) and vertical prototyping (providing a lot of detail for only a few functions).

Other compromises won't be obvious to a user of the system. For example, the internal structure of the product may not have been carefully designed, and the prototype may contain spaghetti code or be badly partitioned. One of the dangers of producing functional prototypes, i.e. ones that users can interact with automatically, is that the prototype can appear to be the final product. Another is that developers may consider fewer alternatives because the prototype works and users like it. However, the compromises made in order to produce the prototype must not be ignored, particularly those that are less obvious from the outside. For a good-quality product, good engineering principles must be adhered to.

"THEN IN HERE WE DO A CLAY MOCK-UP OF THE COMPUTER MODEL"

BOX 11.2

When to use high fidelity and when to use low fidelity prototypes

Table 11.3 summarizes proclaimed advantages and disadvantages of high- and low-fidelity prototyping. Component kits and pattern libraries for interface components (see Section 11.7 and Chapter 12) make it quite easy to develop polished functional prototypes quickly, but there is a strong case for the value of low-fidelity prototypes, such as paper-based sketches, sticky note designs, and storyboarding, to explore initial ideas. Paper prototyping, for example, is used in game design (Gibson, 2014), website development, and product design (Case study 11.1). Both high- and low-fidelity prototypes provide useful feedback during evaluation and design iterations. For example, Dhillon et al (2011) found that a low-fidelity video prototype elicited comparable user feedback as a high-fidelity one, but was quicker and cheaper to produce. In the context of usability testing, most studies have found that there is no difference between the low- and high-fidelity approach in terms of user feedback (Sauer and Sonderegger, 2009), although they present some evidence that medium-fidelity prototypes are viewed as being less attractive than high- or low-fidelity ones. When exploring issues of content and structure, low-fidelity prototyping is preferable simply on the basis of cost, with the caveat that designers must be careful not to design technically infeasible capabilities on paper (Holmquist, 2005). The overriding consideration is the purpose of the prototype, and what level of fidelity is needed in order to get useful feedback. □

Table 11.3 Advantages and disadvantages of low- and high-fidelity prototypes

Type	Advantages	Disadvantages
Low-fidelity prototype	Lower development cost Evaluates multiple design concepts Useful communication device Addresses screen layout issues Useful for identifying market requirements Proof of concept	Limited error checking Poor detailed specification to code to Facilitator-driven Limited utility after requirements established Limited usefulness for usability tests Navigational and flow limitations
High-fidelity prototype	Complete functionality Fully interactive User-driven Clearly defines navigational scheme Use for exploration and test Look and feel of final product Serves as a living specification Marketing and sales tool	More resource-intensive to develop Time-consuming to create Inefficient for proof-of-concept designs Not effective for requirements gathering

Case Study 11.1

Paper prototyping as a core tool in the design of cell phone user interfaces

Paper prototyping is being used by cell phone and tablet companies as a core part of their design process (see Figure 11.6). There is much competition in the industry, demanding ever more new concepts. Mobile devices are feature-rich. They include mega-pixel cameras, music players, media galleries, downloaded applications, and more. This requires designing interactions that are complex, but are clear to learn and use. Paper prototyping offers a rapid way to work through every detail of the interaction design across multiple applications.

Figure 11.6 Prototype developed for cell phone user interface

Mobile device projects involve a range of disciplines – all with their own viewpoint on what the product should be. A typical project may include programmers, project managers, marketing experts, commercial managers, handset manufacturers, user experience specialists, visual designers, content managers, and network specialists. Paper prototyping provides a vehicle for everyone involved to be part of the design process – considering the design from multiple angles in a collaborative way.

The case study on the website describes the benefits of using paper prototyping from the designer's viewpoint, while considering the bigger picture of its impact across the entire project lifecycle. It starts by explaining the problem space and how paper prototyping is used as an integrated part of user interface design projects for European and US-based mobile operator companies. The case study uses project examples to illustrate the approach and explains step by step how the method can be used to include a range of stakeholders in the design process – regardless of their skill set or background. The case study offers exercises so you can experiment with the approach yourself. ■

Although prototypes will have undergone extensive user evaluation, they will not necessarily have been subjected to rigorous quality testing for other characteristics such as robustness and error-free operation. Building a product to be used by thousands or millions of people running on various platforms and under a wide range of circumstances requires a different testing regime than producing a quick prototype to answer specific questions.

The next Dilemma box discusses two different development philosophies. In evolutionary prototyping, a prototype evolves into the final product. Throwaway prototyping uses the prototypes as stepping stones towards the final design. In this case, the prototypes are thrown away and the final product is built from scratch. If an evolutionary prototyping approach is to be taken, the prototypes should be subjected to rigorous testing along the way; for throwaway prototyping such testing is not necessary.

DILEMMA

Prototyping versus engineering

Low-fidelity prototypes are not integrated into the final product. In contrast, high-fidelity prototypes can be and so present developers with a dilemma. They can choose to either build the prototype with the intention of throwing it away after it has fulfilled its immediate purpose, or build a prototype with the intention of evolving it into the final product.

The compromises made when producing a prototype must not be ignored – whatever those compromises were. However, when a project team is under pressure, it can become tempting to pull together a set of existing prototypes as the final product. After all, many hours of development will have been spent developing them, and evaluation with the client has gone well, so isn't it a waste to throw it all away? Basing the final product on prototypes in this way will simply store up testing and maintenance problems for later on: in short, this is likely to compromise the quality of the product.

Evolving the prototype into the final product through a defined process of engineering can lead to a robust final product, but this must be clearly planned from the beginning.

On the other hand, if the device is an innovation, then being first to market with a 'good enough' product may be more important for securing your market position than having a very high-quality product that reaches the market two months after your competitors'. ■

11.3 Conceptual Design

Conceptual design is concerned with transforming requirements into a conceptual model. Designing the conceptual model is fundamental to interaction design, yet the idea of a conceptual model can be difficult to grasp. One of the reasons for this is that conceptual models take many different forms and it is not possible to provide a definitive detailed characterization of one. Instead, conceptual design is best understood by exploring and experiencing different approaches to it, and the purpose of this section is to provide you with some concrete suggestions about how to go about doing this.

A conceptual model is an outline of what people can do with a product and what concepts are needed to understand how to interact with it. The former will emerge from the current functional requirements; possibly it will be a subset of them, possibly all of them, and possibly an extended version of them. The concepts needed to understand how to interact with the product depend on a variety of issues related to who the user will be, what kind of interaction will be used, what kind of interface will be used, terminology, metaphors, application domain, and so on. The first step in getting a concrete view of the conceptual model is to steep yourself in the data you have gathered about your users and their goals and try to empathize with them. From this, a picture of what you want the users' experience to be when using the new product will emerge and become more concrete. This process is helped by considering the issues in this section, and by using scenarios and prototypes to capture and experiment with ideas. Mood boards (traditionally used in fashion and interior design) may be used to capture the desired feel of a new product (see Figure 11.7). This is informed by results from the requirements activity and considered in the context of technological feasibility.

Figure 11.7 An example mood board

Source: Image courtesy of The Blog Studio www.theblogstudio.com.

There are different ways to achieve empathy with users. For example, Beyer and Holtzblatt (1998), in their method Contextual Design, recommend holding review meetings within the team to get different peoples' perspectives on the data and what they observed. This helps to deepen understanding and to expose the whole team to different aspects. Ideas will emerge as this extended understanding of the requirements is established, and these can be tested against other data and scenarios, discussed with other design team members, and prototyped for testing with users. Other ways to understand the users' experience are described in Box 11.3.

Key guiding principles of conceptual design are:

- Keep an open mind but never forget the users and their context.
- Discuss ideas with other stakeholders as much as possible.
- Use prototyping to get rapid feedback.
- Iterate, iterate, and iterate.

11.3.1 Developing an Initial Conceptual Model

Some elements of a conceptual model will derive from the requirements for the product. For example, the requirements activity will have provided information about the concepts involved in a task and their relationships, e.g. through task descriptions and analysis. Immersion in the data and attempting to empathize with the users as described above will, together with the requirements, provide information about the product's user experience goals, and give you a good understanding of what the product should be like. In this section we discuss approaches which help in pulling together an initial conceptual model. In particular, we consider:

- Which interface metaphors would be suitable to help users understand the product?
- Which interaction type(s) would best support the users' activities?
- Do different interface types suggest alternative design insights or options?

BOX 11.3

How to really understand the users' experience

Some design teams go to great lengths to ensure that they come to empathize with the users' experience. This box introduces two examples of this approach.

Buchenau and Suri (2000) describe experience prototyping, which is intended to give designers some insight into a user's experience that can only come from first-hand knowledge. They describe a team designing a chest-implanted automatic defibrillator. A defibrillator is used with victims of cardiac arrest when their heart muscle goes into a chaotic arrhythmia and fails to pump blood, a state called fibrillation. A defibrillator delivers an electric shock to the heart, often through paddle electrodes applied externally through the chest wall; an implanted defibrillator does this through leads that connect directly to the heart muscle. In either case, it's a big electric shock intended to restore the heart muscle to its regular rhythm that can be powerful enough to knock people off their feet.

This kind of event is completely outside most people's experience, and so it is difficult for designers to gain the insight they need to understand the user's experience. You can't fit a prototype pacemaker to each member of the design team and simulate fibrillation in them! However, you can simulate some critical aspects of the experience, one of which is the random occurrence of a defibrillating shock. To achieve this, each team member was given a pager to take home over the weekend (see Figure 11.8). The pager message simulated the occurrence of a defibrillating shock. Messages were sent at random, and team members were asked to record where they were, who they were with, what they were doing, and what they thought and felt knowing that this represented a shock. Experiences were shared the following week, and example insights ranged from anxiety around everyday happenings such as holding a child and operating power tools, to being in social situations and at a loss how to communicate to onlookers what was happening. This first-hand experience brought new insights to the design effort.

Figure 11.8 The patient kit for experience prototyping

Source: Buchenau, M. and Suri, J. F. (2000) Experience prototyping. In Proceedings of DIS 2000, Design Interactive Systems: Processes, Practices, Methods, Techniques, pp. 17–19.

Another instance is the Third Age suit, an empathy suit designed so that car designers can experience what it is like for people with some loss of mobility or declining sensory perception to drive their cars. The suit restricts movement in the neck, arms, legs, and ankles. Originally developed by Ford Motor Company and Loughborough University (see Figure 11.9) it has been used to raise awareness within groups of car designers, architects, and other product designers. □

Figure 11.9 **The Third Age empathy suit helps designers experience the loss of mobility and sensory perception**

Source: Ford Motor Co.

It is not the case that one way of approaching a conceptual design is right for one situation and wrong for another; all of these approaches provide different ways of thinking about the product and help in generating potential conceptual models.

Interface metaphors.

Interface metaphors combine familiar knowledge with new knowledge in a way that will help the user understand the product. Choosing suitable metaphors and combining new and familiar concepts requires a careful balance between utility and fun, and is based on a sound understanding of the users and their context. For example, consider an educational system to teach 6-year-olds mathematics. One possible metaphor is a classroom with a teacher standing at the blackboard. But if you consider the users of the system and what is likely to engage them, a metaphor that reminds the children of something they enjoy would be more suitable, such as a ball game, the circus, a playroom, and so on.

Erickson (1990) suggests a three-step process for choosing a good interface metaphor. The first step is to understand what the system will do, i.e. identifying the functional requirements. Developing partial conceptual models and trying them out may be part of the process. The second step is to understand which bits of the product are likely to cause users problems, i.e. which tasks or subtasks cause problems, are complicated, or are critical. A metaphor is only a partial mapping between the software and the real thing upon which the metaphor is based. Understanding areas in which users are likely to have difficulties means that the metaphor can be chosen to support those aspects. The third step is to generate metaphors. Looking for metaphors in the users' description of the tasks is a good starting point. Also, any metaphors used in the application domain with which the users may be familiar may be suitable.

When suitable metaphors have been generated, they need to be evaluated. Erickson (1990) suggests five questions to ask.

1. How much structure does the metaphor provide? A good metaphor will provide structure, and preferably familiar structure.
2. How much of the metaphor is relevant to the problem? One of the difficulties of using metaphors is that users may think they understand more than they do and start applying inappropriate elements of the metaphor to the product, leading to confusion or false expectations.
3. Is the interface metaphor easy to represent? A good metaphor will be associated with particular visual and audio elements, as well as words.
4. Will your audience understand the metaphor?
5. How extensible is the metaphor? Does it have extra aspects that may be useful later on?

For the shared travel organizer introduced in Chapter 10, one metaphor we could use is a printed travel brochure, which is commonly found in travel agents. This familiarity could be combined with facilities suitable for an electronic brochure such as videos of locations, and searching. To evaluate this metaphor, apply the five questions listed above.

1. Does it supply structure? Yes, it supplies structure based on the familiar paper-based brochure. This is a book and therefore has pages, a cover, some kind of binding to hold the pages together, an index, and table of contents. Travel brochures are often structured around destinations but are also sometimes structured around activities, particularly when the company specializes in adventure trips. However, a traditional brochure focuses on the details of the vacation and accommodation and has little structure to support visa or vaccination information (both of which change regularly and are therefore not suitable to include in a printed document).
2. How much of the metaphor is relevant? Having details of the accommodation, facilities available, map of the area, and supporting

illustrations is relevant for the travel organizer, so the content of the brochure is relevant. Also, structuring that information around types of vacation and destinations is relevant, and preferably both sets of grouping could be offered. But the physical nature of the brochure, such as page turning, is less relevant. The travel organizer can be more flexible than the brochure and should not try to emulate its book nature. Finally, the brochure is printed maybe once a year and cannot be kept up-to-date with the latest changes whereas the travel organizer should be capable of offering the most recent information.

3. Is the metaphor easy to represent? Yes. The vacation information could be a set of brochure-like pages. Note that this is not the same as saying that the navigation through the pages will be limited to page-turning.
4. Will your audience understand the metaphor? Yes.
5. How extensible is the metaphor? The functionality of a paper-based brochure is fairly limited. However, it is also a book, and we could borrow facilities from ebooks (which are also familiar objects to most of our audience), so yes, it can be extended.

Activity 11.2

Another possible interface metaphor for the travel organizer is the travel consultant. A travel consultant takes a set of requirements and tailors the vacation accordingly, offering maybe two or three alternatives, but making most of the decisions on the travelers' behalf. Ask the five questions above of this metaphor.

Comment

Show/Hide

Interaction types.

Chapter 2 introduced four different types of interaction: instructing, conversing, manipulating, and exploring. Which is best suited to the current design depends on the application domain and the kind of product being developed. For example, a computer game is most likely to suit a manipulating style, while a drawing package has aspects of instructing and conversing.

Most conceptual models will include a combination of interaction types, and it is necessary to associate different parts of the interaction with different types. For example, in the travel organizer, one of the user tasks is to find out the visa regulations for a particular destination. This will require an instructing approach to interaction as no dialog is necessary for the system to show the regulations. The user simply has to enter a predefined set of information, e.g. country issuing the passport and destination. On the other hand, trying to identify a vacation for a group of people may be conducted more like a conversation. For example, the user may begin by selecting some characteristics of the destination and some time constraints and preferences, then the organizer will respond with several options, and the user will provide more information or preferences and so on. (You may like to refer back to the scenario of this task in Chapter 10 and consider how well it matches this type of interaction.) Alternatively, for users who don't have any clear requirements yet, they might prefer to explore availability before asking for specific options.

Interface types.

Considering different interfaces at this stage may seem premature, but it has both a design and a practical purpose. When thinking about the conceptual model for a product, it is important not to be unduly influenced by a predetermined interface type. Different interface types prompt and support different perspectives on the product under development and suggest different possible behaviors. Therefore considering the effect of different interfaces on the product at this stage is one way to prompt alternatives.

Before the product can be prototyped, some candidate alternative interfaces will need to have been chosen. These decisions will depend on the product constraints, arising from the requirements you have established. For example, input and output devices will be influenced particularly by user and environmental requirements. Therefore, considering interfaces here also takes one step towards producing practical prototypes.

