
Thesis

How to cite:

© 2001 The Author

Version: Version of Record

Link(s) to article on publisher’s website:
http://ethos.bl.uk/OrderDetails.do?did=9&uin=uk.bl.ethos.365980

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
Non-visual Access to the World Wide Web: Investigations of Design Guidelines and Haptic Interfaces

Chaitya Colwell

A thesis submitted in partial fulfilment of the requirements of the
University of Hertfordshire
for the degree of
Doctor of Philosophy.

The programme of research was carried out in the

Sensory Disabilities Research Unit
of the
Department of Psychology, University of Hertfordshire.

In collaboration with MA Systems and Control Ltd., Eastleigh, Hants.

September 2001
Acknowledgements

I wish to acknowledge the on-going support and guidance of my Supervisors, Helen Petrie and Diana Kornbrot. I also acknowledge the financial support of the Economic and Social Research Council. In addition, I acknowledge the technical expertise of Andrew Hardwick of British Telecommunications plc and his assistance with the haptic device. Finally, I acknowledge the people who took part in the investigation of the haptic device, the evaluation of the Web Content Accessibility Guidelines, and the evaluation of Web site, without whom there would be no empirical data.

For their personal support, I acknowledge my partner Jeff Diggines, my Mum Dhamma Colwell, and the rest of my family and friends.

Thank you all.

Dedication

I dedicate this thesis to my partner, Jeff Diggines, without whose continued support and encouragement I would not have been able to complete this work.
# Table of Contents

Each chapter includes its own table of contents.

Abstract 5  
List of Figures 6  
List of Tables 6  
Chapter 1 — Introduction to Thesis 8  
Chapter 2 — Review of Literature on the use of Guidelines in User Interface Design 41  
Chapter 3 — Evaluation of Guidelines in Development of Web Pages 77  
Chapter 4 — Evaluation of Web Pages by Blind and Partially Sighted People 131  
Chapter 5 — Review of the Literature on Touch Perception and the use of Touch in Computer Interfaces 214  
Chapter 6 — Investigations of the Perception of Virtual Textures and Objects with a Haptic Interface 251  
Chapter 7 — Conclusions to the Thesis 288  
References 299  
Web addresses 305

# Appendices

Appendix 1: Purpose of new HTML tags (Chapter 3) 309  
Appendix 2: Questions sent to participants in the evaluation of the Guidelines (Chapter 3). 310  
Appendix 3: File given to participants working with tables and images (Chapter 3). 311  
Appendix 4: Questions asked of page author participants (Chapter 3). 315  
Appendix 5: Discussion with WAI on results of study (Chapter 3) 317  
Appendix 6: The software and hardware used by the visually impaired participants (Chapter 4). 325  
Appendix 7: Messages and questionnaire sent to visually impaired people (Chapter 4). 327  
Appendix 8: Screenshots of examples evaluated by visually impaired people (Chapter 4). 331  
Appendix 9: Extracts of the HTML code of the examples presented to the participants (Chapter 4). 352  
Appendix 10: Comparison of syntax used for the checkbox 374
Appendix 11: Order in which virtual objects were presented (Chapter 6).

Appendix 12: Visual representations of virtual objects shown to sighted participants (Chapter 6)

Addendum: Explanation of the calibration error with the IE3000 (Chapter 6).
ABSTRACT
This thesis investigates two different approaches to improving access to the Web for visually impaired people: the design of Web content; and the presentation of content. The potential for improving the design of Web content was investigated in an evaluation of the usability of the Web Content Accessibility Guidelines of the Web Accessibility Initiative. Student Web page authors used the Guidelines to adapt Web pages containing various elements (such as images and tables). These pages were collated into a Web site, which was evaluated by visually impaired people. The page authors found it difficult to find the information they required in the Guidelines document, and had difficulties with implementing the advice of the guidelines. The visually impaired people who evaluated the pages found that the extent of the accessibility of the different elements varied depending on the individual’s experience of using the Web, and the software they used. The accessibility of some elements was not improved by the implementation of the guidelines.

The potential for improving the presentation of Web content was investigated using a haptic device. The perception of virtual textures and objects by blind and sighted people via this device was examined. It was found that the virtual textures were perceived differently to the real textures examined in the literature, and that the blind people could better discriminate between the textures than the sighted people could. The virtual objects were explored from the inside and from the outside. It was found that objects generally felt larger from the inside than from the outside. This has been termed the ‘Tardis’ effect.

The thesis concludes that it is difficult to define what we mean by ‘accessibility’. Without a clear definition it is not possible to judge whether a Web site is ‘truly’ accessible. The difficulties in making Web content fully accessible mean that additional methods are required for presenting the content in different ways. The researcher believes that haptic devices offer one such method, and could be particularly useful in presenting information that is visual in nature, such as information laid out in columns.
LIST OF FIGURES

Figure 1.1: A braille display. 14
Figure 1.2: The QuickTips card produced by the WAI. 23
Figure 4.1: The header rows from example 1. 153
Figure 4.2: The first four rows of data from examples 1, 2, 4, and 5. 153
Figure 5.1: Psychophysical functions for electric shock, apparent length, and brightness. 224
Figure 5.2: Representation of aluminium plates used by Lederman (1974; 1981; Lederman & Taylor 1972) 226
Figure 5.3: Cross-section of the square waveform of plates 226
Figure 6.1: Different mental models of the location of virtual space: a) outside and b) inside. 255
Figure 6.2: The Impulse Engine 3000. 257
Figure 6.3: Cross-section of sinusoidal grooves of the virtual textures. 259
Figure 6.4: Representations of the rotated and sheared cubes. 278
Figure 6.5: Graph showing the estimation of the size of virtual cubes: inside and outside presentation. 269
Figure 6.6: Graph showing the estimation of the size of virtual spheres: inside and outside presentation. 270
Figure 6.7: Graph showing estimation of the rotation of the virtual cubes. 271
Figure 6.8: Graph showing estimation of shear of the sheared cubes by blind and sighted participants. 272
Figure 6.9: Squares rotated anti-clockwise by a) 30° and b) 70°. 278
Figure 6.10: Development of probe control. 282

In Appendices:

Figure 10.7: Four rows from example 1. 331
Figure 10.8: First three rows of example 2 with no association. 332
Figure 10.9: Row headers used in both tables in example 3, showing association between header cells. 333
Figure 10.10: First two rows of example 4, showing association between data cells and headers. 334
Figure 10.11: Four rows from example 5, showing association between header and data cells. 335

LIST OF TABLES

Table 1.1: The characteristics of the sets of guidelines discussed in Section 1.3.3.4. 31
Table 3.1: The authoring tools, editors and browsers used by the participants.
Table 3.2: The number of participants who used each combination of editor or tool and browser.
Table 3.3: The Operating Systems with which the participants worked.
Table 4.1: Key to abbreviations used for browsers and screenreaders.
Table 4.2: The browsers used by the participants.
Table 4.3: Screenreaders used.
Table 4.4: Speech synthesizers used.
Table 4.5: The braille displays used.
Table 4.6: The operating systems used.
Table 4.7: Number of examples in each set of examples.
Table 4.8: Number of participants who evaluated each set of examples, and order in which elements were presented.
Table 4.9: Number of participants who found each example of tables the most accessible.
Table 4.10: Answers given by each participant for each example and the percentage correct.
Table 4.11: Number of correct answers given for each example.
Table 4.12: Number of participants who found each example of forms to be the most accessible.
Table 4.13: Number of participants who found each example of frames to be the most accessible.
Table 4.14: Number of participants who found each example of imagemaps the most accessible.
Table 4.15: Number of participants who found each example of headers the most accessible.
Table 4.16: Number of participants who found each example of references the most accessible.
Table 4.17: Number of participants who found each example of the three figures the most accessible.
Table 4.18: Conformance of examples to Bobby's priorities.
Table 4.19: Summary of findings from study.
Table 6.1: Groove widths of the virtual textures.
Table 6.2: The type and size / angle of the simple objects.
Table 6.3: Results of analysis of data for blind and sighted participants.
1. **CHAPTER 1 — INTRODUCTION TO THESIS**

1.1 **Introduction**

1.1.1 Studies of accessibility guidelines

1.1.2 Studies of a haptic device

1.2 **Web access for blind people**

1.2.1 Access to computers

1.2.1.1 Screenreaders

1.2.1.2 Speech synthesisers

1.2.1.3 Braille displays

1.2.1.4 A brief history of access to computers

1.2.2 Access to the Web

1.2.2.1 Using mainstream browsers with screenreaders

1.2.2.2 Using special browsers

1.2.2.3 Mainstream vs. special applications and browsers

1.3 **Improving accessibility**

1.3.1 Defining accessibility

1.3.2 Approaches to improving accessibility

1.3.3 Design of content

1.3.3.1 Education

1.3.3.2 Motivation

1.3.3.3 Legislation

1.3.3.4 Guidelines

1.3.3.5 Logos and symbols

1.3.3.6 Authoring tools

1.3.3.7 Evaluation of accessibility

1.3.4 Presentation of content

1.3.4.1 Novel interfaces

1.3.4.2 Reformatting of Web pages

1.3.4.3 Future developments in access to the Web

1.3.5 Discussion on improving accessibility

1.4 **Overview of thesis chapters**

1.1 **Introduction**

"The power of the Web is in its universality. Access by everyone regardless of disability is an essential aspect."

(Tim Berners-Lee, Director of the W3C and inventor of the World Wide Web (quoted by the Web Accessibility Initiative).)

---

This thesis presents research conducted on the accessibility of the World Wide Web (WWW or Web) for blind people. In this thesis the term 'blind' is used to refer to people who are totally blind and people who have some sight, but no useful sight. In general this group use non-visual technology to access a computer (such as speech or braille output). This type of technology is described in detail later in this introductory chapter. The term 'partially sighted' is used to refer to people who have some useful sight who, in general, use magnification technology to access the computer, which is also described below. The term ‘visually impaired’ is used to refer to blind and partially sighted people. An obvious question is how many people in the UK are blind or partially sighted. A survey by the Royal National Institute for the Blind (RNIB) estimates that there were just under 1 million blind and partially sighted people in the UK over the age of 16 in 1991 (including those living in residential institutions) (Bruce, McKennell, & Walker, 1991, p43).

In this thesis two series of studies are presented. On the surface these studies appear to be quite different, although in fact they illustrate the different approaches that can be taken to improve the accessibility of the Web. The first series of studies involved the evaluation of the Web Content Accessibility Guidelines for developing accessible Web sites. The second series of studies involved a haptic device. An overview of these two series of studies is presented below.

1.1.1 Studies of accessibility guidelines
The Web Content Accessibility Guidelines\(^2\) (WCA Guidelines) were developed as an online resource by the Web Accessibility Initiative\(^3\) (WAI) which is a programme of the World Wide Web Consortium\(^4\) (W3C). The Guidelines document is accompanied by a Checklist, and a Techniques document containing techniques for implementing the Guidelines. The first study focused on the use of the Guidelines by Web page authors to develop Web pages. The second study evaluated the accessibility of the pages produced by the page authors for visually impaired people.


\(^3\) Web Accessibility Initiative home page: http://www.w3.org/WAI (checked Nov 2000).

This series of two studies was the first of the use of the WCA Guidelines and the first known study of online guidelines or accessibility guidelines.

Guidelines for designing usable user interfaces are regarded as useful tools for transferring knowledge from human factors specialists to interface designers (Mosier & Smith, 1986; de Souza, Long, & Bevan, 1990; Tetzlaff & Schwartz, 1991). However, previous studies of the use of guidelines have found that this type of tool has limitations. Studies have found that designers had difficulty finding the relevant guidelines and applying guidelines to the current design (Mosier & Smith, 1986), had difficulty understanding the advice given, especially when this concerned new concepts (Tetzlaff & Schwartz, 1991), and had various other difficulties and made errors (de Souza et al., 1990). In a study of the use of a standard, designers were found to be more influenced by real systems than by the standard (Thovtrup & Nielsen, 1991). In two of these studies the importance of the role of examples was emphasised. Designers were found to follow examples to the exclusion of the accompanying guidelines text (Tetzlaff & Schwartz, 1991) and examples were found to have as much impact on designers as a standard (Thovtrup & Nielsen, 1991).

None of the studies cited above involved user testing of designs produced using a set of guidelines or a standard, although some of the designs were evaluated in terms of their conformance to the guidelines or standard using some form of checklist. Therefore this series of studies is the first known evaluation of a design produced using guidelines that involves users.

1.1.2 Studies of a haptic device

The second series of studies presented in this thesis involved a novel, 'haptic' device with which users can 'feel' virtual objects and textures. It has been suggested that devices of this type could provide blind people with additional support in using the Web (Hardwick, Furner, & Rush, 1998). Such devices could provide haptic or tactile information to the supplement synthesised speech output or braille output that would normally be used by a blind Web user.

---

5 Haptic perception is the combination of tactile and kinaesthetic perception.
Before such haptic devices can be used as interfaces to the Web it is important to investigate the abilities of devices to display objects and textures in an appropriate way, and to investigate the abilities of blind (and sighted) people in interacting with the device and the artefacts it displays. The researcher therefore conducted a series of psychophysical studies in order to investigate the capabilities of the device and the abilities of potential users of the device.

This was the first known study of the perception of virtual textures by blind people and of the perception of the size and angle of virtual objects by blind or sighted people. Previous research had been conducted on the perception of real textures by sighted people (Stevens & Harris, 1962; Lederman, 1974; 1981; 1982; Lederman & Abbott, 1981; Lederman & Taylor, 1972) and on the perception of virtual textures by sighted people (Minsky, 1995).

It was clear that this particular device was not sophisticated enough to provide information of sufficient complexity to support Web navigation. A great deal of work would be required on the hardware and the software before a feasible application could be developed. However in the near future it is likely that force-feedback will be of great benefit for blind people in the form of force-feedback mice and similar devices. However, devices like the one used in these studies are likely to remain in the domain of research.

1.2 Web access for blind people

In order to discuss access to the Web for blind people it is necessary to describe how this group use computers. Once a blind person has access to a computer, for which they require a combination of hardware and software, they can access the Web using either a mainstream browser or one designed especially for blind people.

1.2.1 Access to computers

Blind people access computers using a combination of hardware and software that present the visual contents of the screen in another form, mainly in synthesised speech, but also in braille. A piece of software known as a screenreader directs the information that is sent to the screen to either a speech synthesiser or a braille
display. To input commands to the computer blind people do not use a mouse, but enter commands via the keyboard (they learn to touch-type as sighted people do). The three types of hardware and software used to access computer output are described below.