To illustrate this, we consider a subset of the interfaces introduced in Chapter 6, and the different perspectives they bring to the travel organizer:

- Shareable interface. The travel organizer has to be shareable as it is intended to be used by a group of people, and it should be exciting and fun. The design issues for shareable interfaces which were introduced in Chapter 6 will need to be considered for this system. For example how best (whether) to use the individuals’ own devices such as smartphones in conjunction with a shared interface.
- Tangible interface. Tangible interfaces are a form of sensor-based

interaction, where blocks or other physical objects are moved around. Thinking about a travel organizer in this way conjures up an interesting image of people collaborating, maybe with the physical objects representing themselves traveling, but there are practical problems of having this kind of interface, as the objects may be lost or damaged.

- Augmented and mixed reality. The travel organizer is not the kind of product that is usually designed for an augmented or mixed reality interface. The question is what would the physical object be in this case, that the virtual element could enhance? One possibility might be to enhance the physical brochure to provide more dynamic and easily changed information.

Activity 11.3

Consider the movie rental subscription service introduced in Chapter 10.

1. Identify tasks associated with this product that would best be supported by each of the interaction types instructing, conversing, manipulating, and exploring.
2. Pick out two interface types from Chapter 6 that might provide a different perspective on the design.

Comment

Show/Hide

11.3.2 Expanding the Initial Conceptual Model

Considering the issues in the previous section helps the designer to produce a set of initial conceptual model ideas. These ideas must be thought through in more detail and expanded before being prototyped or tested with users. For example, concrete suggestions of the concepts to be communicated between the user and the product and how they are to be structured, related, and presented are needed. This means deciding which functions the product will perform (and which the user will perform), how those functions are related, and what information is required to support them. These decisions will be made initially only tentatively and may change after prototyping and evaluation.

What functions will the product perform?

Understanding the tasks the product will support is a fundamental aspect of developing the conceptual model, but it is also important to consider which elements of the task will be the responsibility of the user and which will be carried out by the product. For example, the travel organizer may suggest specific vacation options for a given set of people, but is that as much as it should do? Should it automatically reserve the booking, or wait until it is told that this travel arrangement is suitable? Developing scenarios, essential use cases, and use cases will help clarify the answers to these questions. Deciding what the system will do and the user will do is sometimes called task allocation. This trade-off has cognitive implications (see Chapter 3), and is linked to social aspects of collaboration (see Chapter 4). If the cognitive load is too high for the user, then the device may be too stressful to use. On the other hand, if the product has too much control and is too inflexible, then it may not be used at all.

Another decision is which functions to hard-wire into the product and which to leave under software control, and thereby indirectly in the control of the human user.

How are the functions related to each other?

Functions may be related temporally, e.g. one must be performed before another, or two can be performed in parallel. They may also be related through any number of possible categorizations, e.g. all functions relating to privacy on a smartphone, or all options for viewing photographs in a social networking site. The relationships between tasks may constrain use or may indicate suitable task structures within the product. For example, if one task depends on another, the order in which tasks can be completed may need to be restricted.

If task analysis has been performed, the breakdown will support these kinds of decision. For example, in the travel organizer example, the task analysis performed in Section 10.7 shows the subtasks involved and the order in which the subtasks can be performed. Thus, the system could allow potential travel companies to be found before or after investigating the destination's facilities. It is, however, important to identify the potential travel companies before looking for availability.

What information is needed?

What data is required to perform a task? How is this data to be transformed by the system? Data is one of the categories of requirements identified and captured through the requirements activity. During conceptual design, these requirements are considered to ensure that the model provides the information necessary to perform the task. Detailed issues of structure and display, such as whether to use an analog display or a digital display, will more likely be dealt with during the concrete design activity, but implications arising from the type of data to be displayed may impact conceptual design issues.

For example, identifying potential vacations for a set of people using the travel organizer requires the following information: what kind of vacation is required; available budget; preferred destinations (if any); preferred dates and duration (if any); how many people it is for; and any special requirements (such as disability) that this group has. In order to perform the function, the system needs this information and must have access to detailed vacation and destination descriptions, booking availability, facilities, restrictions, and so on.

Initial conceptual models may be captured in wireframes – a set of documents that show structure, content, and controls. Wireframes may be constructed at varying levels of abstraction, and may show a part of the product or a complete overview. Case Study 11.2 and Chapter 12 include more information and some examples.

11.4 Concrete Design

Conceptual design and concrete design are closely related. The difference between them is rather a matter of changing emphasis: during design, conceptual issues will sometimes be highlighted and at other times, concrete detail will be stressed. Producing a prototype inevitably means making some concrete decisions, albeit tentatively, and since interaction design is iterative, some detailed issues will come up during conceptual design, and vice versa.

Designers need to balance the range of environmental, user, data, usability, and user experience requirements with functional requirements. These are sometimes in conflict. For example, the functionality of a wearable interactive product will be restricted by the activities the user wishes to perform while wearing it; a computer game may need to be learnable but also challenging.

There are many aspects to the concrete design of interactive products: visual appearance such as color and graphics, icon design, button design, interface layout, choice of interaction devices, and so on. Chapter 6 introduces several interface types and their associated design issues; these issues represent the kinds of decision that need to be made during concrete design. Case study 11.2 illustrates the impact that different-sized devices may have on the same application, and the need to explicitly design for different form factors. Chapter 6 also introduces some guidelines, principles, and rules for different interface types to help designers ensure that their products meet usability and user experience goals.

Case Study 11.2

Designing mobile applications for multiple form factors

Trutap is a social networking service designed for more than 350 different models of mobile device, which was built for a UK startup between 2007 and 2009. It aggregates online blogging, instant messaging, and social services like Facebook, allowing its users to interact with them (see Figures 11.10 and 11.11). The design of the Trutap application, which took place over two major releases, posed significant challenges in terms of how to integrate disparate sources of data onto small-screen devices, and produce a design which would scale between form factors, i.e. different physical handset designs.

Figure 11.10 Trutap: version 2.0 design concepts

Source: © Trutap, Reproduced with permission.

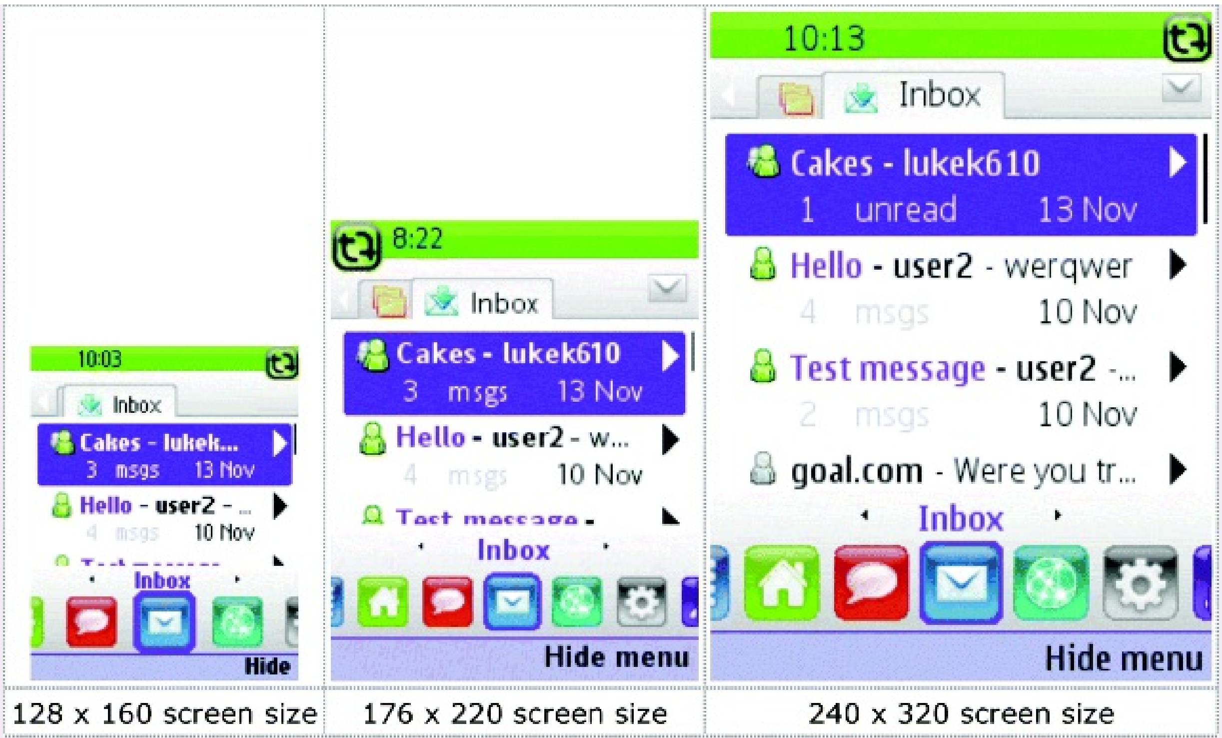


Figure 11.11 Trutap: version 2.0 screenshots, inbox

Source: © Trutap, Reproduced with permission.

The product was designed with a clear goal: teenagers and young adults were spending much of their social lives online. Trutap would help them keep connected, wherever they were.

Two versions of the product were launched: Trutap 1.0 offered its own mechanisms for managing people's contacts and communicating with them, and tied into a range of existing instant messaging networks (Yahoo!, MSN, AOL, and the like). Launched in 2008, this version saw far greater take-up in India and Indonesia than with its original target audience of UK students.

This take-up, combined with the successful launch of the iPhone in July 2008 and the increasing prominence of Facebook as the dominant site for personal social networking, led to a change in emphasis for the 2.0 release of Trutap. Launched a year after 1.0, and technically an evolution rather than a reworking, release 2.0 emphasized the aggregation of existing online services, tying into Facebook, weblogging software, and photo management, and extending the number of instant messaging services covered. Publicly, the product was presented as a means for aspirational middle classes in the developing world to experience many of the same benefits that the iPhone promised, but on their conventional mobile devices.

This case study, by Tom Hume, Johanna Hunt, Bryan Rieger, and Devi

Lozdan from Future Platforms Ltd, explores the impact that different form factors had on the design of Trutap. ■

Concrete design also deals with issues related to user characteristics and context, and two aspects that have drawn particular attention for concrete design are accessibility and national culture. Accessibility was discussed in Box 1.2. Researchers, designers, and evaluators have investigated a range of techniques, toolkits, and interaction devices to support individuals with different accessibility needs. More recently, there has been a reaction to this approach that challenges the ‘rhetoric of compassion’ in favor of a ‘rhetoric of engagement', and suggests that users be empowered rather than designed for (Rogers and Marsden, 2013).

Aspects of cross-cultural design include use of appropriate language(s), colors, icons and images, navigation, and information architecture (Rau et al, 2013). Example design guidelines include ensuring that the product supports different formats for dates, times, numbers, measurements, and currencies, and that generic icons are designed where possible (Esselink, 2000). However, Marsden et al (2008) warn of the problems in seeing a user's need and attempting to meet that need without first asking the community if they, too, recognize that need (see also Case Study 11.3 and Gary Marsden's interview at the end of this chapter).

Guidelines, although seemingly attractive, can be misguided. One of the most well-known sets of guidelines for cultural web design was proposed by Marcus and Gould (2000), building on the cultural dimensions proposed by Hofstede (1994). However, Hofstede's work, and its application in interaction design, has been challenged (see Box 11.4), and designing for a cross-cultural audience is now recognized as more than a translation exercise. As Carlson (1992, p. 175) has put it, successful products “are not just bundles of technical solutions; they are also bundles of social solutions. Inventors succeed in a particular culture because they understand the values, institutional arrangements, and economic notions of that culture.”

BOX 11.4

Using Hofstede's dimensions in interaction design

One of the most influential pieces of work on characterizing national culture differences was carried out by a management theorist called Geert Hofstede around 1970. He was given access to responses from a survey of IBM employees in over 50 countries worldwide and from this he identified four dimensions of national culture: power distance (PD), individualism (IND), masculinity–femininity (MAS), and uncertainty avoidance (UA). As a result of work done in Hong Kong at a later date by a Canadian, Michael Bond, a fifth dimension was added that deals with time orientation.

Although influential, Hofstede's work does have limitations. For example, he admits that the people involved in designing the original questionnaire were all from Western cultures. In addition, his studies have been discussed and challenged over the intervening years: e.g. Oyserman et al (2002) challenged his finding that European Americans are more individualistic than people from other ethnic groups. The application of his ideas in interaction design has also been challenged – e.g. work by Oshlyansky (2007) found that Hofstede's model does not help explain cultural differences in affordance; nor does it seem to apply to technology acceptance. So, although popular, Hofstede's dimensions may not be the best approach to accommodating national culture differences in interaction design. □

11.5 Using Scenarios

Scenarios are informal stories about user tasks and activities. Scenarios can be used to model existing work situations, but they are more commonly used for expressing proposed or imagined situations to help in conceptual design. Often, stakeholders are actively involved in producing and checking through scenarios for a product. Bødker suggests four roles (Bødker, 2000, p. 63):

1. As a basis for the overall design.
2. For technical implementation.
3. As a means of cooperation within design teams.
4. As a means of cooperation across professional boundaries, i.e. as a basis of communication in a multidisciplinary team.

In any one project, scenarios may be used for any or all of these. More specifically, scenarios have been used as scripts for user evaluation of prototypes, as the basis of storyboard creation (see Section 11.6.1), and to build a shared understanding among team members. Scenarios are good at selling ideas to users, managers, and potential customers.

Bødker proposes the notion of plus and minus scenarios. These attempt to capture the most positive and the most negative consequences of a particular proposed design solution (see Figure 11.12), thereby helping designers to gain a more comprehensive view of the proposal. This idea has been extended by Mancini et al (2010) who use positive and negative video scenarios to explore futuristic technology.

Scenario 3: Hyper-wonderland

This scenario addresses the positive aspects of how a hypermedia solution will work.

The setting is the Lindholm construction site sometime in the future.

Kurt has access to a portable PC. The portables are hooked up to the computer at the site office via a wireless modem connection, through which the supervisors run the hypermedia application.

Action: During inspection of one of the caissons[1] Kurt takes his portable PC, switches it on and places the cursor on the required information. He clicks the mouse button and gets the master file index together with an overview of links. He chooses the links of relevance for the caisson he is inspecting.

Kurt is pleased that he no longer needs to plan his inspections in advance. This is a great help because due to the 'event-driven' nature of inspection, constructors never know where and when an inspection is taking place. Moreover, it has become much easier to keep track of personal notes, reports etc. because they can be entered directly on the spot.

The access via the construction site interface does not force him to deal with complicated keywords either. Instead, he can access the relevant information right away, literally from where he is standing.

A positive side-effect concerns his reachability. As long as he has logged in on the computer, he is within reach of the secretaries and can be contacted when guests arrive or when he is needed somewhere else on the site. Moreover, he can see at a glance where his colleagues are working and get in touch with them when he needs their help or advice.

All in all, Kurt feels that the new computer application has put him more in control of things.

[1] Used in building to hold water back during construction.

Scenario 4: Panopticon

This scenario addresses the negative aspects of how a hypermedia solution will work.

The setting is the Lindholm construction site sometime in the future.

Kurt has access to a portable PC. The portables are hooked up to the computer at the site office via a wireless modem connection, through which the supervisors run the hypermedia application.

Action: During inspecting one of the caissons Kurt starts talking to one of the builders about some reinforcement problem. They argue about the recent lab tests, and he takes out his portable PC in order to provide some data which justify his arguments. It takes quite a while before he finds a spot where he can place the PC: either there is too much light, or there is no level surface at a suitable height. Finally, he puts the laptop on a big box and switches it on. He positions the cursor on the caisson he is currently inspecting and clicks the mouse to get into the master file. The table of contents pops up and from the overview of links he chooses those of relevance – but no lab test appears on the screen. Obviously, the file has not been updated as planned.

Kurt is rather upset. This loss of prestige in front of a contractor engineer would not have happened if he had planned his inspection as he had in the old days.

Sometimes. he feels like a hunted fox especially in situations where he is drifting around thinking about what kind of action to take in a particular case. If he has forgotten to log out, he suddenly has a secretary on the phone: "I see you are right at caisson 39, so could you not just drop by and take a message?"

All in all Kurt feels that the new computer application has put him under control.