1.2.1.1 Screenreaders

A screenreader captures the information sent from the computer to the screen. It passes this information to either a speech synthesiser or a braille display. There are currently around eight screenreaders available. Many developers produce separate screenreaders for use with text-based applications and applications with graphical user interfaces (GUIs) such as those used under the Windows operating system (OS). Most of the screenreaders available for GUIs only run under Windows, only one is available for the Mac OS. Screenreaders are also available for applications running under the Unix and OS/2 operating systems. The screenreaders for text-based applications mainly run under the MS-DOS operating system. All screenreaders support speech synthesisers, but not all support braille displays.

Bosher\(^6\) provides an overview of the major screenreaders available, including their developers, the platform they run under, and whether they support speech and braille output\(^7\).

- ASAW from Microtalk, runs under DOS and Windows 95/98, supports speech.
- Hal 95 from Dolphin, runs under DOS, Windows 95/98 and NT, supports speech and braille.
- Jaws For Windows from Henter-Joyce, runs under DOS, Windows 95/98 and NT, supports speech and braille.
- OutSpoken from Alva, runs under Windows 95/98 and Macintosh, supports speech and braille.
- Simply Talker from Econonet, runs under Windows 95/98, supports speech.
- Slimware Window Bridge from Synthavoice, runs under DOS, Windows 3.x and 95/98, supports speech and braille.
- Window-Eyes from GWMicro, runs under DOS, Windows 3.x and 95/98, supports speech.
- WinVision from Artic, runs under Windows 3.x and 95/98, supports speech.

---

\(^6\) In a WAI Web page on assistive technology available at: http://www.w3.org/WAI/References/Browsing (checked Nov 2000).

\(^7\) The URLs for these screenreaders are provided after the Reference list.
As well as presenting the information on the screen, screenreaders also provide additional functionality to allow the user to interact with the information. For example screenreaders have commands for pausing or repeating the speech, and echoing back keyboard input. Screenreaders often have two modes, one for reviewing information and another for interacting with the information.

When a screenreader is used with a speech synthesiser the user can control the amount of information that is read, for example a paragraph, a line, a word, or a character at a time. The user can also move the focus backwards and forwards in the text, control the rate and pitch of the speech, and may be able to choose between different voices.

### 1.2.1.2 Speech synthesisers

A speech synthesiser can be either a piece of hardware or software. The former is usually a small box that sits on the desktop and has its own speaker. The latter is usually a card that fits inside the computer and uses external speakers or those of the computer. Users working in shared offices may need to wear headphones. The main advantage of using a speech synthesiser over a braille display is that they are cheap, and often come as standard with new computers. The main disadvantages are that synthesised speech can be very tiring to listen to over a long period of time, more so than human speech, and no information is available about the spatial layout of the screen.

### 1.2.1.3 Braille displays

The braille displays used with desktop computers are similar in size and shape to a standard keyboard (see Figure 1.1, below). Those used with portable computers are approximately the same size as such computers. A braille display sits underneath the keyboard so that the user has access to both. A braille display has a number of braille cells (usually 20, 40 or 80) in a row. These cells are groups of eight plastic pins arranged in a 2 by 4 matrix. The cells are 'refreshable' which means that the pins can move up and down to create braille characters. The row of cells displays
one line, or part of a line, of the text on the screen. As the user moves the cursor up and down the screen, the braille display presents the rows of text.

![Figure 1.1: A braille display](http://www.humanware.com)

The use of a braille display requires the user to be able to read braille. Contrary to a common belief, only a small percentage of blind people can read braille proficiently. The majority of blind people lose their sight later in life, either accidentally, or progressively, and therefore do not learn to read braille as a child. Most blind people who have lost their sight later in life do learn to read braille, but do not usually reach a stage at which they can read it efficiently enough to use braille display with a computer (Bruce et al., 1991).

There are several advantages of using a braille display. The user can read as fast as they like and is not restricted by the rate of the synthesised speech. Also the user gets a better idea of the layout of items on the screen. The main disadvantage is that braille displays are very expensive (for example in 1999 prices in the UK started at around £2000). They are therefore mainly used in the workplace.

---

8 Picture from HumanWare Web site (http://www.humanware.com) (checked Nov 2000).
9 A survey conducted by the RNIB published in 1991 found that only 3% of visually impaired people are sufficiently fluent in braille to read a book or magazine (Bruce et al., 1991).
1.2.1.4 A brief history of access to computers

The applications that ran on early personal computers had text-based interfaces and were accessible to blind people because the early screenreaders read the screen line by line. More recently the advent of the graphical user interface (GUI) meant that the existing screenreaders could not present the screen in an effective way. Those screenreaders did not recognise the icons and buttons that made up the interface and reading the screen line by line was not necessarily appropriate (Petrie & Gill, 1993).

For some years blind people were left behind in terms of using applications with GUIs. This was a particular problem for those blind people whose job involved the use of computers because as their colleagues changed over to using applications with GUIs, blind people were not able to make this shift. In some cases this was a threat to maintaining employment. Eventually the developers of applications and screenreaders began to improve the situation. The developers of applications gave the icons, buttons, menus and other interaction objects labels that were not displayed on the screen but were available to screenreaders. At the same time the developers of screenreaders brought out new products that could read these labels.

A vital aspect of this transition from text-based to GUI applications was that blind people needed to learn about the use of GUIs, which are, of course, inherently visual in nature and intended to be used with a mouse. Morley has written two books (1995; 1997) that describe the visual layout of the screen, and explain the purpose of the various interaction objects and how to interact with such objects via the keyboard.

The current situation is that blind people who need or want to use applications with GUIs can do so. Others still use older screenreaders with text-based interfaces. Blind people are currently using software with both types of interface to access the Web.

1.2.2 Access to the Web

Blind people use a variety of different browsers to access the Web. Some use a mainstream browser with a screenreader; others use specially designed browsers,
many use both at different times for different tasks. Examples of GUI and text-based mainstream browsers, and browsers specially designed for blind are described below.

Bosher\(^{10}\) provides an overview of the mainstream browsers often used by people with disabilities\(^{11}\).

- Lynx is a popular text-based browser for UNIX, Windows 95/NT and MS-DOS. It allows flexible and powerful text-based access from older platforms.
- Internet Explorer 4.01 runs under Windows. It includes many features to enhance accessibility.
- Net-Tamer runs under MS-DOS and includes both text-based and graphical browsing capabilities.
- Netscape Navigator enables enlargement of fonts.
- Opera is a compact browser for Windows 95/98 offers enhanced keyboard navigation and screen magnification.

The main difference between using a special browser and a mainstream browser with a screenreader is that special browsers announce the different HTML elements, whereas most screenreaders cannot interpret the content of the Web page and just present it as text, links and images.

1.2.2.1 Using mainstream browsers with screenreaders

Many blind people who want or need to use a mainstream browser with a GUI use Internet Explorer. This requires a Windows screenreader. They use Internet Explorer for two main reasons. Firstly, its developers, Microsoft Corp., have improved its ability to work with screenreaders to present the content of Web pages in a useful way. Secondly, the developers of some screenreaders, in particular Jaws, have adapted their product to improve the way in which they work with Internet Explorer. Eight of the twenty participants in the second study presented in this thesis (see Chapter 4) used Internet Explorer with Jaws, and this was the most common combination of browser and screenreader. Other blind people use Netscape with a screenreader.