Figure 11.12 Example plus and minus scenarios

Source: S. Bødker (2000) Scenarios in user-centered design – setting the stage for reflection and action. Interacting with Computers, 13(1), Fig. 2, p. 70.

11.6 Generating Prototypes

In this section we illustrate how prototypes may be used in design, and demonstrate one way in which prototypes may be generated from the output of the requirements activity: producing a storyboard from a scenario and a card-based prototype from a use case. Both of these are low-fidelity prototypes and they may be used as the basis to develop more detailed interface designs and higher-fidelity prototypes as development progresses.

11.6.1 Generating Storyboards from Scenarios

A storyboard represents a sequence of actions or events that the user and the product go through to achieve a task. A scenario is one story about how a product may be used to achieve the task. It is therefore possible to generate a storyboard from a scenario by breaking the scenario into a series of steps which focus on interaction, and creating one scene in the storyboard for each step. The purpose for doing this is two-fold: first, to produce a storyboard that can be used to get feedback from users and colleagues; second, to prompt the design team to consider the scenario and the product's use in more detail. For example, consider the scenario for the travel organizer developed in Chapter 10. This can be broken down into five main steps:

1. The Thomson family gather around the organizer and enter a set of initial requirements.
2. The system's initial suggestion is that they consider a flotilla trip but Sky and Eamonn aren't happy.
3. The travel organizer shows them some descriptions of the flotillas written by young people.
4. Will confirms this recommendation and asks for details.
5. The travel organizer emails the details.

Activity 11.4

Consider an augmented reality in-car navigation system that takes information from a GPS and displays routes and traffic information directly onto the car windscreen. Suggest one plus and one minus scenario. For the plus scenario, think of the possible benefits of the system. For the minus scenario, imagine what could go wrong.

Figure 11.13 An example car-navigation system based on augmented reality

Source: The Aeon Project, courtesy of Michaël Harboun

Comment

Show/Hide

The first thing to notice about this set of steps is that it does not have the detail of a use case but identifies the key events or activities associated with the scenario. The second thing to notice is that some of these events are focused solely on the travel organizer's screen and some are concerned with the environment. For example, the first one talks about the family gathering around the organizer, while the third and fifth are focused on the travel organizer. We therefore could produce a storyboard that focuses on the screens or one that is focused on the environment. Either way, sketching out the storyboard will prompt us to think about design issues.

For example, the scenario says nothing about the kind of input and output devices that the system might use, but drawing the organizer forces you to think about these things. There is some information about the environment within which the system will operate, but again drawing the scene makes you stop and think about where the organizer will be. You don't have to make any decisions about, e.g. using a trackball, or a touch screen, or whatever, but you are forced to think about it. When focusing on the screens, the designer is prompted to consider issues including what information needs to be available and what information needs to be output. This all helps to explore design decisions and alternatives, but is also made more explicit because of the drawing act.

We chose to draw a storyboard that focuses on the environment of the travel organizer, and it is shown in Figure 11.14. While drawing this, various questions relating to the environment came to mind such as how can the interaction be designed for all the family? Will they sit or stand? How confidential should the interaction be? What kind of documentation or help needs to be available? What physical components does the travel organizer need? And so on. In this exercise, the questions it prompts are just as important as the end product.

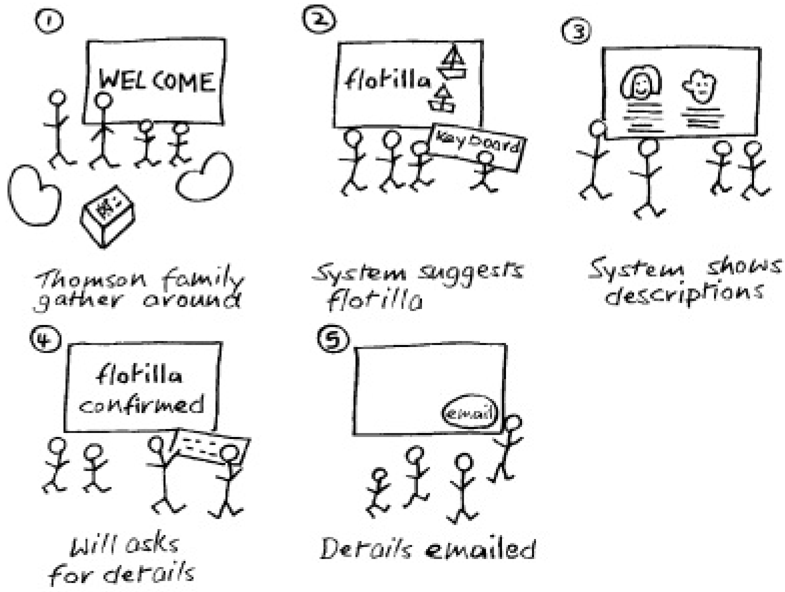


Figure 11.14 The storyboard for the travel organizer focusing on environmental issues

Note that although we have used the scenario as the main driver for producing the storyboard, there is other information from the requirements activity that also informs development.

Activity 11.5

Activity 10.3 asked you to develop a futuristic scenario for the one-stop car shop. Using this scenario, develop a storyboard that focuses on the environment of the user. As you are drawing this storyboard, write down the design issues you are prompted to consider.

Comment

Show/Hide

11.6.2 Generating Card-Based Prototypes from Use Cases

The value of a card-based prototype lies in the fact that the screens or interaction elements can be manipulated and moved around in order to simulate interaction with a user or to explore the user's end-to-end experience. Where a storyboard focusing on the screens has been developed, this can be translated into a card-based prototype and used in this way. Another way to produce a card-based prototype is to generate one from a use case output from the requirements activity.

For example, consider the use case generated for the travel organizer in Section 10.6.2. This focused on the visa requirements part of the system. For each step in the use case, the travel organizer will need to have an interaction component to deal with it, e.g. a button or menu option, or a display. By stepping through the use case, it is possible to build up a card-based prototype to cover the required behavior. For example, the cards in Figure 11.16 were developed by considering each of the steps in the use case. Card one covers step 1; card two covers steps 2, 3, 4, 5, 6, and 7; and card three covers steps 8, 9, 10, and 11 (notice the print button that is drawn into card three to allow for steps 10 and 11). As with the storyboards, drawing concrete elements of the interface like this forces the designer to think about detailed issues so that the user can interact with the prototype. In card two you will see that I chose to use a drop-down menu for the country and nationality. This is to avoid mistakes. However, the flaw in this is that I may not catch all of the countries in my list, and so an alternative design could also be incorporated where the user can choose an 'enter below' option and then type in the country or nationality (see Figure 11.17).

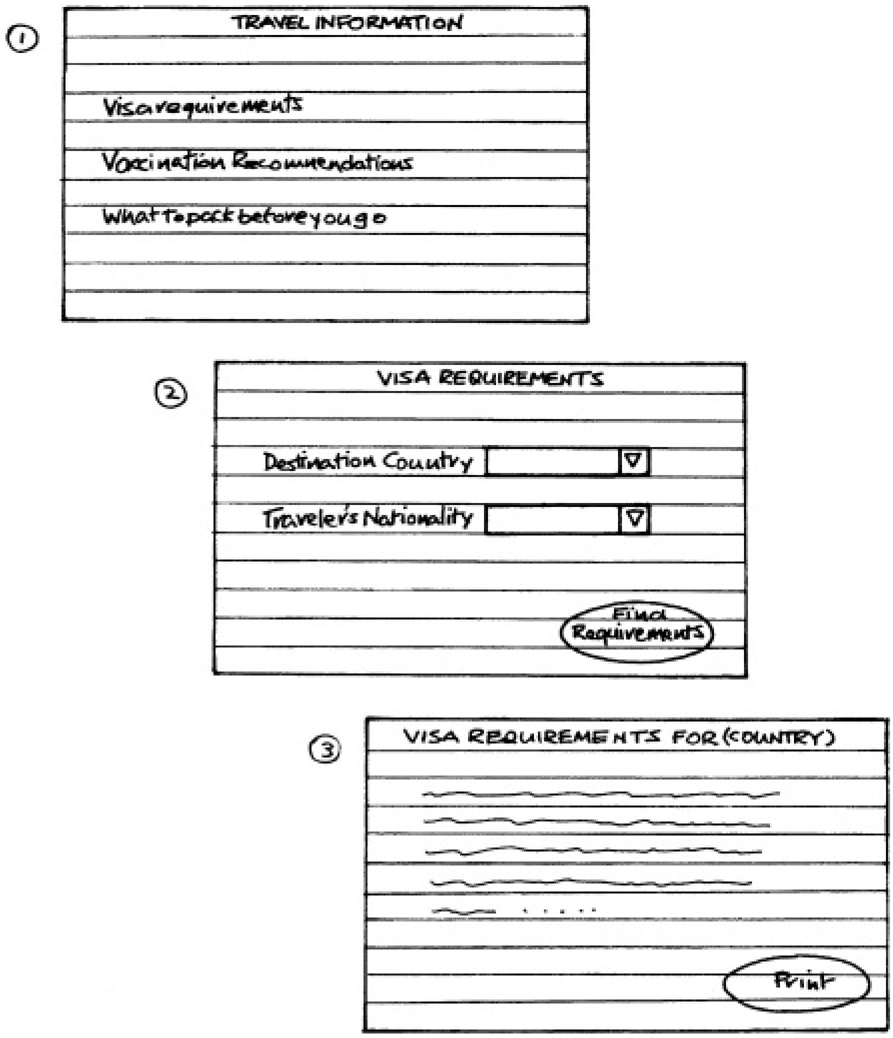


Figure 11.16 Cards one to three of a card-based prototype for the travel organizer

Figure 11.17 Card four of a card-based prototype for the travel organizer

These cards can then be shown to potential users of the system or fellow designers to get informal feedback. In this case, I showed these cards to a colleague, and through discussion of the application and the cards, concluded that although the cards represent one interpretation of the use case, they focus too much on an interaction model that assumes a WIMP/GUI interface. Our discussion was informed by several things including the storyboard and the scenario. One alternative would be to have a map of the world, and users can indicate their destination and nationality by choosing one of the countries on the map; another might be based around national flags. These alternatives could be prototyped using cards and further feedback obtained.

Activity 11.6

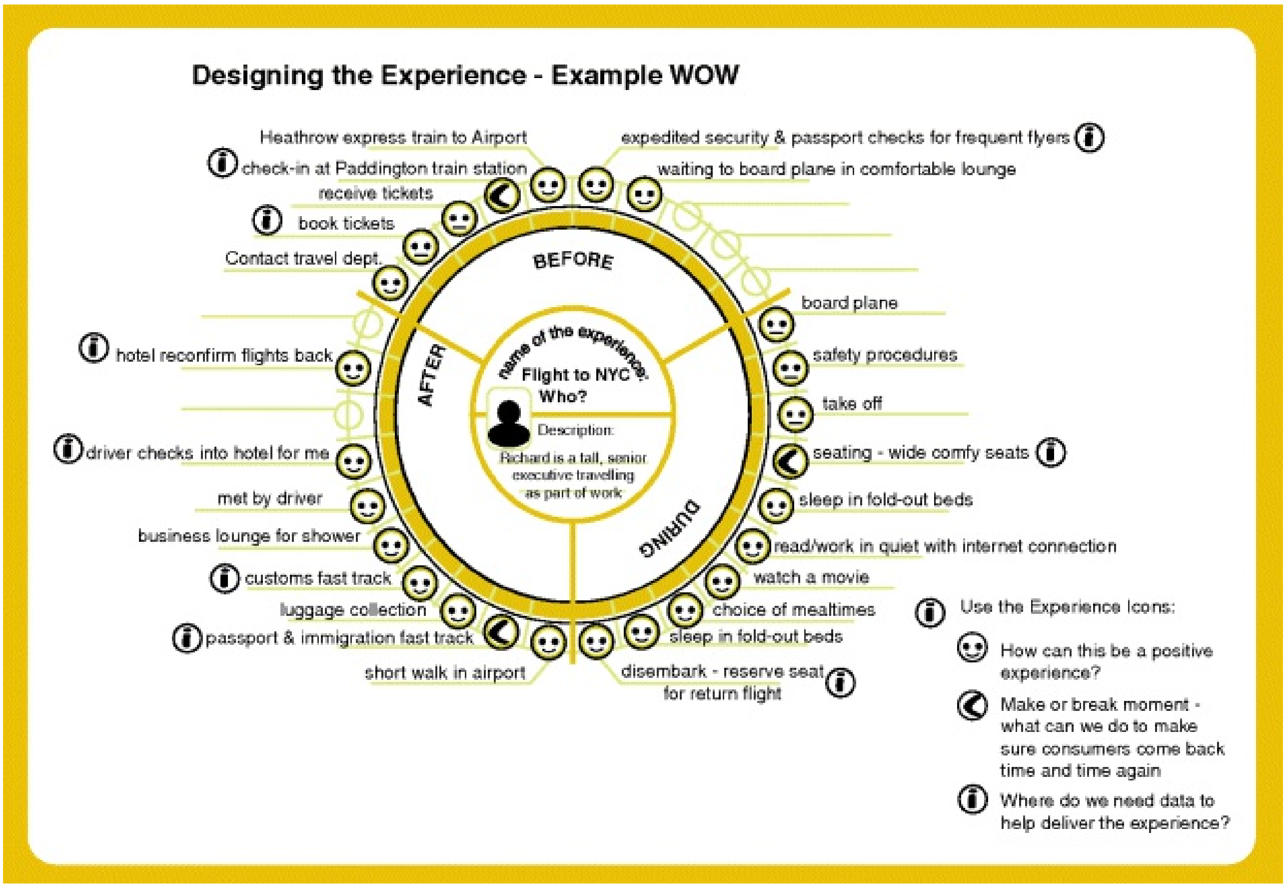
Produce a card-based prototype for the movie rental subscription service and the task of renting a movie as described by the use case in Activity 10.4. You may also like to ask one of your peers to act as a user and step through the task using the prototype.

Comment

Show/Hide

A set of card-based prototypes that cover a scenario from beginning to end may be the basis of a more detailed prototype, such as an interface or screen sketch, or it may be used in conjunction with personas to explore the user's end-to-end experience. This latter purpose is achieved by creating a visual representation of the user's experience. These representations are variably called a design map (Adlin and Pruitt, 2010) or a customer journey map (Ratcliffe and McNeill, 2012), or an experience map. They illustrate a user's path or journey through the product or service, and are usually created for a particular persona, hence giving the journey sufficient context and detail to bring the discussions to life. They support designers in considering the user's overall experience when achieving a particular goal and are used to explore and question the designed experience and to identify issues that have not been considered so far. They may be used to analyze existing products and to collate design issues, or as part of the design process.

There are many different types of representation, of varying complexities. Two main ones are: the wheel and the timeline. The wheel representation is used when an interaction phase is more important than an interaction point, e.g. for a flight (see Figure 11.19(a) for an example). The timeline is used where a service is being provided that has a recognizable beginning and end point, such as purchasing an item through a website (an example of a timeline representation is in Figure 10.4(b)). Figure 11.19(b) illustrates the structure of a timeline and how different kinds of issues may be captured, e.g. questions, comments, and ideas.



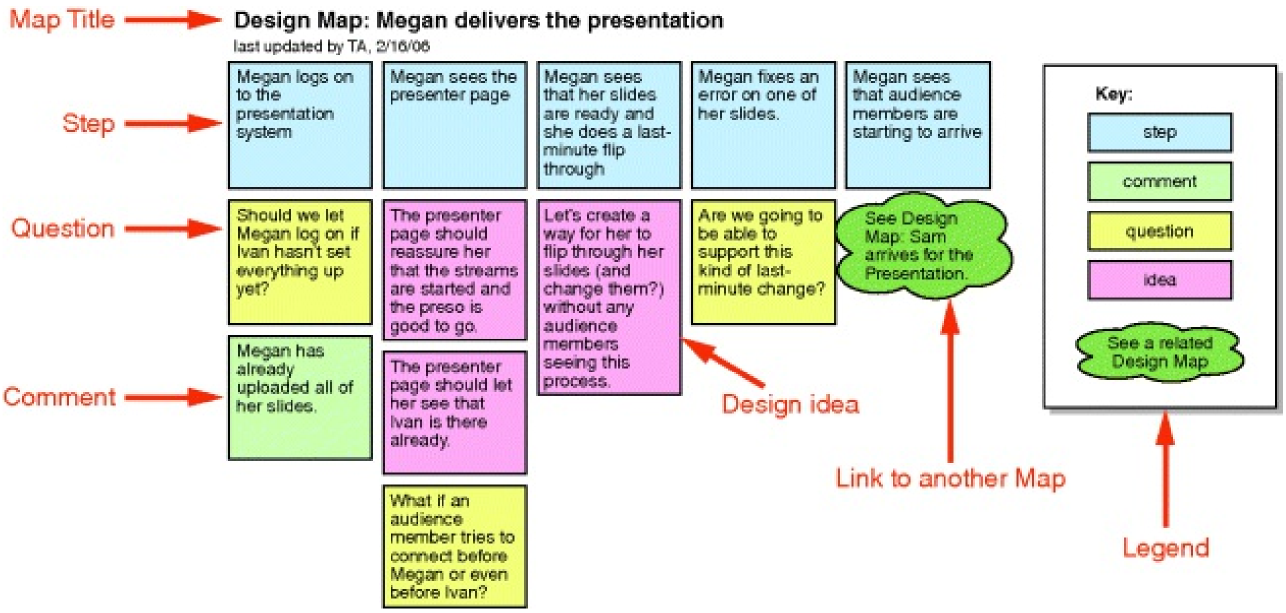


Figure 11.19 (a) An experience map using a wheel representation. (b) An example timeline design map illustrating how to capture different issues.

Source: (a) http://www.ux-lady.com/experience-maps-user-journey-and-more-exp-map-layout/ (b) Adlin, T. and Pruitt, J. (2010) The Essential Persona Lifecycle: Your guide to building and using personas. Morgan Kaufmann p. 134.

To generate one of these representations, take one persona and two or three scenarios. Draw a timeline for the scenario and identify the interaction points for the user. Then use this as a discussion tool with colleagues to identify any issues or questions that may arise. Some people consider the user's mood and identify pain points, sometimes the focus will be on technical issues, and sometimes this can be used to identify missing functionality or areas of under-designed interaction.

Video illustrating the benefits of experience mapping using a timeline at http://youtu.be/eLT_Q8sRpyI

BOX 11.5

Involving users in design: participatory design

The idea of participatory design emerged in Scandinavia in the late 1960s and early 1970s. There were two influences on this early work: the desire to be able to communicate information about complex systems, and the labor union movement pushing for workers to have democratic control over changes in their work. In the 1970s, new laws gave workers the right to have a say in how their working environment was changed, and such laws are still in force today. A fuller history of the movement is given in Ehn (1989) and Nygaard (1990).

Several projects at this time attempted to involve users in design and focus on work rather than on simply producing a product. One of the most discussed is the UTOPIA project, a cooperative effort between the Nordic Graphics Workers Union and research institutions in Denmark and Sweden to design computer-based tools for text and image processing.

Involving users in design decisions is not simple, however. Cultural differences can become acute when users and designers are asked to work together to produce a specification for a system. Bødker et al (1991) recount the following scene from the UTOPIA project: “Late one afternoon, when the designers were almost through with a long presentation of a proposal for the user interface of an integrated text and image processing system, one of the typographers commented on the lack of information about typographical code-structure. He didn't think that it was a big error (he was a polite person), but he just wanted to point out that the computer scientists who had prepared the proposal had forgotten to specify how the codes were to be presented on the screen. Would it read ‘<bf/’ or perhaps just ‘\b’ when the text that followed was to be printed in boldface?”

In fact, the system being described by the designers was a WYSIWYG (what you see is what you get) system, and so text that needed to be in bold typeface would appear as bold (although most typographic systems at that time did require such codes). The typographer was unable to link his knowledge and experience with what he was being told. In response to this kind of problem, the project started using mockups. Simulating the working situation helped workers to draw on their experience and tacit knowledge, and designers to get a better understanding of the actual work typographers needed to do.

Case Study 11.3 describes an extension to the participatory design idea, called community-based design. □

Case Study 11.3

Deaf telephony

This case study by Edwin Blake, William Tucker, Meryl Glaser, and Adinda Freudenthal discusses their experiences of community-based design in South Africa. The process of community-based co-design is one that explores various solution configurations in a multidimensional design space whose axes are the different dimensions of requirements and the various dimensions of designer skills and technological capabilities. The bits of this space that one can ‘see’ are determined by one's knowledge of the user needs and one's own skills. Co-design is a way of exploring that space in a way that alleviates the myopia of one's own viewpoint and bias. As this space is traversed, a trajectory is traced according to one's skills and learning, and according to the users’ expressed requirements and their learning.

The project team set out to assist South African deaf people to communicate with each other, with hearing people, and with public services. The team has been working for many years with a deaf community that has been disadvantaged due to both poverty and hearing impairment. The story of this wide-ranging design has been one of continual fertile (and on occasion frustrating) co-design with this community. The team's long-term involvement has meant they have transformed aspects of the community and that they have themselves been changed in what they view as important and in how they approach design.

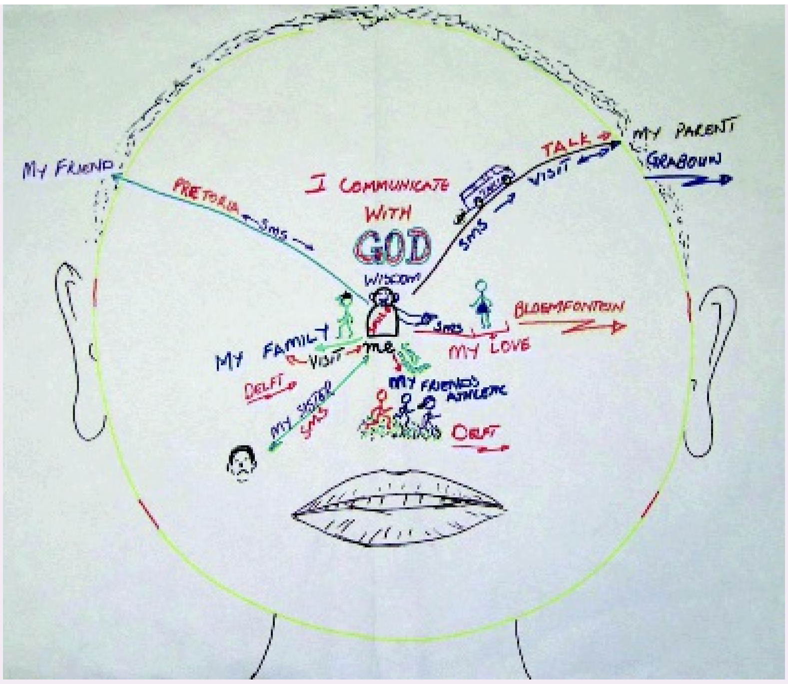


Figure 11.20 One participant's view of communication

Source: Copyright Edwin Blake et al.

Figure 11.21 Participants discussing design in sign language

Deaf users in this community started out knowing essentially nothing about computers. Their first language is South African Sign Language (SASL) and this use of SASL is a proud sign of their identity as a people. Many are also illiterate or semi-literate. There are a large number of deaf people using SASL; in fact there are more than some of the smaller official languages. Since the advent of democracy in 1994, there has been an increasing empowerment of deaf people and it is accepted as a distinct language in its own right.

In this case study, a brief historical overview of the project and the various prototypes that formed nodes in a design trajectory are presented. The methodology of Action Research and its cyclical approach to homing in on an effective implementation is reviewed. An important aspect of the method is how it facilitates learning by both the researchers and the user community so that together they can form an effective design team. Lastly, such a long-term intimate involvement with a community raises important ethical issues, which are fundamentally concerns of reciprocity. ■

11.7 Construction

As prototyping and building alternatives progresses, development will focus more on putting together components and developing the final product. This may take the form of a physical product, such as a set of alarms, sensors and lights, or a piece of software, or both. Whatever the final form, it is very unlikely that you will develop anything from scratch as there are many useful (in some cases essential) resources to support development. Here we introduce two kinds of resource: physical computing kits and software development kits (SDKs).

11.7.1 Physical Computing

Physical computing is concerned with how to build and code prototypes and devices using electronics. Specifically, it is the activity of “creating physical artifacts and giving them behaviors through a combination of building with physical materials, computer programming and circuit building”(Gubbels and Froehlich, 2014). Typically, it involves designing things, using a printed circuit board (PCB), sensors (e.g. accelerometers, infrared, temperature) to detect states, and actuators (e.g. motors, valves) that cause some effect. An example is a ‘friend or foe’ cat detector that senses, via an accelerometer, any cat (or anything else for that matter) that tries to push through a family's catflap. The movement triggers an actuator to take a photo of what came through the catflap using a webcam positioned on the back door. The photo is uploaded to a website that alerts the owner if there are cats or other objects that do not match their own cat's image.

A number of physical computing toolkits have been developed for educational and prototyping purposes. One of the earliest is Arduino (see Banzi, 2009). The aim was to enable artists and designers to learn how to make and code physical prototypes using electronics in a couple of days, having attended a workshop. The toolkit is composed of two parts: the Arduino board (see Figure 11.22), which is the piece of hardware that is used to build objects, and the Arduino IDE (integrated development environment), which is a piece of software that makes it easy to program and upload a sketch (Arduino's name for a unit of code) to the board. A sketch, for example, might turn on an LED when a sensor detects a change in the light level. The Arduino board is a small circuit that contains a tiny chip (the microcontroller). It has a number of protruding 'legs' that provide input and output pins – which the sensors and actuators are connected to. Sketches are written in the IDE using a simple processing language, then uploaded to the board and translated into the 'C' programming language.

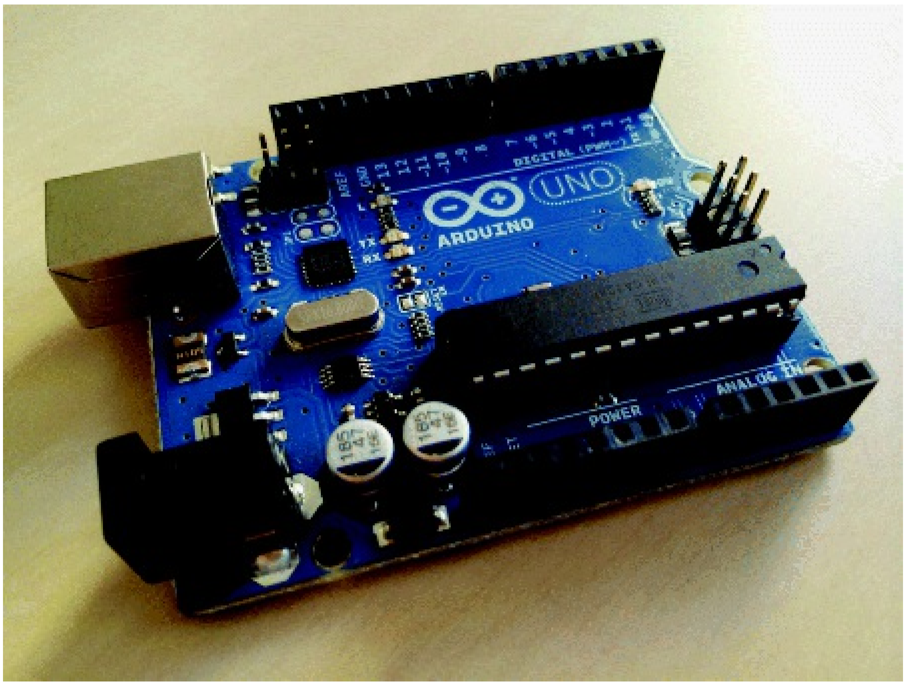


Figure 11.22 The Arduino board

Source: Courtesy of Nicolai Marquardt

There are other toolkits that have been developed, based on the basic Arduino kit. The most well known is the LilyPad, which was co-developed by Leah Beuchley (see Figure 11.23 and her interview at the end of Chapter 6). It is a set of sewable electronic components for building fashionable clothing and other textiles. The Engduino® is a teaching tool based on the Arduino LilyPad; it has 16 multicolour LEDs and a button, which can be used to provide visual feedback, and simple user input. It also has a thermistor (that senses temperature), a 3D accelerometer (that measures accelerations), and an infrared transmitter/receiver that can be used to transmit messages from one Engduino® to another.

Figure 11.23 The Lilypad Arduino kit

Source: Photo courtesy of Leah Buechley http://web.media.mit.edu/~leah/LilyPad/build/turn_signal_jacket.html.

Video introducing MakeMe (Marquardt et al, 2015), a novel toolkit that is assembled from a flat electronic sheet, where six sides are snapped out then slotted together to become an interactive cube that lights up in different colors, depending on how vigorously it is shaken. Intended to encourage children to learn, share, and fire their imagination to come up with new games and other uses, see it in action at **http://www.codeme.io/**

Other kinds of easy-to-use and quick-to-get-started physical toolkits – intended to provide new opportunities for people to be inventive and creative with – are Senseboard (Richards and Woodthorpe, 2009), LittleBits, and MaKey MaKey (Silver and Rosenbaum, 2012). The MaKey MaKey toolkit comprises a printed circuit board with an Arduino microcontroller, alligator clips, and a USB cable (see Figure 11.24). It communicates with a computer to send key presses, mouse clicks, and mouse movements. There are six inputs (the four arrow keys, the space bar, and a mouse click) positioned on the front of the board that alligator clips are clipped onto in order to connect with a computer via the USB cable. The other ends of the clips can be attached to any non-insulating object, such as a vegetable or piece of fruit. Thus, instead of using the computer keyboard buttons to interact with the computer, external objects such as bananas are used. The computer thinks MaKey MaKey is just like a keyboard or mouse. An example is to play a digital piano app using bananas as keys rather than keys on the computer keyboard. When they are touched, they make a connection to the board and MaKey MaKey sends the computer a keyboard message.

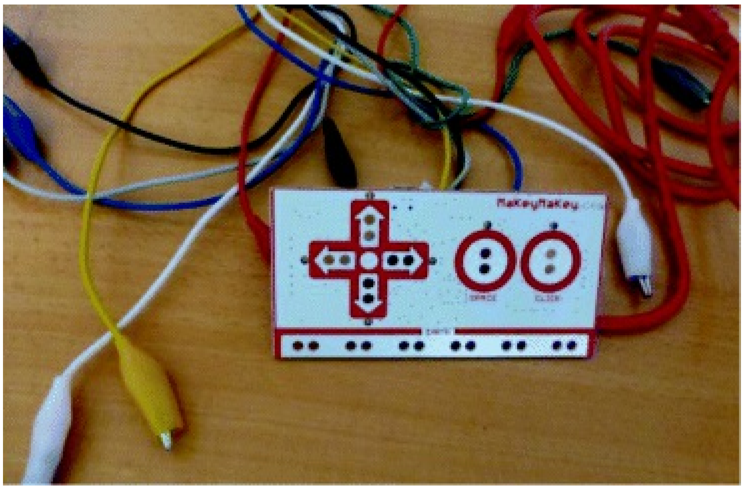


Figure 11.24 The MaKey MaKey toolkit

So far, physical toolkits have been aimed at children or designers to enable them to start programming through rapid creation of small electronic gadgets and digital tools (e.g. Hodges et al, 2013). However, Rogers et al (2014) demonstrated how retired people were equally able to be creative using the kit, turning "everyday objects into touchpads." They ran a series of workshops where small groups of retired friends, aged between early 60s and late 80s, assembled and played with the MaKey MaKey toolkit (see Figure 11.25). After playing music using fruit and vegetables as input, they saw many new possibilities for innovative design. Making and playing together, however childlike it might seem at first, can be a catalyst for imagining, free thinking, and exploring. People are sometimes cautious to volunteer their ideas, fearing that they are easily squashed, but in a positive environment they can flourish. The right kind of shared experience can create a positive and relaxed atmosphere, in which people from all walks of life can freely bounce ideas around.