The more recent versions of such browsers support all the functionality available on the Web: Java, scripts, frames, forms, tables, images etc. They have many settings

\(^{10}\) In a WAI Web page at: \text{http://www.w3.org/WAI/References/Browsing (checked Nov 2000)}.
\(^{11}\) The URLs for these browsers are provided after the Reference list (Section 9).
with which the user can choose what elements are displayed and how. For example, the user can choose whether or not they wish images to be displayed. Any alternative text associated with the image is displayed if the image is not.

Experienced blind Web users who use GUI browsers with a screenreader can navigate most Web pages. The main difficulties for all users are caused by the screenreader reading the page line by line, and advanced technologies such as scripts or Java applications. Most screenreaders require the user to find the best way of reading a page that contains frames or tables because the software cannot detect these elements. The developers of some screenreaders are finding ways of overcoming this problem. For example, the Jaws screenreader has a function that can remove columns that have been used to layout text and images, for example in a table, and present the information in one column. Until all screenreaders are able to present more complex Web pages in a useful manner, most users will have to continue to rely on their experience of trial and error to successfully navigate such pages.

Blind people who want or need to use a mainstream text-based browser mostly use Lynx. Lynx does not display tables or frames, but does support forms. It does not display images but will display any alternative text associated with them. This browser runs under the Unix OS but can be run remotely so that users of other operating systems can use it. A Web page is presented as text which a screenreader reads line by line. This is a very effective method of accessing the basic functionality of the Web, but again, not for more advanced features, such as scripts or Java applications.

1.2.2.2 Using special browsers

There are currently around ten ‘special’ browsers designed for use by blind people, and others are under development. These browsers produce synthesised speech themselves via a speech synthesiser and therefore do not require a screenreader. Whereas a screenreader would read the contents of the screen line by line, many special browsers attempt to interpret the underlying HTML and present the different elements to the user.
The first special browser developed was WebSpeak. As shall be seen later in this thesis, WebSpeak can present the various elements found in Web pages, such as tables, forms and frames, in a very useful way. For example, WebSpeak announces the elements within a page, such as the title, the headings and the links. It may say, "This page is entitled Chetz Colwell's home page. Heading: Welcome" and then read the following paragraphs of text. When WebSpeak encounters an image, it announces "image" and when it encounters a link it announces it "link". The user can choose whether to read the elements in the order they appear, or they can read through the links, through the headings, or through the paragraphs. The user can easily switch between these modes as required. Bosher\textsuperscript{12} provides an overview of browsers designed specially for people with disabilities\textsuperscript{13}.

- Braillesurf is a braille and speech browser under development by the Braillenet project, currently only available in French. Features include speech output and braille support.
- Emacspeak is a speech-enabled environment for EMACS, which runs under UNIX or LINUX. Features include full web browsing capabilities through W3, speech output, and a simple keyboard interface.
- HomePage Reader is the new speech-based browser from IBM, which uses Netscape as its engine. Features include speech output, and a simple keyboard interface based on number-pad.
- Marco Polo is a plug-in for Netscape Navigator from Sonicon. Features include speech output, auditory icons and a simple keyboard interface.
- MultiWeb is a disability-specific browser developed at Deakin University. Features include speech output, screen magnification, and scanning for switch devices.
- WebSpeak was the first low-vision browser and is developed by the Productivity Works. Features include speech output, synchronised speech and screen magnification, and a simple keyboard interface.
- Sensus Internet Browser is a low-vision Internet browser from Sensus in Denmark. Features include speech output, braille support, and special screen fonts.
- Simply Web is a talking interface using the Internet Explorer engine. Features include speech output, and a simple keyboard interface.
- VIP is a low-vision browser from JBliss Inc. Features include screen magnification and speech output.

1.2.2.3 Mainstream vs. special applications and browsers

There is debate about whether it is better for blind people to use mainstream applications or special applications, and there are advantages and disadvantages to

\textsuperscript{12} In a WAI Web page at: http://www.w3.org/WAI/References/Browsing (checked Nov 2000).
\textsuperscript{13} The URLs for these browsers are provided after the Reference list (Section 9).
both. This debate naturally extends to include Web browsers. In terms of the Web, the main advantage for a blind person using a mainstream browser is that they are using the same browser as their sighted colleagues, and thus those providing technical support are familiar with the software. Also they do not have to wait for a special browser to catch up with the latest Web technology. However, a screenreader working with a mainstream browser may not be able to present the contents of a Web page in a useful way. For example, most screenreaders do not recognise particular HTML elements, especially those that force the visual layout, such as tables and frames. A screenreader may read the screen line by line even when this is not appropriate, which can be very confusing. Also, screenreaders may not immediately catch up with the latest Web technology. However, some screenreader developers are working with the browser developers so that the screenreaders can improve the way in which they present certain HTML elements, particularly those that change the visual layout of the page, such as tables and frames.

The main advantage of using a special browser is that the basic HTML elements are presented in a useful way, which can help people learn about the Web. However, the disadvantages are that technical support personnel are unlikely to be familiar with the software, and there is often a delay before such browsers support new Web technologies.

A common conclusion to this debate is that special browsers are excellent for blind people who are new to the Web because they present the content so clearly, but that for those with more experience who want to access the full functionality of the Web, mainstream browsers are better. Many blind people have more than one browser installed on their computer and use a different browser for different tasks. If a Web page is not easy to use with one browser they will try it using another browser. Four of the 22 blind participants in the second study reported in this thesis (see Chapter 4) used more than one browser.
1.3 Improving accessibility

1.3.1 Defining accessibility

The problem of access to the Web for visually impaired people has many levels. The design of pages can cause accessibility problems, but the design of browsers and screenreaders also contributes to the problems in the way they present information to the user. A recent discussion conducted via the email discussion list of the British Computer Association of the Blind provides a good example of the different levels of accessibility problems. One person had difficulty accessing a site that is of particular interest to visually impaired people because it contained frames. Another correspondent found the site was completely accessible to them and stated that this was because they were using a 'better' screenreader. Several other people then tried the site with different screenreaders. The result was that the site seemed to be easier to navigate with some screenreaders than with others.

Some correspondents said that as long as the site was accessible to some screenreaders it was acceptable, others said that the site should cater for those who do not have the top-of-the-range screenreaders. The discussion also touched on whether the site was accessible with different browsers. The main issue was whether the site was accessible with text-based browsers or only with GUI browsers. Text-based browsers do not display frames, but do support no-frames versions if the page author provides these. This discussion provides a good example of the complexity of accessibility and how it occurs at different levels.

Another interesting related question is whether it is the actual HTML element that is inaccessible, or the lack of alternatives that makes an element inaccessible, or whether it is the way that an element is presented to the user. For example, an image in a Web page is not accessible to a visually impaired person per se, but the information provided by the image can be made accessible by the provision of alternative text. This is what is usually meant by making an image accessible. If an image has alternative text, that text may be announced automatically by some screenreaders or special browsers, but may not be announced by other screenreaders. Therefore the image is accessible to some users, but not to others.
The situation is different with elements that present information in a certain way. For example, most screenreaders and special browsers will read the data within the cells of a table. However, some screenreaders will read the cells one by one, row by row, but without differentiating between the different cells and different rows. In contrast, other screenreaders will pause between cells and between rows so that the user can differentiate between the cells and rows. This discussion highlights the complex interaction between the source (the HTML) the screenreader (if used), and the browser.