Figure 11.25 A group of retired friends playing with a MaKey MaKey toolkit

BOX 11.6

The rise of the Maker Movement

The Maker Movement emerged in the mid 2000s. Following in the footsteps of the computer revolution and the Internet, it is viewed as the next big revolution that will transform manufacturing and production (Hatch, 2014). Whereas the explosion of the web was all about what it could do for us virtually, with a proliferation of apps, social media, and services, the Maker Movement is transforming how we make, buy, consume, and recycle things, from houses to clothes and food to bicycles. At its core is DIY – crafting physical things using a diversity of machines, tools, and methods collaboratively in workshops and makespaces. In a nutshell, it is about inventing the future through connecting technologies, the Internet, and physical things.

While there have always been hobbyists tinkering away making radios, clocks, and other devices, the world of DIY making has been opened up to many more people. Affordable, powerful, and easy-to-use tools, coupled with a renewed focus on locally sourced products and community-based activities, and a desire for sustainable, authentic, and ethically produced products, has led to a ground swell in interest in how to make. Fablabs (fabrication laboratories) first started appearing in cities throughout the world, offering a large physical space containing electronics and manufacturing equipment, including 3D printers, CNC milling machines, and laser cutters. Individuals bring their digital files to print and make things – which would have been impossible for them to do previously – such as large 3D models, furniture, and installations. Then smaller makerspaces started appearing in their thousands across the world, from Shanghai to rural India, again sharing production facilities for all to use and make. While some are small, for example sharing the use of a 3D printer, others are much larger and well resourced, offering an array of manufacturing machines, tools, and workspaces to make in.

Another development has been to build and program e-textiles using sewing machines and electronic thread. E-textiles comprise fabrics that are embedded with electronics, such as sensors, LEDs, and motors that are stitched together using conductive thread and conductive fabrics (Buechley and Qiu, 2014). An early example is the turn-signal biking jacket (developed by Leah Buechley and illustrated in Figure 1.3). Other e-textiles include interactive soft toys, wallpaper that sings when touched, and fashion clothing that reacts to the environment or events.

A central part of the Maker Movement involves tinkering (as discussed in Section 11.2.4) and the sharing of knowledge, skills, know-how, and what you have made. The Instructables.com website is for anyone to explore, document, and share their DIY creations. Go to the Instructables site and take a look at a few of the projects that have been uploaded by makers. How many of them are a combination of electronics, physical materials, and pure invention? Are they fun, useful, or gadgety? How are they presented? Do they inspire you to make? Another site, Etsy.com, is a popular online marketplace for people who make things to sell their crafts and other handmade items. It is designed to be easy for makers to use and set up their store to sell to family, friends, and strangers across the world. Unlike the corporate online sites, (e.g. Amazon, eBay), it is a place for craft makers to reach others and to show off their wares in ways they feel best fit what they have made.

In essence, the Maker Movement is about taking the DIY movement online to make it public and, in doing so, massively increase who can take part and how it is shared (Anderson, 2013). ◻

Activity 11.7

Watch the video of Lady Gaga in the Voltanis, the first flying dress, developed by the e-textile company XO. What do you think of this fusion of fashion and state-of-the-art electronics and technology?

Video of Lady Gaga in the Voltanis at http://vimeo.com/91916514

Comment

Show/Hide

11.7.2 SDKs: Software Development Kits

A software development kit (SDK) is a package of programming tools and components that supports the development of applications for a specific platform, e.g. for iOS on iPad, iPhone, and iPod touch, for the Kinect device, and for the Windows phone. Typically an SDK includes an IDE (integrated development environment), documentation, drivers, and sample programming code to illustrate how to use the SDK components. Some also include icons and buttons that can easily be incorporated into the design. While it is possible to develop applications without using an SDK, it is so much easier using such a powerful resource, and so much more can be achieved.

For example, the availability of Microsoft's Kinect SDK makes the device's powerful gesture recognition and body motion tracking capabilities accessible. This has led to the exploration of, for example, motion tracking in immersive games (Manuel et al, 2012), user identification using body lengths (Hayashi et al, 2014), and robot control (Wang et al, 2013).

An SDK will include a set of application programming interfaces (APIs) that allows control of the components without knowing the intricacies of how they work. In some cases, access to the API alone is sufficient to allow significant work to be udertaken, e.g. Hayashi et al (2014) only needed access to the APIs. The difference between APIs and SDKs is explored in Box 11.7.

BOX 11.7

APIs and SDKs

SDKs (software development kits) consist of a set of programming tools and components while an API (application programming interface) is the set of inputs and outputs, i.e. the technical interface to those components. To explain this further, an API allows different-shaped building blocks of a child's puzzle to be joined together, while an SDK provides a workshop where all of the development tools are available to create whatever size and shape blocks you fancy, rather than using pre-shaped building blocks. An API therefore allows the use of pre-existing building blocks, while an SDK removes this restriction and allows new blocks to be created, or even to build something without blocks at all. An SDK for any platform will include all the relevant APIs, but it adds programming tools, documentation, and other development support as well. □

Assignment

This assignment continues work on the online booking facility introduced at the end of Chapter 10. The work will be continued in the assignments for Chapters 12, 14, and 15.

a. Based on the information gleaned from the assignment in Chapter 10, suggest three different conceptual models for this system. You should consider each of the aspects of a conceptual model discussed in this chapter: interface metaphor, interaction type, interface type, activities it will support, functions, relationships between functions, and information requirements. Of these conceptual models, decide which one seems most appropriate and articulate the reasons why.

b. Produce the following prototypes for your chosen conceptual model:

   i. Using the scenarios generated for the online booking facility, produce a storyboard for the task of booking a ticket for one of your conceptual models. Show it to two or three potential users and get some informal feedback.

   ii. Now develop a card-based prototype from the use case for the task of booking a ticket, also incorporating feedback from part (i). Show this new prototype to a different set of potential users and get some more informal feedback.

c. Consider your product's concrete design. Sketch out the application's landing page. Consider the layout, use of colors, navigation, audio, animation, etc. While doing this, use the three main questions introduced in Chapter 6 as guidance: Where am I? What's here? Where can I go? Write one or two sentences explaining your choices, and consider whether the choice is a usability consideration or a user experience consideration.

d. Sketch out an experience map for your product. Use the scenarios and personas you have already generated to explore the user's experience. In particular, identify any new interaction issues that you had not considered before, and suggest what you could do to address them.

e. How does your product differ from applications that typically might emerge from the Maker Movement? Do software development kits have a role? If so, what is that role? If not, why do you think not?

Take a Quickvote on Chapter 11:
www.id-book.com/quickvotes/chapter11

Summary

This chapter has explored the activities of design, prototyping, and construction. Prototyping and scenarios are used throughout the design process to test out ideas for feasibility and user acceptance. We have looked at different forms of prototyping, and the activities have encouraged you to think about and apply prototyping techniques in the design process.

Key points

- Prototyping may be low fidelity (such as paper-based) or high fidelity (such as software-based).
- High-fidelity prototypes may be vertical or horizontal.
- Low-fidelity prototypes are quick and easy to produce and modify and are used in the early stages of design.
- Ready-made software and hardware components support the creation of prototypes.
- There are two aspects to the design activity: conceptual design and concrete design.
- Conceptual design develops an outline of what people can do with a product and what concepts are needed to understand how to interact with it, while concrete design specifies the details of the design such as layout and navigation.
- We have explored three approaches to help you develop an initial conceptual model: interface metaphors, interaction styles, and interface styles.
- An initial conceptual model may be expanded by considering which functions the product will perform (and which the user will perform), how those functions are related, and what information is required to support them.
- Scenarios and prototypes can be used effectively in design to explore ideas.
- Physical computing kits and software development kits facilitate the transition from design to construction.

Further Reading

BANZI, M. and SHILOH, M. (2014) Getting started with Arduino (3rd edn). Maker Media Inc. This hands-on book provides an illustrated step-by-step guide to learning about Arduino with lots of ideas for projects to work on. It outlines what physical computing is in relation to interaction design and the basics of electricity, electronics, and prototyping using the Arduino hardware and software environment.

CARROLL, J. M. (ed.) (1995) Scenario-based Design. John Wiley & Sons, Inc. This volume is an edited collection of papers arising from a three-day workshop on use-oriented design. The book contains a variety of papers including case studies of scenario use within design, and techniques for using them with object-oriented development, task models, and usability engineering. This is a good place to get a broad understanding of this form of development.

GREENBERG, S., CARPENDALE, S., MARQUARDT, N. and BUXTON, B. (2012) Sketching User Experiences. Morgan Kaufman. This is a practical introduction to sketching. It explains why sketching is important and provides very useful tips to get the reader into the habit of sketching. It is a companion book to Buxton, B. (2007) Sketching User Experiences. Morgan Kauffman, San Francisco.

LAZAR, J. (ed.) (2007) Universal Usability: Designing information systems for diverse user populations. John Wiley & Sons Ltd. This book provides an interesting selection of case studies that demonstrate how developers can design for diverse populations to ensure universal usability.

Interview with the late Gary Marsden

Gary Marsden died suddenly and unexpectedly in December 2013. He was only 43. He was a professor in the Computer Science Department at the University of Cape Town. His research interests spanned mobile interaction, computer science, design and ICT for Development. He is a co-author of a book published in 2015, with Matt Jones and Simon Robinson, entitled, There's Not an App for That: Mobile User Experience Design for Life. He was also a co-author of Mobile Interaction Design, which was published in 2006. He won the 2007 ACM SIGCHI Social Impact Award for his research in using mobile technology in the developing world. He made a big impression on the HCI world. We have decided to keep his interview from the 3rd edition.

Gary, can you tell us about your research and why you do it?

My work involves creating digital technology for people living in Africa. Most of this work is based on designing software and interfaces for mobile cellular handsets as this is currently the most prevalent digital technology within Africa.

Because the technology is deployed in Africa, we work within a different design space than those working in more developed parts of the world. For instance, we assume that users have no access to personal computers or high-speed Internet connections. We must also take into account different literacy levels in our users and the cultures from which they come. Not only does this affect the technology we create, but the methods we use to create it.

As a computer science professional, I want to understand how to create digital systems that are relevant and usable by the people purchasing them. For many people here, buying a cellular handset is a significant investment and I want to make sure that the discipline of interaction design is able to help deliver a product which maximizes the purchaser's investment.

How do you know if the systems that you build are what people want and need?

This is currently a hotly debated topic in the field and it is only recently that there has been sufficient work from which to draw conclusions.

The first challenge crops up in designing a system for people who have very little exposure to technology. For many of our users, they have no experience of digital technology beyond using a simple cellular handset. Therefore, participatory techniques, where users are asked to become co-designers, can be problematic as they have no abstract notions of basic ideas like the separation between hardware and software. To overcome this, we often take a technology probe approach, allowing users to comment on a high-fidelity prototype rather than require them to make abstract decisions about a series of paper sketches.

For many of the systems we build, we are interested in more than simple measures of efficiency and effectiveness. Sure, it is important that technology is usable, but in the resource-constrained environment, it is critical that the technology is useful; money is too scarce to spend on something that does not significantly improve livelihood.

To measure impact on people and communities we often borrow from the literature on development and measure issues like domestification – the extent to which a technology is appropriated into someone's day-to-day living. In a lot of our work we also partner with non-governmental organizations (NGOs) who are based in a community and are looking for research partners to provide digital solutions to problems they meet – for instance, we have worked with a voter education NGO that wanted to use digital technology to better inform voters about their choices in an upcoming election. In that project we would adopt the goals of the NGO (how much people understand their voting choices) as part of the success criteria for our project. Often NGOs have sophisticated instruments to measure the impact they are having, as their funding relies on it. We can use those instruments to measure our impact.

To understand how our participants truly feel about a system, we use 'polyphonic' assessment, as reported by Bill Gaver. The method employs unbiased journalists who interview users and report their assessment of the system. We have adopted this approach in our work and found it to be highly effective in gaining feedback on our systems. Furthermore, it overcomes a strong Hawthorne effect experienced by researchers who work in resource poor environments – users are so grateful for the attention and resources being given them, they rate any system highly in an attempt to please the researchers and keep them investing in that community.

At present, there is no clear consensus about how best to evaluate technology deployments in developing world communities, but it is clear that the technology cannot be evaluated solely on a human–computer interaction level, but needs to be considered on a livelihoods and community impact level.

Have you encountered any big surprises in your work?

My work seems to be endlessly surprising which, as a researcher, is highly stimulating. The first surprise when I moved here 12 years ago, was the penetration of mobile handsets. In an era when handsets were considered a luxury in Europe (1999), I saw people living in shacks talking on their mobile handsets. Clearly domestification was not an issue for cellular technology.

When I started to run research projects in Africa, I was surprised by the extent to which much HCI research and methods incorporated assumptions based in the developed world – for example, the issue I mentioned earlier around participatory design. Also, the early HCI literature I read on the internationalization of interfaces did not stand me in good stead. For example, my colleague, Marion Walton, built one interface consisting of a single button on a screen. We asked participants to click on the button, but one participant was unable to do this. When we pointed out the button to him, he said, 'That is not a button, that is a picture of a button.' Of course, he was correct and we learnt something valuable that day about visual culture.

Finally, the environment in Africa leads to surprises. The strangest problem I have had was trying to fix a computer in rural Zambia that had suddenly stopped working. On taking the casing off, I discovered white ants had eaten the green resin out of the circuit board and used it to build a nest over the power supply (where it was warm). Although it now looked like a beautiful lace, the motherboard could not be salvaged.

What are your hopes for the future?

My hope and my passion are to create a new generation of African computer scientists who create technology for their continent. Whilst the work I am engaged in may be helping to some small degree, it is not sustainable for outside people or teams to create new technology for everyone who lives in the developing world. As an educator, I believe the solution is to teach interaction design in African universities and empower Africans to create the technology that is most appropriate to them and their environment. ☐

Chapter 13
Introducing Evaluation

Objectives

The specific aims of this chapter are to:

- Explain the key concepts and terms used in evaluation.
- Introduce a range of different types of evaluation methods.
- Show how different evaluation methods are used for different purposes at different stages of the design process and in different contexts of use.
- Show how evaluators mix and modify methods to meet the demands of evaluating novel systems.
- Discuss some of the practical challenges that evaluators have to consider when doing evaluation.
- Illustrate through short case studies how methods discussed in more depth in Chapters 7 and 8 are used in evaluation and describe some methods that are specific to evaluation.

13.1 Introduction

Imagine you have designed an app for teenagers to share music, gossip, and photos. You have prototyped your first design and implemented the core functionality. How would you find out whether it would appeal to them and if they will use it? You would need to evaluate it – but how? This chapter presents an introduction to the main types of evaluation and the methods you can use.

Evaluation is integral to the design process. Evaluators collect information about users' or potential users' experiences when interacting with a prototype, an app, a computer system, a component of a computer system, an application, or a design artifact such as a screen sketch. They do this in order to improve its design. Evaluation focuses on both the usability of the system (e.g. how easy it is to learn and to use) and on the users' experience when interacting with it (e.g. how satisfying, enjoyable, or motivating the interaction is).

Devices like smartphones, iPods, iPads, and e-readers have heightened awareness about usability, but many designers still assume that if they and their colleagues can use a product and find it attractive, others will, too. The problem with this assumption is that designers may design only for themselves. Evaluation enables them to check that their design is appropriate and acceptable for the wider user population.

There are many different evaluation methods. Which to use depends on the goals of the evaluation. Evaluations can occur in a range of places such as laboratories, people's homes, outdoors, and work settings. Evaluations usually involve observing participants and measuring their performance – in usability testing, experiments, or field studies. There are other methods, however, that do not involve participants, such as modeling user behavior.

These tend to be approximations of what users might do when interacting with an interface, often done as a quick and cheap way of assessing different interface configurations. The level of control on what is evaluated varies; sometimes there is none, such as in field studies, and in others there is considerable control over which tasks are performed and the context, such as in experiments.