1.3.2 Approaches to improving accessibility

The distinction made above between the accessibility of Web content and the way in which it is presented to the user is also found in the different approaches to improving accessibility. These approaches can roughly be divided into two categories. The first category includes techniques for improving the accessibility of the content of the Web, such as the use of guidelines for creating accessible content. The second category includes different approaches to the presentation of the content of the Web to the user, for example the development of special browsers. This chapter discusses the issues in these two main categories of the design of content and the presentation of content.

The activities of the Web Accessibility Initiative (WAI) of the W3C fall into both of these categories. The WAI has five working groups, which are listed below with their current focus of activity\(^4\). The activities of the WAI will be referred to at various points in this chapter.

- Web Content Working Group: development of guidelines for the design of accessible Web content.
- User Agent Working Group: development of guidelines for the design of accessible user agents (including browsers).
- Authoring Tool Working Group: development of guidelines for the design of authoring tools which support accessibility features (and are themselves accessible).

\(^4\) For details of each group see http://www.w3.org/WAI/ (checked Nov 2000)
• Education and Outreach Working Group: increasing awareness of accessibility issues of designers, developers, managers, governments etc.

1.3.3 Design of content

There are several different approaches to improving the accessibility of the content of the Web. These include increasing awareness of accessibility issues, providing education about specific techniques, improving authoring tools, and providing techniques for evaluating accessibility. These different approaches are discussed in the following sections.

1.3.3.1 Education

The WAI is committed to increasing awareness of accessibility issues and providing education on improving accessibility via its Education and Outreach Working Group. This group is involved in a range of activities. For example, the development and distribution of a ‘QuickTips’ card which summarises the techniques for making a page accessible (as shown in Figure 1.2, below). The group also gives presentations at conferences, and provides training.

The resources available from the AWARE centre of the HTML Writers Guild (HWG) complement those of the WAI. It provides on-line courses on Web accessibility, maintains a database of resources on Web development, and raises awareness with a ‘Web accessibility month’ every year. The HWG represents page authors in the W3C.

16 Information on QuickTips card available at: http://www.w3.org/WAI/References/QuickTips/ (checked Nov 2000).
1.3.3.2 Motivation

An important aspect of educating people, such as page authors, managers, etc. about accessibility is to provide them with some motivation for improving the accessibility of their page or site. This can be difficult when authors or managers see this task as one that requires many person-hours to make changes for a small proportion of their potential users/customers. However, there are a number of points that can be used to support the argument for making a site more accessible.

- Some page authors may believe that a site that contains many images, tables or frames can only be accessible if a text-only version is created. This approach is obsolete as there are now techniques available to allow access to the above elements.
- Some page authors may believe that they do not have any customers who have disabilities and it therefore not worth the cost of making a site accessible for users of special browsers. However, it could be argued that some of their existing customers may not be using standard browsers or devices and such organisations may not want to exclude any potential customers from their site. For example,
- Mobile devices may not be able to display images, scripts, or frames.
• Non-disabled users may be using the Web via the telephone and are therefore using voice input, and do not have a visual display.
• Users may be accessing the Web over a slow modem connection and have to wait for images to load.
• Some organisations who already make information available in different formats, such as on tape or in braille, may prefer to make information available to visually impaired people via the Web, rather than to prepare audio or braille versions.
• Authors of ‘shopping’ sites may believe that they do not have disabled customers. However, some people with disabilities may prefer to shop via the Web if they have mobility difficulties.

1.3.3.3 Legislation

In October 1999 Part III of the UK Disability Discrimination Act (DDA) came into force. “This means that service providers – such as banks, social services departments, places of worship, leisure centres – will have to make ‘reasonable adjustments’ to ensure that disabled people can access their services” (Casserley, 1999, p8).

A Code of Practice has been published, which provides details of what the Act covers and what may be expected from it. Service providers now have certain obligations to take reasonable steps where it is impossible or unreasonably difficult for people with disabilities to access their services.

• To provide a reasonable alternative means of service where there are physical barriers to services
• To provide an auxiliary aid or service, such as information on tape.

While it is relatively easy to define ‘impossible’, it is less easy to define ‘unreasonably difficult’. Although the Act does not provide a definition, the Code of Practice says that,

“When considering if services are unreasonably difficult for disabled people to use, service providers should take account of whether the time, inconvenience, effort or discomfort entailed in using the service would considered unreasonable by other people if they had to endure similar difficulties” (Casserley, 1999, p9).

The Act does not define what are ‘reasonable’ steps, but the Code of Practice provides some information. “What is reasonable will vary according to the type of service being provided; the nature of the service provider and its size and resources; the effect of the disability on the individual disabled person” (Casserley, 1999, p9).
Certain factors may be taken into account when considering what are 'reasonable' steps:

• Whether the steps would be effective in overcoming the difficulty;
• The extent to which it would be practicable for the steps to be taken;
• The extent of any disruption caused by taking the steps;
• The financial and other resources of the service provider;
• The amount of resources already spent on making adjustments;
• Whether financial or other assistance is available.

The Act has provided certain 'justifications' that can be used to justify not making 'reasonable adjustments'.

• If adjustments would endanger health and safety;
• If the disabled person is not able to enter into an enforceable agreement (due to, for example, learning difficulties);
• If the adjustment would fundamentally alter the nature of the service provided.

The DDA represents the first opportunity for people with disabilities in the UK to use the law to change the way in which services are provided, and to make service providers take positive steps to improve services. Casserley believes that “for all its shortcomings, the DDA is still a useful tool in gaining equality of treatment for blind and partially sighted people” (Casserley, 1999, p10).

The fact that the DDA came into effect so recently means that it is not yet known to what extent it will apply to the Web and the Internet. It is hoped that there will soon be test cases that will set precedents for UK-based Web sites to be required to be accessible. While the DDA has only recently come into effect, the US has had such legislation since 1990. There have been some cases that have been successful in stating that US-based Web sites are required to be accessible.

The Americans with Disabilities Act18 (ADA) came into force in 1990. The Act, “Prohibits discrimination on the basis of disability in employment, programs and services provided by state and local governments, goods and services provided by private companies, and in commercial facilities” (quoted in Waddell & Thomason, 1998, p1). This means that 'entities' covered by the Act are required to

---

"Furnish appropriate auxiliary aids and services where necessary to ensure effective communication with individuals with disabilities, unless doing so would result in a fundamental alteration to the program or service or in an undue burden. ...Auxiliary aids include taped texts, Brailled materials, large print materials, captioning and other methods of making audio and visual media available to people with disabilities" (Waddell, 1998) (emphasis added).

It is the concept of 'effective communication' that has been found to relate to the accessibility of Web sites. The US Department of Justice is quoted by Waddell & Thomason (1998) as saying that entities covered by the Act "that use the Internet for communications regarding their programs, goods, or services must be prepared to offer those communications through accessible means as well" (Waddell & Thomason, 1998, p2).

In defining 'effective communication' the US Office of Civil Rights is quoted by Waddell & Thomason (1998) as stating that

"The issue is not whether the [person] with the disability is merely provided access, but the issue is rather the extent to which the communication is actually as effective as that provided to others" (Waddell & Thomason, 1998, p2).

An interesting issue on the topic of legislation that is not addressed by the available literature is the definition of accessibility. For example, if a Web page is accessible to a user using the latest versions of a browser and screenreader, but is not accessible to a user of an older text-based browser, can the page be judged to be accessible? A reasonable suggestion would be to use the conformance measures of the WCA Guidelines to assess the accessibility of a Web page. However, the highest rating of conformance requires the use of certain HTML tags that are not yet supported by most mainstream or special browsers. Therefore, if conformance to the Guidelines was a measure of accessibility, Web sites that conform may not currently be accessible to some or all users. This is an important issue to be resolved if UK-based Web sites are to be required to be accessible by law.

1.3.3.4 Guidelines

Guidelines are a popular method for informing page authors about accessibility. There are several sets of guidelines available on the Web, and at least one set has
been published in a journal. These sets of guidelines differ on many levels. Some are very long while others are just one page long. Some focus on accessibility for people with one specific disability while others cover all groups. Some include examples of HTML code and some include guidelines for usability as well as accessibility. Some are produced by organisations, while others are produced by individuals. A broad selection of these sets of guidelines is discussed below and the salient features are summarised in Table 1.1, below.

1.3.3.4.1 Web Content Guidelines

One of the key strategies of the WAI for educating page authors about accessibility is the development of the Web Content Accessibility Guidelines. This document is a Recommendation of the World Wide Web Consortium (W3C). The Guidelines document presents a set of guidelines for improving accessibility. The Guidelines document is accompanied by two other related documents. The first is the Techniques for Web Content Accessibility Guidelines 1.0\textsuperscript{19} which contains techniques for the implementation of the checkpoints in the Guidelines document. While the Guidelines document is a stable document, the Techniques document is expected to evolve as Web technologies develop.

The second related document is the Checklist of Checkpoints for Web Content Accessibility Guidelines 1.0\textsuperscript{20}. This is formally an appendix of the Guidelines document. It presents all the checkpoints grouped by priority (see below). The purpose of the Checklist is to review the accessibility of a page or site. It has columns where the page author could indicate on a printout whether each checkpoint has been satisfied, has not been satisfied, or is not applicable.

The Guidelines, Techniques and Checklist documents were developed by the WAI Web Content Working Group. This group had public discussions via email, teleconferences, and face to face meetings.

\textsuperscript{19} Techniques for Web Content Accessibility Guidelines 1.0 available at: http://www.w3.org/TR/WAI-WEBCONTENT-TECHS/ (checked Nov 2000).

Each guideline in the Guidelines document has an explanation of its purpose, and states which groups of users would benefit from its implementation. Each guideline has a list of checkpoints associated with it. Within the list of checkpoints associated with each guideline there may be links to techniques in the Techniques document, and links to other related checkpoints listed under other guidelines.

Each checkpoint has an associated priority level. Priority 1 applies to checkpoints that must be satisfied; otherwise one or more groups will find access 'impossible'. Priority 2 applies to checkpoints that should be satisfied; otherwise one or more groups might find access 'difficult'. Priority 3 applies to checkpoints that may be satisfied; otherwise one or more groups will find access 'somewhat difficult'.

The WCA Guidelines document describes three levels of conformance that are associated with the priorities. To achieve level A conformance all priority 1 checkpoints must be satisfied. To achieve level 'Double A' conformance all checkpoints of priority 1 and 2 must be satisfied. To achieve level 'Triple A' conformance all checkpoints of priority 1, 2 and 3 must be satisfied.

The document defines how page authors can indicate which level of conformance they have satisfied. Three logos are available, one for each level of conformance. To use these a page author must state the title of the Guidelines, the URL of the Guidelines, which level of conformance has been satisfied, and the scope of the claim (a page, a whole site, or a definite portion of a site).

1.3.3.4.2 *Hints for accessible Web sites*

The Royal National Institute for the Blind (RNIB) has developed 'Hints for designing accessible Websites'. This document is divided into sections that cover the main aspects of pages that can cause accessibility problems. Each section contains approximately five guidelines. The guidelines provide advice on making pages accessible to partially sighted people as well as blind people. They are brief; each guideline consists of one or two sentences. The document does not contain any

---

examples and does not provide any information about how to implement the advice. There is no indication of the priority of any of the guidelines and no 'RNIB approved' logo for sites to display.

1.3.3.4.3 ‘AUS’ Standards

The Attorney General’s department of New South Wales produced the ‘Attorney General’s Universal Specifications (AUS) Standards’. This document is divided into sections that cover the common aspects of pages, such as tables, frames and forms. The document provides a rationale for each guideline and examples of HTML code for the implementation of the advice. The document also contains information on the validation of HTML code and the use of META tags to improve the way in which search engines display information about the page or site. The document does not state whether any of the guidelines have higher priority over others. Web sites developed by the Australian government that comply with the Standards can display the AUS logo. It is assumed that there is no assessment process for sites that claim complicity and no system of ‘policing’ such claims.

1.3.3.4.4 Harmony Project

The Harmony project was funded by the European Union. The project developed the HARMONY Guidelines for the Creation of Accessible WWW Documents. These guidelines are divided into sections that cover the main causes of accessibility problems. The document does not include any examples of HTML code to assist with the implementation of the advice, nor does it indicate if any of the guidelines have a higher priority over others.

1.3.3.4.5 Gallagher’s accessibility test

Guidelines have also been published in journals, for example Gallagher’s article (1998) features list of seven guidelines in a ‘quick accessibility test’. The test is a


combination of questions about aspects of a Web page with advice for improving accessibility.

1.3.3.4.6 O'Brien's guidelines

Stella O'Brien produced her own guidelines, 'Ten hazards in accessible web design'\(^{24}\). The document has a personal quality to it in that it is written in the first person. For example, in one guideline the author states,

"I don't need to know that the image I have loaded is 'a rainbow coloured horizontal line'. I am interested to learn that it represents 'Section 2' or 'Part 3'.'" (O'Brien, 1998, p1)

The document does not contain any examples or any indication of priorities.

1.3.3.4.7 Comparison of existing guidelines

Table 1.1, below, provides an overview of the characteristics of the different sets of guidelines discussed above. The use of guidelines is discussed in more detail in Chapter 3 on the evaluation of the WCA Guidelines. Here, it can be concluded that there is a range of sets of guidelines available that are different on a range of levels. One thing they all have in common is that they all appear to lack any evaluation of their usability or their validity. It may not be appropriate for guidelines produced by individuals to be subject to rigorous testing, but it is certainly appropriate for guidelines produced by organisations who may have a wider audience that may expect the guidelines to be valid.

1.3.3.5 Logos and symbols

An example of a 'kitemark' approach is the 'Web access symbol'\(^{25}\). This symbol was developed in a competition run by the National Center for Accessible Media of the US. The symbol can be used in Web sites if they meet certain accessible design criteria. Sites that display the symbol are intended to be used as exemplar sites for others to view and use to follow the example. It is acknowledged that it will not be possible to guarantee that all the sites that use the symbol will be totally accessible.


It is intended that visually impaired Web users will let the designers know if they encounter problems and offer suggestions for improving accessibility.

Table 1.1: The characteristics of the sets of guidelines discussed in Section 1.3.3.4.