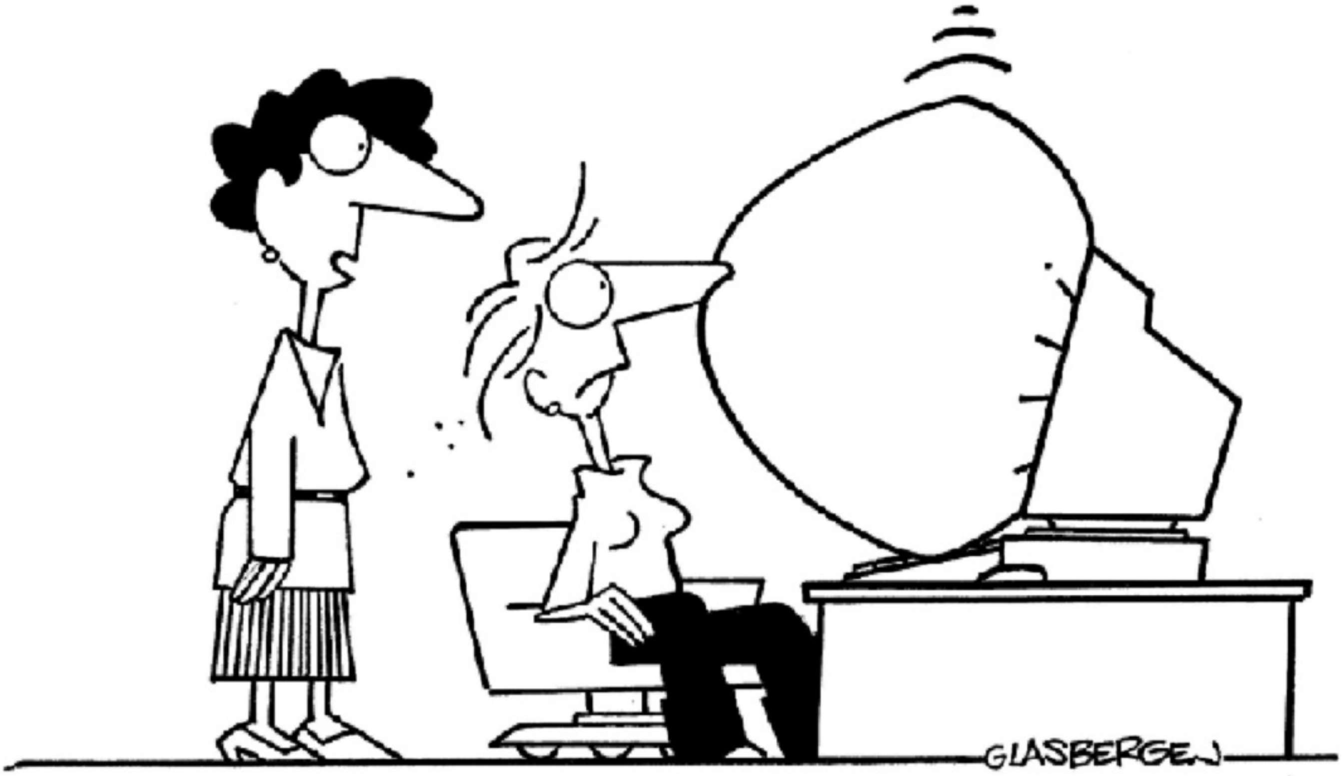
In this chapter we discuss why evaluation is important, what needs to be evaluated, where evaluation should take place, and when in the product lifecycle evaluation is needed. The different types are then illustrated by short case studies.

13.2 The Why, What, Where, and When of Evaluation

Conducting evaluations involves understanding not only why evaluation is important but also what aspects to evaluate, where evaluation should take place, and when to evaluate.

13.2.1 Why Evaluate?

Nowadays users expect much more than just a usable system; they also look for a pleasing and engaging experience. This means it is even more important to carry out an evaluation. As the Nielsen Norman Group (www.nngroup.com) notes, "User experience encompasses all aspects of the end-user's interaction . . . the first requirement for an exemplary user experience is to meet the exact needs of the customer, without fuss or bother. Next come simplicity and elegance, which produce products that are a joy to own, a joy to use."



"it's the latest innovation in office safety. When your computer crashes, an air bag is activated so you won't bang your head in frustration."

From a business and marketing perspective, well-designed products sell. Hence, there are good reasons for companies investing in evaluation. Designers can focus on real problems and the needs of different user groups rather than debating what each other likes or dislikes. It also enables problems to be fixed before the product goes on sale.

Activity 13.1

Identify one adult and one teenager prepared to talk with you about their Facebook usage (these may be family members or friends). Ask them questions such as: How often do you look at Facebook each day? How many photos do you post? What kind of photos do you have in your albums? What photo do you have as your profile picture? How often do you change it? How many friends have you got? What books and music do you list? Are you a member of any groups?

Comment

Show/Hide

13.2.2 What to Evaluate

What to evaluate ranges from low-tech prototypes to complete systems; a particular screen function to the whole workflow; and from aesthetic design to safety features. For example, developers of a new web browser may want to know if users find items faster with their product, whereas developers of an ambient display may be interested in whether it changes people's behavior, and game app developers will want to know how engaging and fun their games are and how long users will play them for. Government authorities may ask if a computerized system for controlling traffic lights results in fewer accidents or if a website complies with the standards required for users with disabilities. Makers of a toy may ask if 6-year-olds can manipulate the controls and whether they are engaged by its furry case, and whether the toy is safe for them to play with. A company that develops personal, digital music players may want to know if the size, color, and shape of the casing are liked by people from different age groups living in different countries. A software company may want to assess market reaction to its new homepage design.

Activity 13.2

What aspects would you want to evaluate for the following systems:

1. a personal music player (e.g. a smartphone app)?
2. a website for selling clothes?

Comment

Show/Hide

13.2.3 Where to Evaluate

Where evaluation takes place depends on what is being evaluated. Some characteristics, such as web accessibility, are generally evaluated in a laboratory, because it provides the control necessary to systematically investigate whether all requirements are met. This is also true for design choices, such as choosing the size and layout of keys for a smartphone. User experience aspects, such as whether children enjoy playing with a new toy and for how long before they get bored, can be evaluated more effectively in natural settings which are often referred to as ‘in the wild studies'. Remote studies of online behavior, such as social networking, can be conducted to evaluate natural interactions for a lot of participants in their own homes, cheaply and quickly. Living laboratories have also been built which are somewhere in between labs and in the wild settings; providing the setting of being in an environment, such as the home, while also giving the ability to control, measure, and record activities.

Activity 13.3

A company is developing a new car seat to monitor if a person starts to fall asleep when driving and to provide a wake-up call using olfactory and haptic feedback. Where would you evaluate it?

Comment

Show/Hide

13.2.4 When to Evaluate

At what stage in the product lifecycle evaluation takes place depends on the type of product. For example, the product being developed could be a brand-new concept or it could be an upgrade of an existing product. If the product is new, then considerable time is usually invested in market research and establishing user requirements. Once these requirements have been established, they are used to create initial sketches, a storyboard, a series of screens, or a prototype of the design ideas. These are then evaluated to see if the designers have interpreted the users' requirements correctly and embodied them in their designs appropriately. The designs will be modified according to the evaluation feedback and new prototypes developed and subsequently evaluated.

When evaluations are done during design to check that a product continues to meet users' needs, they are known as formative evaluations. Formative evaluations cover a broad range of design processes, from the development of early sketches and prototypes through to tweaking and perfecting an almost finished design.

Evaluations that are done to assess the success of a finished product are known as summative evaluations. If the product is being upgraded then the evaluation may not focus on establishing a set of requirements, but may evaluate the existing product to ascertain what needs improving. Features are then often added, which can result in new usability problems. Other times, attention is focused on improving specific aspects, such as enhanced navigation.

Many agencies such as the National Institute of Standards and Technology (NIST) in the USA, the International Standards Organization (ISO), and the British Standards Institute (BSI) set standards by which particular types of products, such as aircraft navigation systems and consumer products that have safety implications for users, have to be evaluated.

13.3 Types of Evaluation

We classify evaluations into three broad categories, depending on the setting, user involvement, and level of control. These are:

1. Controlled settings involving users (examples are laboratories and living labs): users' activities are controlled in order to test hypotheses and measure or observe certain behaviors. The main methods are usability testing and experiments.

2. Natural settings involving users (examples are online communities and products that are used in public places): there is little or no control of users' activities in order to determine how the product would be used in the real world. The main method used is field studies.
3. Any settings not involving users: consultants and researchers critique, predict, and model aspects of the interface in order to identify the most obvious usability problems. The range of methods includes inspections, heuristics, walkthroughs, models, and analytics.

There are pros and cons of each type. For example, lab-based studies are good at revealing usability problems but poor at capturing context of use; field studies are good at demonstrating how people use technologies in their intended setting but are expensive and difficult to conduct (Rogers et al, 2007, Rogers et al, 2013); and modeling and predicting approaches are cheap and quick to perform but can miss unpredictable usability problems and subtle aspects of the user experience.

A key concern for deciding on which approach to use is how much control is needed in order to find out how an interface or device is used. For the example of the teenager music app mentioned earlier, we would need to find out how teenagers use it, whether they like it, what problems they experience with the functions, and so on. This requires determining how they carry out various tasks using the interface operations, which will involve a degree of control in designing the evaluation study to ensure they try out all the tasks and operations the app has been designed for.

13.3.1 Controlled Settings Involving Users

Controlled settings enable evaluators to control what users do, when they do it, and for how long. They also enable evaluators to reduce outside influences and distractions, such as friends or colleagues talking. The approach has been extensively and successfully used to evaluate software applications running on PCs and other technologies where participants can be seated in front of them to perform a set of tasks.

Usability testing.

Typically this approach to evaluating user interfaces involves collecting data using a combination of methods – i.e. experiments, observation, interviews, and questionnaires – in a controlled setting. It is generally done in laboratories although increasingly it is being done remotely or in natural settings. The primary goal is to determine whether an interface is usable by the intended user population to carry out the tasks for which it was designed.

This involves investigating how typical users perform on typical tasks. By typical we mean the users for whom the system is designed (e.g. teenagers) and the things that it is designed for them to be able to do (e.g. buy the latest fashions). It often involves comparing the number and kinds of errors that the users make and recording the time that it takes them to complete the task. As users perform the tasks, they may be recorded on video; their interactions with the software may also be recorded, usually by logging software. User satisfaction questionnaires and interviews can also be used to elicit users' opinions about how they found the experience of using the system. It can be supplemented by observation at product sites to collect evidence about how the product is being used in work or other environments.

Observing users' reactions to an interactive product has helped developers understand usability issues that would be extremely difficult for them to glean simply through reading reports, or listening to presentations. For many years usability testing has been a staple of HCI, being used in the development of standard products that go through many generations, such as word processing systems, databases, and spreadsheets (Johnson, 2014; Krug, 2014; Redish, 2012; Koyani et al, 2004). The findings from the usability testing are often summarized in a usability specification that enabled developers to test future prototypes or versions of the product against it. Optimal performance levels and minimal levels of acceptance are generally specified and current levels noted. Changes in the design can then be implemented, such as navigation structure, use of terms, and how the system responds to the user.

Madrigal and McClain (2010) provide practical guidance including a list of the dos and don'ts of doing usability testing. They also point out how "usability testing is one of the least glamorous, but most important aspects of user experience research."

Experiments are typically conducted in research labs in universities or industry to test hypotheses. They are the most controlled setting, where researchers try to remove any extraneous variables that may interfere with the participant's performance. The reason for this is so that they can reliably say that the findings arising from the experiment are due to the particular interface feature being measured. For example, an experiment comparing which is the best way for users to enter text when using a tablet interface would control all other aspects of the setting to ensure they do not affect the performance. These include providing the same instructions to the participants, using the same tablet interface, and asking the participants to do the same tasks. The conditions that might be compared could be: typing using a virtual keyboard, typing using a physical keyboard, and swiping using a virtual keyboard (sliding across the keys to select letters, see Chapter 6). The aim of the experiment would be to test whether one is better than the other in terms of speed of typing and number of errors. A number of participants would be brought into the lab separately to carry out a predefined set of text entry tasks and their performance measured in terms of time taken and any errors made, e.g. selecting the wrong letter. The data collected would then be analyzed to determine if the scores for each condition were significantly different. If the performance measures obtained for the virtual keyboard were significantly faster than the other two and had the least number of errors, it could be possible to say that this method of text entry is the best.

BOX 13.1

Living labs

Living labs have been developed to evaluate people's everyday lives – that would be simply difficult to assess in usability labs. For example, research has been conducted to investigate people's habits and routines over a period of several months. An early example of a living lab was the Aware Home, which was a technology-rich, experimental house that enabled research into people's lives to be conducted in an authentic yet experimental setting (Abowd et al, 2000). The house was embedded with a complex network of sensors and audio/video recording devices that recorded their movements throughout the house and their use of technology. This enabled the occupants' behavior to be monitored and analyzed e.g. their routines and deviations.

A primary motivation was to evaluate how real families would respond and adapt to such a setup, over a period of several months. However, it has proved difficult to get participant families to agree to leave their own homes and live in a living lab home for a period of time. Nowadays, many Living Labs have become more like commercial enterprises, which offer facilities, infrastructure, and access to participating communities, bringing together users, developers, researchers, and other stakeholders. For example, MIT's Living Labs "brings together interdisciplinary experts to develop, deploy and test – in actual living environments – new technologies and strategies for design that respond to the changing world" (livinglabs.mit.edu). This enables researchers and companies to observe and record a number of activities related to health, energy, and the environment. The lab is set up to simulate a setting (e.g. a home, a dance floor, a classroom) and fully functional prototypes are installed in order to let users experience future product and service ideas in these experimental authentic settings. Ambient assisted living homes have also been developed where a network of sensors is embedded throughout someone's home rather than in a special house. The rationale is to enable disabled and elderly people to lead safe and independent lives by providing a non-intrusive system that can remotely monitor and provide alerts to caregivers in the event of an accident, illness, or unusual activities (e.g. Fernández-Luque et al, 2009). The term 'living lab' is also used to describe innovation networks in which people gather in person and virtually to explore and form commercial research and development collaborations. ■

Dilemma

Is a living lab really a lab?

The concept of a living lab differs from a traditional view of a laboratory insofar as it is trying to be both natural and experimental, and where the goal is to bring the lab into the home (or other natural setting). The dilemma is how artificial do you make the more natural setting; where does the balance lie in setting it up to enable the right level of control to conduct evaluation without losing the sense of it being natural?■

13.3.2 Natural Settings Involving Users

The aim of field studies is to evaluate people in their natural settings. They are used primarily to: (i) help identify opportunities for new technology; (ii) establish the requirements for a new design; and (iii) facilitate the introduction of technology, or inform deployment of existing technology in new contexts. Methods that are typically used are observation, interviews and focus groups, and interaction logging (see Chapter 7). The data takes the form of events and conversations that are recorded by the researchers as notes, or by audio or video recording, or by the participants, as diaries and notes. A goal is to be unobtrusive and not to affect what people do during the evaluation. However, it is inevitable that some methods will influence how people behave. For example, diary studies require people to document their activities or feelings at certain times and this can make them reflect on and possibly change their behavior.

More recently, there has been a trend towards conducting in the wild studies in HCI and ubiquitous computing. These are essentially field studies that look at how new technologies or prototypes have been deployed and used by people in various settings, such as the outdoors, public places, and homes. In moving into the wild, researchers inevitably have to give up control of what is being evaluated in order to observe how people approach and use, or don't use technologies in their everyday lives. For example, a researcher might be interested in observing how a new mobile navigation device will be used in urban environments. To conduct the study they would need to recruit people who are willing to use the device for a few weeks or months in their natural surroundings. They might then tell the participants what they can do with the device. Other than that, it is up to the participants to decide how to use it and when, as they move between their work or school, home, and other places.

The downside of handing over control is that it makes it difficult to anticipate what is going to happen and to be present when something interesting does happen. This is in contrast to usability testing where there is always an investigator or camera at hand to record events (Rogers et al, 2013). Instead, the researcher has to rely on the participants recording and reflecting how they use it, by writing up their experiences in diaries, filling in online forms, and/or taking part in intermittent interviews. The participants, however, may get too busy to do this reliably or forget certain events when they come to doing it. In addition, the participants may not try out the full range of functions provided by the device – meaning that the researcher has no idea as to whether they are useful or usable. However, the fact that the participants do not use them can be equally revealing.

Field studies can also be virtual, where observations take place in multiuser games such as World of Warcraft, online communities, chat rooms, and so on (e.g. Bainbridge, 2010). A goal of this kind of field study is to examine the kinds of social processes that occur in them, such as collaboration, confrontation, and cooperation. The researcher typically becomes a participant and does not control the interactions (see Chapters 7 and 8). Similar to in the wild studies, which are a form of field study, it can be difficult to anticipate what is going to happen and where.