<table>
<thead>
<tr>
<th>Guideline sets</th>
<th>WAI</th>
<th>RNIB</th>
<th>Harmony</th>
<th>AUS</th>
<th>Gallagher</th>
<th>O’Brien</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disabilities covered</td>
<td>All</td>
<td>Visual impairment</td>
<td>Visual impairment</td>
<td>All</td>
<td>Visual impairment</td>
<td>All</td>
</tr>
<tr>
<td>Priorities explained?</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Example code provided?</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Logo to use?</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Approx. no. of pages</td>
<td>117*</td>
<td>2</td>
<td>22</td>
<td>11</td>
<td>&lt;1</td>
<td>3</td>
</tr>
<tr>
<td>Background information provided?</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Rationale / Explanation provided?</td>
<td>Y</td>
<td>Some</td>
<td>Y</td>
<td>Some</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>

Key: Number marked with * includes the Guidelines document, the associated techniques document and the Checklist.

This use of exemplar sites is likely to increase the awareness of Web page designers of accessibility issues. However, a problem with it is that it may be difficult for designers to follow the principles that underlie the design of a site, just from looking at it or reading the source code. This is especially true if the accessibility of the exemplar site has not been assessed.
The 'Speech Friendly Ribbon award'\textsuperscript{26} can be displayed on Web sites that have been judged to be accessible to blind and visually impaired people, that is accessible with all browsers and screenreading software. Five criteria have to be met by a Web site in order for the Ribbon to be awarded.

- Links are placed on separate lines;
- Images have associated alternative text;
- Sites that are 'extremely graphical in nature' have a text-only alternative site;
- Forms are 'formatted for ease of navigation' and offer the option of sending the information by email;
- Frames use the no-frames option.

The developers of the Ribbon do not explain why links should be placed on separate lines, or how to format a form for easy navigation. Instead it provides a link to another page containing a list of Web sites relating to accessibility. If a Web page author wishes to display the Ribbon on their site and believe they have met the criteria they can submit their URL via email to the Ribbon owners. The process used to assess a site is not described. Unfortunately, the Ribbon site does not seem to adhere to its own criteria: it has 'mailto' links embedded in paragraphs of text.

1.3.3.6 Authoring tools

There are a number of authoring tools for developing Web pages currently available, such as DreamWeaver\textsuperscript{27}, FrontPage\textsuperscript{28} and HotMetal Pro\textsuperscript{29}. In addition many word processing applications enable page authors to save files in HTML format. Only one authoring tool is currently known to include any accessibility features, HotMetal Pro.

The Authoring Tool Working Group\textsuperscript{30} of the WAI had developed guidelines for the developers of authoring tools who wish to enable page authors to develop accessible pages. This work is essential as it has potential to improve the accessibility of the

\textsuperscript{26}The Web page for the Speech Friendly Ribbon award is no longer available. However, examples of Web pages that have been awarded the Ribbon include Blind links available at: http://www.seidata.com/~marriage/ribbon.html; Disability Resource Directory available at: http://www.kansas.net/~cbaslock/awards.html; Electric edge available at: http://www.ragged-edge-mag.com/archive/index1.htm (all checked Nov 2000).

\textsuperscript{27}Details of DreamWeaver available at: http://www.macromedia.com/software/dreamweaver/ (checked Nov 2000).


\textsuperscript{29}Details of HotMetalPro available at: http://www.hotmetalpro.com/ (checked Nov 2000).

many Web pages that are created using dedicated tools or word processors. The Authoring Tool Accessibility Guidelines state that most of the content of the Web is created using authoring tools\textsuperscript{31}. It is therefore crucial that such tools help authors produce accessible content.

1.3.3.7 Evaluation of accessibility

1.3.3.7.1 Automatic checker - Bobby

The concept of an automatic accessibility checker is an extension of that of an HTML checker. These automatic accessibility checkers are online services for which the user enters the URL of the page to be checked and the checker returns a report. An advantage of an automatic accessibility checker is that the user is not required to have any knowledge of accessibility. A disadvantage of such a checker is that it can only make decisions using its built-in rules and cannot make 'intelligent' decisions.

The Bobby accessibility checker\textsuperscript{32} was developed by the Center for Applied Special Technology (CAST). CAST has worked closely with the WAI and the latest version of Bobby incorporates the principles of the WCA Guidelines. The report returned by Bobby contains feedback on the accessibility of the page and its compatibility with a range of browsers. The report contains explanations of any problems found and suggestions for solving the problems.

1.3.3.7.2 Self-administered tests

The Public Service Commission (PSC) of Canada developed the 'Webpage Accessibility Self-Evaluation Test'\textsuperscript{33} (WPASET). The test consists of 20 questions for evaluating the accessibility of a page for a visually impaired user. Each question has a brief description of why a particular feature may be inaccessible. The user of the test scores points for their answer to each question. The test addresses many of the features that can cause problems for visually impaired users.


In some cases the technique for solving an accessibility problem is implied within the question, for example, 'use the ALT text part of the HTML image tag to provide description'. However, in other cases the user of the test may not be able to identify a solution to a problem they have found on their page, for example. The test does not take a prohibitive approach: it does not say that designers should not use particular HTML features, but says that designers should consider how users would access the information.

The WPASET test addresses some usability issues that affect sighted users as well as visually impaired users. For example, it emphasises the importance of consistency of document layout and other design strategies. It also briefly mentions how navigational links help 'everyone' find their way through Web documents.

Petrie, Colwell, Evenepoel & Engelen (1998) performed an evaluation of the WPASET test. Twenty-one students used the test to evaluate the accessibility of three Web sites. First the students decided whether each question in the test was relevant to each site, and then allocated a score for each relevant question. It was found that the students disagreed about when particular questions would be applicable, and about the scores allocated to different sites. A number of inconsistencies were also found with the scoring systems of the test. This evaluation was conducted on version 1 of the WPASET, which is no longer available. The current version is number 2.

1.3.3.7.3 Involvement of users

While the approaches discussed above tend to only involve those who have a professional interest in accessibility, there is one approach that directly involves users with disabilities. Users often report accessibility problems they experience to Webmasters, but in a very informal way. The Webwatch email list is a forum in which users discuss the difficulties they have with particular sites and report on their

34 The discussion of the Webwatch list is archived at: http://lyris1.telelists.com/htbin/lyris.pl?enter=webwatch&text_mode=0 (checked Nov 2000)
correspondence with Webmasters when this is both successful and unsuccessful in prompting changes to be made.

1.3.3.7.3.1 Report card

The possibility of formalising the process of users reporting accessibility problems to Webmasters has been considered by the WAI. The WAI were developing a formal process in which users who encounter accessibility problems with a particular site can complete an online 'report card'. The user would be able to enter the details of the site and a description of the problems. The card would then be sent from the WAI to the Webmaster with an explanatory note. Records would be kept of the number of times sites are reported, and the types of problem reported. The WAI would follow up reports to check whether the sites have changed and keep records of this. There is debate of whether sites should be 'named and shamed' if they receive high numbers of reports and do not make any changes. The report card system is not yet in operation. Its success in improving accessibility will depend on whether users use the card and whether page authors respond favourably.

1.3.4 Presentation of content

While there are many ways of encouraging page authors to make their sites accessible to visually impaired people, it would be unrealistic to expect all Web sites to be completely accessible. Furthermore, it is unlikely that all aspects of Web pages can be made accessible with existing browsers and screenreaders because they do not support special tags and may not present information in an accessible way. Therefore a complementary approach is required, and this is occurring with the development of alternative ways of presenting the content of Web pages.