13.3.3 Any Settings Not Involving Users

Evaluations that take place without involving users are conducted in settings where the researcher has to imagine or model how an interface is likely to be used. Inspection methods are commonly employed to predict user behavior and to identify usability problems, based on knowledge of usability, users' behavior, the contexts in which the system will be used, and the kinds of activities that users undertake. Examples include heuristic evaluation that applies knowledge of typical users guided by rules of thumb and walkthroughs that involve stepping through a scenario or answering a set of questions for a detailed prototype. Other techniques include analytics and models.

The original heuristics used in heuristic evaluation were for screen-based applications (Nielsen and Mack, 1994; Nielsen and Tahir, 2002). These have been adapted to develop new sets of heuristics for evaluating web-based products, mobile systems, collaborative technologies, computerized toys, and other new types of systems. One of the problems of using heuristics is that designers can sometimes be led astray by findings that are not as accurate as they appeared to be at first (Cockton et al, 2002; Tomlin, 2010 and for an overview see: http://www.nngroup.com/topic/heuristic-evaluation/).

Cognitive walkthroughs, which were the first walkthroughs developed, involve simulating a user's problem-solving process at each step in the human–computer dialog, and checking to see how users progress from step to step in these interactions (see Wharton et al, 1994 in Nielsen and Mack, 1994). A key feature of cognitive walkthroughs is that they focus on evaluating designs for ease of learning.

As discussed in Chapters 7 and 8, analytics is a technique for logging data either at a customer's site or remotely. Web analytics is the measurement, collection, analysis, and reporting of Internet data in order to understand and optimize web usage (Arikan, 2008). For example, Google provides a commonly used approach for collecting analytics data (http://www.google.com/analytics/). The kind of data collected depends on the reasons why the evaluation is being conducted. Analytics is a particularly useful method for evaluating design features of a website. For example, a company may want to evaluate its audience potential or the buzz (comments) that are happening about the products described on its website. Using specially developed tools, transactions on the website can be analyzed from log files of interaction data (also now called big data). Analysis of which web pages are visited (i.e. page hits) can also be analyzed, and so can direct mail campaign data, sales and lead information, and user performance data. Using a variety of automated analysis tools, including statistical analysis and data visualization, variations in web traffic can be identified and so can the most popular pages and products. Recently, with an increased focus on web-based learning through Massive Open Online Courses (MOOCs) and other forms of distance learning, there is more focus on learning analytics to see how well people really do learn using online technology.

Models have been used primarily for comparing the efficacy of different interfaces for the same application, for example, the optimal arrangement and location of features. A well-known approach uses Fitts' Law to predict the time it takes to reach a target using a pointing device (MacKenzie, 1995).

13.3.4 Choosing and Combining Methods

The three broad categories identified above provide a general framework to guide the selection of evaluation methods. Often combinations of methods are used across the categories to obtain a richer understanding. For example, sometimes usability testing in labs is combined with observations in natural settings to identify the range of usability problems and find out how users typically use a product. Figure 13.1 illustrates one way in which laboratory-based usability testing and field studies in natural settings can be combined to evaluate a mobile application – e.g. for a smartphone.

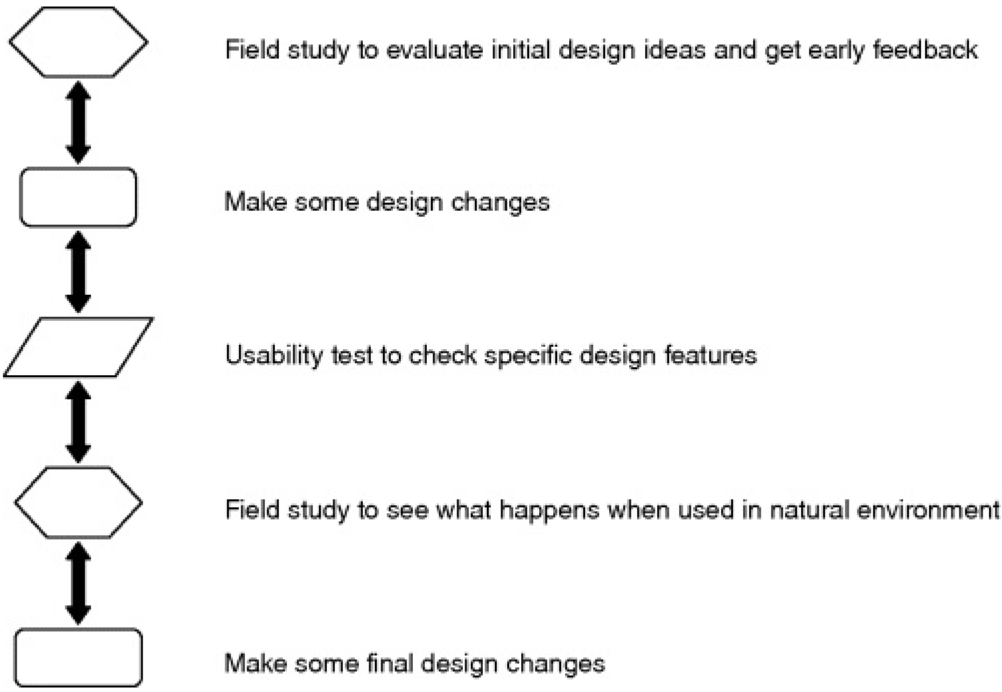


Figure 13.1 Example of the way laboratory-based usability testing and field studies can complement each other

How do you choose between using a controlled and uncontrolled setting? There are obviously pros and cons for each one. The benefits of controlled settings include being able to test hypotheses about specific features of the interface, where the results can be generalized to the wider population. A benefit of uncontrolled settings is that unexpected data can be obtained that provides quite different insights into people's perceptions and their experiences of using, interacting, or communicating through the new technologies in the context of their everyday and working lives.

13.3.5 Opportunistic Evaluations

Evaluations may be detailed, planned studies or opportunistic explorations. The latter are generally done early in the design process to provide designers with feedback quickly about a design idea. Getting this kind of feedback early in the design process is important because it confirms whether it is worth proceeding to develop an idea into a prototype. Typically, these early evaluations are informal and do not require many resources. For example, the designers may recruit a few local users and ask their opinions. Getting feedback this early in design can help save time and money if an idea needs to be modified or abandoned. Opportunistic evaluations with users can also be conducted in addition to more formal evaluations.

13.4 Evaluation Case Studies

Two contrasting case studies are described below to illustrate how evaluations can take place in different settings with different amounts of control over users' activities. The first case study (Section 13.4.1) describes an experiment that tested whether it was more exciting playing against a computer versus playing against a friend for a collaborative computer game (Mandryk and Inkpen, 2004). The focus was on how to measure user experience, which they note is more difficult to achieve than traditional usability measures. The second case study (Section 13.4.2) describes an in the wild evaluation of skiers (Jambon and Meillon, 2009) that used a combination of methods to assess how a mobile device providing personal feedback, on how well users are doing, helped improve their performance. The focus of the evaluation was whether and how often the participants consulted the feedback.

13.4.1 Case Study 1: An Experiment Investigating a Computer Game

For games to be successful they must engage and challenge users. Ways of evaluating this aspect of the user experience are therefore needed and, in this case study (Mandryk and Inkpen, 2004), physiological responses were used to evaluate users' experiences when playing against a friend and when playing alone against the computer. The researchers conjectured that physiological indicators could be an effective way of measuring a player's experience. Specifically, they designed an experiment to evaluate the participants' experience of playing an online ice-hockey game.

Ten participants, who were experienced game players, took part in the experiment. During the experiment sensors were placed on the participants to collect physiological data. These included measures of the moisture produced by sweat glands in the hands and feet, and changes in heart rate and breathing rate. In addition, they videoed participants and asked them to complete user satisfaction questionnaires at the end of the experiment. In order to reduce the effects of learning, half of the participants played first against a friend and then against the computer, and the other half played against the computer first. Figure 13.2 shows the setup for recording data while the participants were playing the game.

Source: Mandryk and Inkpen (2004) Physiological Indicators for the Evaluation of Co-located Collaborative Play, *CSCW'2004*, pp. 102–111. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

Figure 13.2 The display shows the physiological data (top right), two participants, and a screen of the game they played

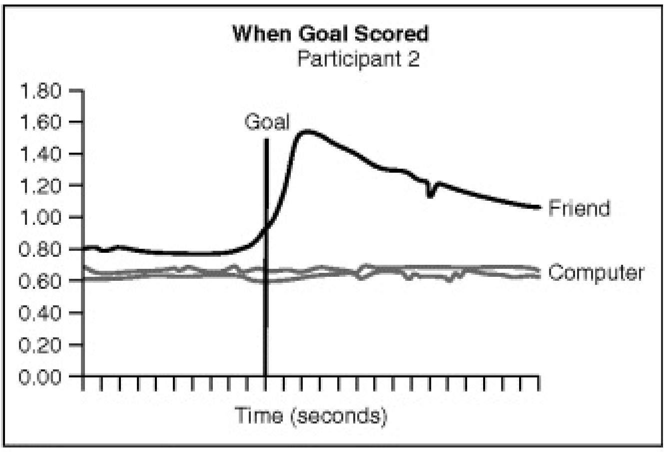
Source: Mandryk and Inkpen (2004) Physiological Indicators for the Evaluation of Co-located Collaborative Play, CSCW'2004, pp. 102–111. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

Results from the user satisfaction questionnaire revealed that the mean ratings on a 1–5 scale for each item indicated that playing against a friend was the favored experience (Table 13.1). Data recorded from the physiological responses was compared for the two conditions and in general revealed higher levels of excitement when participants played against a friend than when they played against the computer. The physiological recordings were also compared across participants and, in general, indicated the same trend. Figure 13.3 shows a comparison for two participants.

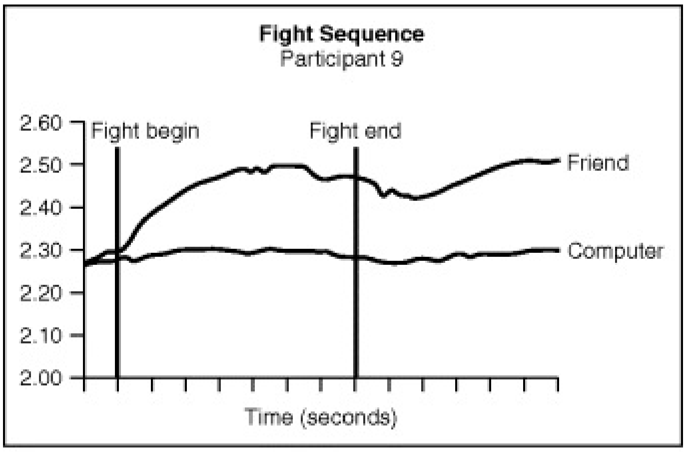
TABLE 13.1

Mean subjective ratings given on a user satisfaction questionnaire using a five-point scale, in which 1 is lowest and 5 is highest for the 10 players. Identifying strongly with an experience state is indicated by a higher mean. The standard deviation indicates the spread of the results around the mean. Low values indicate little variation in participants' responses, high values indicate more variation

	Playing against computer		Playing against friend	
	Mean	St. Dev.	Mean	St. Dev.
Boring	2.3	0.949	1.7	0.949
Challenging	3.6	1.08	3.9	0.994
Easy	2.7	0.823	2.5	0.850
Engaging	3.8	0.422	4.3	0.675
Exciting	3.5	0.527	4.1	0.568
Frustrating	2.8	1.14	2.5	0.850
Fun	3.9	0.738	4.6	0.699



Source: Mandryk and Inkpen (2004) Physiological Indicators for the Evaluation of Co-located Collaborative Play, *CSCW'2004*, pp. 102–111. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.



Source: Mandryk and Inkpen (2004) Physiological Indicators for the Evaluation of Co-located Collaborative Play, *CSCW'2004*, pp. 102–111. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

Figure 13.3 (a) A participant's skin response when scoring a goal against a friend versus against the computer and (b) another participant's response when engaging in a hockey fight against a friend versus against the computer

Source: Mandryk and Inkpen (2004) Physiological Indicators for the Evaluation of Co-located Collaborative Play, CSCW'2004, pp. 102–111. ©2004 Association for Computing Machinery, Inc. Reprinted by permission.

Because of individual differences in physiological data, it was not possible to directly compare the means of the two sets of data collected: subjective questionnaires and physiological measures. However, by normalizing the results it was possible to correlate the results across individuals. This indicated that the physiological data gathering and analysis methods were effective for evaluating levels of challenge and engagement. Although not perfect, these two kinds of measures offer a way of going beyond traditional usability testing in an experimental setting to get a deeper understanding of user experience goals.

Activity 13.4

1. What kind of setting was used in this experiment?
2. How much control did the evaluators exert?
3. Which methods were recorded and when?

Comment

Show/Hide

13.4.2 Case Study 2: In the Wild Study of Skiers

Jambon and Meillon (2009) carried out an in the wild study to evaluate whether and how skiers might use a mobile device the authors designed to help skiers improve their performance. Each skier wore a helmet that had an accelerometer and a mini-camera on top of it (Figure 13.4a). These were used to gather data that could be used to provide feedback of the skiers' performance, which were displayed on a smartphone (Figure 13.4b). The skiers had access to the smartphones while on the slopes – which they kept in their pockets.

[Figure 13.4](#) **(a) A skier wearing a helmet with an accelerometer (dark red box) and a mini-camera (black cylinder) placed on it for assessing the skier's performance and (b) the smartphone that provides feedback to the skier in the form of visualizations**

Source: Jambon and Meillon (2009) User experience in the wild. In: Proceedings of CHI '09, ACM Press, New York, p. 4070.

The study examined how the mobile system was used by the participants while skiing. A series of trials were run in which skiers descended the mountain. Video clips from the mini-camera and data from the accelerometers were collected for each skier's descent. The skiers were then asked to enter a chalet where the research team downloaded this data. The skiers then received SMS messages telling them that their data could be viewed on their smartphones. This included: maps of their ski runs, distance covered, duration of descent, maximum speed, and the video recorded.

Figure 13.5 shows how the different components were linked together.

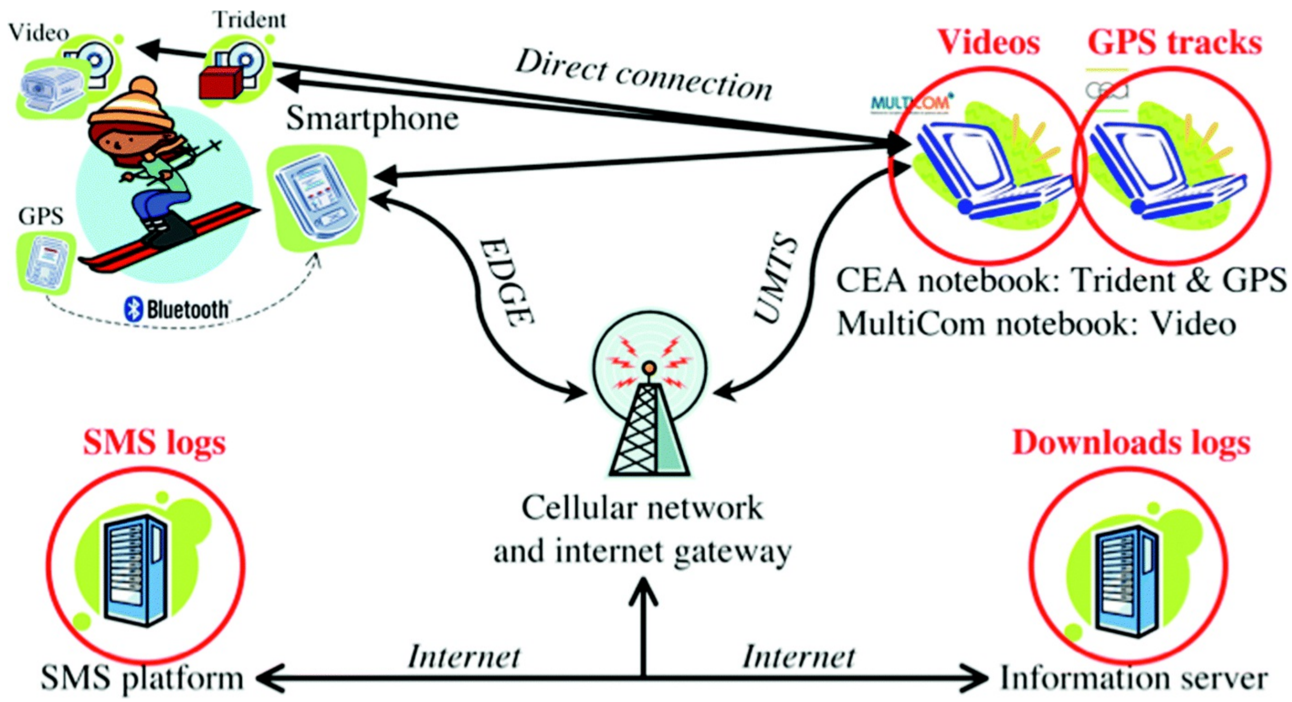


Figure 13.5 Components of the e-skiing system. Back arrows indicate the data transfers between devices, servers, and linking systems. Arrow shapes indicate different types of communications and the red circles indicate the data collection points

Source: Jambon et al (2009) User experience in the wild. In: Proceedings of CHI '09, ACM Press, New York, p. 4070.

When and how often the skiers consulted their smartphones for feedback was logged. To the great surprise of the evaluators, the skiers did not check their performance on the slopes. Instead they preferred to wait and review it in the bar during breaks. This shows how in the wild studies can reveal unexpected findings.

Approximately a week after the ski trials, the evaluators ran a focus group with the skiers in order to learn how they felt about the system. This was organized as an informal dinner at which the skiers confirmed that they preferred to get their feedback after skiing on the slopes, so that their time on the slopes was not interrupted. The skiers also discussed the problems associated with using the equipment on the slopes. For example, the Bluetooth links between the GPS system and the smartphones were not reliable and there were other technical problems too.

Activity 13.5

1. What kind of setting was used in this evaluation?
2. How much control did the evaluators exert?
3. Which types of data were collected?

Comment

Show/Hide

BOX 13.2

Crowdsourcing

Imagine having access to hundreds of thousands of participants who will perform tasks or provide feedback on a design or experimental task quickly and almost immediately. The service Mechanical Turk, hosted by Amazon, has thousands of people registered (known as Turkers), who have volunteered to take part by performing various activities, online, known as human intelligence tasks (HITs) for a very small reward. HITs are submitted by researchers or companies who pay from $0.01 for simple tasks (such as tagging pictures) to a few dollars (for taking part in an experiment). An advantage of using crowdsourcing in HCI is that it can be less expensive to run than traditional lab studies (where typically a participant is paid between $15 and $30, depending on the length of the experiment). Another benefit is that potentially many more participants can be recruited.

Heer and Bostock (2010) used crowdsourcing to determine how reliable it was to conduct an experiment in this way, i.e. asking random people over the Internet to take part. Using Mechanical Turk, they asked the Turkers signed up to the service to take part in a series of perception tasks using different visual display techniques. A large number agreed, enabling them to analyze their results statistically and generalize from their findings. They also developed a short set of test questions which generated 2880 responses. They then compared their findings from using crowdsourcing with those reported in published lab-based experiments. They found that while the results from their study using Turkers showed wider variance than in the reported study, the overall results across the studies were the same. They also found that the total cost of their experiment was a sixth of the cost of a typical laboratory study involving the same number of people. Crowdsourcing has become popular in a wide variety of applications ranging from planned experiments like Heer and Bostock's study (2010) to collecting design ideas and data from visitors at a nature park (Maher et al, 2014). ■

13.5 What Did We Learn from the Case Studies?

The case studies provide examples of how different evaluation methods are used in a variety of physical settings that involve users in different ways to answer various kinds of questions. They demonstrate how evaluators exercise control in different settings and also how they need to be creative when working with innovative systems and dealing with constraints created by the evaluation setting, and the robustness of the technology being evaluated. In addition, the case studies demonstrated:

- How to observe users in natural settings.
- Unexpected findings resulting from in the wild studies.
- Having to develop different data collection and analysis techniques to evaluate user experience goals such as challenge and engagement.
- The ability to run experiments on the Internet that are quick and inexpensive using crowdsourcing.
- How to recruit a large number of participants using Mechanical Turk.

BOX 13.3

The language of evaluation

Sometimes terms describing evaluation are used interchangeably and have different meanings. In order to avoid confusion we define some of these terms below in alphabetical order. (You may find that other texts use different terms).

Analytics: Data analytics refers to examining large volumes of raw data with the purpose of drawing inferences about that information. Web analytics is commonly used to measure website traffic through analyzing users' click data.

Analytical evaluation: Evaluation methods that model and predict user behavior. This term has been used to refer to heuristic evaluation, walkthroughs, modeling, and analytics.

Bias: The results of an evaluation are distorted. This can happen for several reasons. For example, selecting a population of users that has already had experience with the new system and describing their performance as if they were new users.

Controlled experiment: A study that is conducted to test hypotheses about some aspect of an interface or other dimension. Aspects that are controlled typically include the task that participants are asked to perform, the amount of time available to complete the tasks, and the environment in which the evaluation study occurs.

Crowdsourcing: A web-based method that provides the opportunity to enable potentially hundreds, thousands, or even millions of people to evaluate a product or take part in an experiment. The crowd may be asked to perform a particular evaluation task using a new product, or to rate or comment on the product.

Ecological validity: A particular kind of validity that concerns how the environment in which an evaluation is conducted influences or even distorts the results.

Expert review or crit: An evaluation method in which one or more people with usability expertise and knowledge of the user population review a product looking for potential problems.

Field study: An evaluation study that is done in a natural environment such as in a person's home, or in a work or leisure place.

Formative evaluation: An evaluation that is done during design to check that the product fulfills requirements and continues to meet users' needs.

Heuristic evaluation: An evaluation method in which knowledge of typical users is applied, often guided by heuristics, to identify usability problems.

Informed consent form: A form describing what a participant in an evaluation study will be asked to do, what will happen to the data collected about them, and their rights while involved in the study.

In the wild study: A field study in which users are observed using products or prototypes within their everyday context.

Living laboratory: A place that is configured to measure and record people's everyday activities in a natural setting, such as the home.

Predictive evaluation: Evaluation methods in which theoretically based models are used to predict user performance.

Reliability: The reliability or consistency of a method is how well it produces the same results on separate occasions under the same circumstances.

Scope: Refers to how much the findings from an evaluation can be generalized.

Summative evaluation: An evaluation that is done when the design is complete.

Usability laboratory: A laboratory that is specially designed for usability testing.

Usability testing: Involves measuring users' performance on various tasks.

User studies: A generic term that covers a range of evaluations involving users, including field studies and experiments.

Users or participants: These terms are used interchangeably to refer to the people who take part in evaluation studies.

Validity: Validity is concerned with whether the evaluation method measures what it is intended to measure. ■

13.6 Other Issues to Consider when Doing Evaluation

When reading the case studies, you probably thought of other issues, such as the importance of asking a good question to focus the evaluation and help you to decide on the best approach and methods to use. You may also have wondered about how to find suitable participants and, having found them, how to approach them. Can you just ask children in a café to participate or do you need permission from their parents? What do you have to tell participants and what if they decide part way through the study that they don't want to continue? Can they stop or do they have to continue? These and other issues are discussed on the ID-Book 4th Edition website but two things for you to think about now are: informing participants about their rights, and making sure that you take account of biases and other influences that impact how you describe your evaluation findings.

13.6.1 Informing Participants about Their Rights and Getting Their Consent

Most professional societies, universities, government, and other research offices require researchers to provide information about activities in which human participants will be involved. They do this to protect participants by ensuring that they are not endangered physically (e.g. in medical studies) or emotionally and that their right to privacy is protected. This documentation is reviewed by a panel and the researchers are notified whether their plan of work, particularly the details about how human participants and data collected about them will be treated, is acceptable. Drawing up such an agreement is mandatory in many universities and major organizations. Indeed, special review boards generally prescribe the format required and many provide a detailed form that must be completed. Once the details are accepted, the review board checks periodically in order to oversee compliance. In American universities these are known as Institutional Review Boards (IRB). Other countries use different names and different forms for similar processes. Over the years IRB forms have become increasingly detailed, particularly now that much research involves the Internet and people's interaction via communication technologies. Several lawsuits at prominent universities have heightened attention to IRB compliance to the extent that it sometimes takes several months and multiple amendments to get IRB acceptance. IRB reviewers are not only interested in the more obvious issues of how participants will be treated and what they will be asked to do; they also want to know how the data will be analyzed and stored. For example, data about subjects must be coded and stored to prevent linking participants' names with that data. This means that names must be replaced by codes or pseudonyms that must be stored separately from the data and stored in a secure location.

Participants in evaluation studies have to be told what they will be asked to do, the conditions under which data will be collected, and what will happen to their data when they finish the task. Participants are told their rights, e.g. that they may withdraw from the study at any time if they wish. As discussed in Chapter 7, this information is usually presented in a form that the participant reads and signs before the study starts. This form is often called an informed consent form.

13.6.2 Some Things that Influence how You Interpret Data

Decisions have to be made about what data is needed to answer the study questions, how the data will be analyzed, and how the findings will be presented (see Chapter 8). To a great extent the method used determines the type of data collected, but there are still some choices. For example, should the data be treated statistically? Some general questions also need to be asked. Is the method reliable? Will the method measure what is intended, i.e. what is the validity of the method and the data collected? Will the evaluation study be ecologically valid or is the fundamental nature of the process being changed by studying it? Are biases creeping in that will distort the results? Will the results be generalizable, i.e. what is their scope?

Reliability

The reliability or consistency of a method is how well it produces the same results on separate occasions under the same circumstances. Another evaluator or researcher who follows exactly the same procedure should get similar results. Different evaluation methods have different degrees of reliability. For example, a carefully controlled experiment will have high reliability, whereas observing users in their natural setting will be variable. An unstructured interview will have low reliability: it would be difficult if not impossible to repeat exactly the same discussion.

Validity

Validity is concerned with whether the evaluation method measures what it is intended to measure. This encompasses both the method itself and the way it is implemented. If, for example, the goal of an evaluation study is to find out how users use a new product in their homes, then it is not appropriate to plan a laboratory experiment. An ethnographic study in users' homes would be more appropriate. If the goal is to find average performance times for completing a task, then a method that only recorded the number of user errors would be invalid.

Ecological Validity

Ecological validity is a particular kind of validity that concerns how the environment in which an evaluation is conducted influences or even distorts the results. For example, laboratory experiments are controlled, so what the participants do and how they behave is quite different from what happens naturally in their workplace, at home, or in leisure environments. Laboratory experiments therefore have low ecological validity because the results are unlikely to represent what happens in the real world. In contrast, ethnographic studies do not impact the participants or the study location as much, so they have high ecological validity.

Ecological validity is also affected when participants are aware of being studied. This is sometimes called the Hawthorne effect after a series of experiments at the Western Electric Company's Hawthorne factory in the USA in the 1920s and 1930s. The studies investigated changes in length of working day, heating, lighting, and so on; however, eventually it was discovered that the workers were reacting positively to being given special treatment rather than just to the experimental conditions. Similar findings sometimes occur in medical trials. Patients given the placebo dose (a false dose in which no drug is administered) show improvement that is due to receiving extra attention that makes them feel good.

Biases

Bias occurs when the results are distorted. For example, expert evaluators performing a heuristic evaluation may be more sensitive to certain kinds of design flaws than others, and this will be reflected in the results. Evaluators collecting observational data may consistently fail to notice certain types of behavior because they do not deem them important. Put another way, they may selectively gather data that they think is important. Interviewers may unconsciously influence responses from interviewees by their tone of voice, their facial expressions, or the way questions are phrased, so it is important to be sensitive to the possibility of biases.

Scope

The scope of an evaluation study refers to how much its findings can be generalized. For example, some modeling methods, like Fitts' Law (discussed in Chapter 15) which is used to evaluate keypad design, have a narrow, precise scope. (The problems of overstating results are discussed in Chapter 8).

Assignment

In this assignment, think about the case studies and reflect on the evaluation methods used.

1. For the case studies think about the role of evaluation in the design of the system and note the artifacts that were evaluated, when during the design they were evaluated, which methods were used, and what was learned from the evaluations. Note any issues of particular interest. You may find that constructing a table like the one that follows is a helpful approach.

Name of the study or artifact evaluated	When during the design the evaluation occurred	How controlled was the study and what role did users have	Which methods were used	What kind of data was collected and how was it analyzed	What was learned from the study	Notable issues

2. What were the main constraints that influenced the evaluation?
3. How did the use of different methods build on and complement each other to give a broader picture of the evaluation?
4. Which parts of the evaluation were directed at usability goals and which at user experience goals?

Take a Quickvote on Chapter 13:
www.id-book.com/quickvotes/chapter13

Summary

The aim of this chapter was to introduce the main approaches to evaluation and the methods typically used. These will be revisited in more detail in the next two chapters. This chapter stressed how evaluation is done throughout design; collecting information about users' or potential users' experiences when interacting with a prototype, a computer system, a component of a computer system, or a design artifact (e.g. screen sketch) in order to improve its design.

The pros and cons of running lab-based versus in the wild studies were outlined, in terms of cost, effort, constraints, and the types of results that can be elicited. Choosing which approach to use will depend on the aims of the evaluation, and the researcher's or evaluator's expectations and the resources available to them. Crowdsourcing was presented as a creative, and sometimes cost-saving evaluation approach. Finally we briefly mentioned the ethical issues concerned with how evaluation participants are treated and their rights to privacy. We also raised questions about data interpretation including the need to be aware of biases, reliability, data and ecological validity, and the scope of the study.

Key points

- Evaluation and design are very closely integrated.
- Some of the same data gathering methods are used in evaluation as for establishing requirements and identifying users' needs, e.g. observation, interviews, and questionnaires.
- Evaluations can be done in controlled settings such as laboratories, less controlled field settings, or where users are not present.
- Usability testing and experiments enable the evaluator to have a high level of control over what gets tested, whereas evaluators typically impose little or no control on participants in field studies.
- Different methods are usually combined to provide different perspectives within a study.
- Participants are made aware of their rights through an informed consent form.
- It is important not to over-generalize findings from an evaluation.

Further Reading

ALBERT, B., TULLIS, T. and TEDESCO, D. (2010) Beyond the Usability Lab. Morgan Kaufmann. This book describes methods that can be used to plan, organize, conduct, and analyze specifically large-scale online user experience studies targeted at the professional UX market.

LAZAR, J., FENG, J. H. and HOCHHEISER, H. (2010) Research Methods in Human–Computer Interaction. John Wiley & Sons Ltd, Chichester, UK. This book provides a useful overview of qualitative and quantitative methods. Chapter 14, 'Working with Human Subjects,' discusses ethical issues of working with human participants.

ROGERS, Y., YUILL, N. and MARSHALL, P. (2013) Contrasting Lab-Based and In-the-Wild Studies For Evaluating Multi-User Technologies. In B. Price (2013) The SAGE Handbook on Digital Technology Research. SAGE Publications: 359–173. This chapter explores the pros and cons of lab-based and in-the-wild evaluation studies with reference to different types of technology platforms including tabletops and large wall displays.

SHNEIDERMAN, B. and PLAISANT, C. (2010) Designing the User Interface: Strategies for effective human–computer interaction (5th edn). Addison-Wesley. This text provides an alternative way of categorizing evaluation methods and offers a useful overview.

TULLIS, T. and ALBERT, B. (2008) Measuring the User Experience. Morgan Kaufmann. This book provides a more general treatment of usability testing.

WHITESIDE, J., BENNETT, J. and HOLTZBLATT, K. (1988) Usability engineering: our experience and evolution. In M. Helander (ed.) Handbook of Human–Computer Interaction. North Holland, pp. 791–817. Though written many years ago, this seminal article reviews the strengths and weaknesses of usability engineering and explains why ethnographic techniques can provide a valuable alternative, or supplement, in some circumstances.