A number of different approaches are being taken to increase the number of, and improve the quality of, the devices used by visually impaired people to access the Web. For example, the development of mainstream browsers that are accessible, the adaptation of screenreaders to work with mainstream browsers, the development of special browsers, and the development of novel interfaces and devices.

35 The Web page reporting on this process is no longer available. It has not been possible to ascertain whether the WAI are still considering this idea.
The User Agent Working Group of the WAI is in the process of developing a set of guidelines\textsuperscript{36} for use by developers of user agents, such as browsers, screenreaders and new devices. The aim of the guidelines is to ensure that user agents are accessible and make use of the new accessibility tags in HTML Versions 4.0\textsuperscript{37} and 4.01\textsuperscript{38}.

The issues relating to the use of mainstream browsers, screenreaders, and special browsers were discussed earlier in this chapter. The use of novel interfaces is discussed in the following Section.

1.3.4.1 Novel interfaces

One approach to improving the presentation of the content of the Web is the development of novel interfaces or devices. One example is the ACCESS project (Morley, Petrie, O'Neill, & McNally, 1998). This project extended the number of ways in which information can be presented and the modes in which the user interacts with the information. The ACCESS project was funded by the European Union. One aim of the project was to investigate the use of a multi-modal interface for a hypermedia system. It is mentioned here because the designers intended to use the interface for a Web browser, although this did not happen within the lifetime of the project. The interface included a variety of devices. The input devices included a joystick, and a touchtablet, as well as a standard keyboard. The output devices included synthesised speech, and a braille display. Information in a hypermedia system was developed in order to evaluate the interface. The information was stored as a variety of different media, such as text, audio clips, tactile pictures, and digitised speech. Being a hypermedia system there were links between nodes. The input devices allowed the user to navigate within and between the hypermedia nodes. Navigation was also assisted by the use of non-speech sounds to indicate various events and content. For example, sounds were used to indicate the presence of a link, and when a new page was loaded. The system was evaluated with blind

\textsuperscript{36} User Agent Guidelines available at: http://www.w3.org/TR/UAAG10/ (checked Nov 2000)
\textsuperscript{37} HTML 4.0 specification available at: http://www.w3.org/TR/html4/ (checked Nov 2000)
\textsuperscript{38} HTML 4.01 specification available at: http://www.w3.org/TR/html401/ (checked Nov 2000)
students who provided feedback on the different devices, the ease of navigation, and the use of sounds.

1.3.4.2 Reformatting of Web pages

The British Broadcasting Company (BBC) developed a tool called Betsie\textsuperscript{39} to make the Web pages on the BBC's Web sites accessible to visually impaired people. The tool was developed in consultation with the RNIB. When the user requests a Web page, Betsie retrieves the page and reformat it before presenting it to the user. Betsie removes images and columns and moves navigational links to the bottom of the page. When a link is selected Betsie retrieves the relevant page. The developers of the tool acknowledge that it has limitations: it only retrieves pages on the BBC Web sites and a few other selected sites, and it does not work correctly with all pages.

1.3.4.3 Future developments in access to the Web

A key area of development in access to computers and the Web is the use of haptic information available via desktop devices such as mice that provide tactile or haptic feedback. Mice that provide tactile feedback are now being released into the mainstream market, for example the iFeel mouse\textsuperscript{40} developed by Logitech. These mice provide tactile information on the objects on the screen to assist (sighted) users in the recognition of, and interaction with, these objects. The user "can experience tactile sensations via the mouse as the cursor travels over icons and menus, command bars, folders, tool bars, Web buttons and hyperlinks". The product information does not mention the potential benefit for blind users.

The Moose is a haptic device that was developed at Stanford University (O'Modhrain & Gillespie, 1997). An aim of the project is to provide access for blind people to GUIs and the Web. It provides haptic feedback from the objects on the screen to enable blind users to navigate. These objects might be the icons and menus of a GUI, or forms and frames in Web pages.

\textsuperscript{39}Betsie home page available at: http://www.bbc.co.uk/education/betsie/about.html (checked Nov 2000).

The researcher believes that this type of mouse is unlikely to be of immediate benefit for blind users. The efficient use of a standard mouse requires some practice in hand/eye co-ordination. Therefore blind people do not generally use mice. If a blind user were to use a tactile feedback mouse, they would need to practice controlling it in order to make, and maintain, contact with on-screen objects. Also, tactile feedback would not be sufficient on its own for a blind user to access the screen. The mouse software would have to be used in conjunction with a screenreader in order for the user to distinguish between different icons or menu items.

The technology of the tactile feedback mouse is being extended for use on the Web. The Immersion Corporation have developed a toolkit to enable Web developers to add the sense of touch to their Web pages using Shockwave and Flash multimedia software\(^1\). This will in theory allow users to feel simulations of the objects on the screen with a tactile feedback mouse. The researcher believes that in the near future blind people will be excluded from these developments because the multimedia software is not accessible, i.e. its content is not available to screenreaders. The required ‘hooks’ will need to be built into this multimedia technology before blind people could use it with a screenreader, let alone with a tactile mouse. Despite the current limitations of this technology, the researcher expects that it will be of benefit to blind users in the future.

1.3.5 Discussion on improving accessibility

An important limitation of many of the different approaches discussed above is that few have been empirically evaluated (with the exception of the evaluation of the Webpage Accessibility Self-Evaluation Test, described in Section 1.3.3.7.2, above). This means that it is not known whether, for example, blind people find Betsie to be a useful tool, or whether devices such as the iFeel mouse will provide useful information. Ideally, all approaches should be evaluated as part of an iterative design process.

---

Much of the rationale for making the web accessible to visually impaired users is based on the expected widespread use of mobile devices and the telephone for Web access. However, it is not just sighted people that will want access to the Web with these devices: they also need to be accessible to visually impaired people. Furthermore the accessibility of such devices needs to be evaluated. Currently there are no known studies of the accessibility of mobile devices or technology such as WebTV.

An important issue is that it is difficult to define accessibility. A page that is accessible to one person, who uses a particular browser or screenreader, may not be accessible to another person using different software. It is difficult to set standards for accessibility when there is no clear definition or consensus on proposed criteria. This particularly applies to the implementation of legislation. Nevertheless, a high level of accessibility should be aimed for, and page authors should be encouraged to consider accessibility in their designs.

1.4 Overview of thesis chapters

This chapter has provided an introduction to this thesis. It has been shown that there are two overall approaches to improving Web accessibility; improving the design of content, and providing alternative ways of presenting content. These two approaches are reflected in the following chapters of this thesis. Regarding the design of content, Chapter 2 presents a review of the literature on the use of guidelines in user interface design. This review provides a background to the study presented in Chapter 3, in which page authors used the Web Content Accessibility Guidelines to adapt Web pages. In Chapter 4 a further study is presented in which visually impaired people evaluated the accessibility of the Web pages adapted in the previous study. Having examined the use of guidelines in the design of content, this thesis goes on to examine the use of haptic devices to present content. Chapter 5 presents a review of the literature on the perception of textures and objects via touch, including the use of haptic devices. This review provides a background to the studies presented in Chapter 6 in which the perception of virtual textures and objects by blind and sighted people is investigated. Finally, Chapter 7 presents the researcher's
conclusions on improving Web accessibility for visually impaired people through the use of guidelines and the use of haptic devices.
CHAPTER 2 – REVIEW OF LITERATURE ON THE USE OF GUIDELINES IN USER INTERFACE DESIGN

2.1 Introduction ........................................................................................................... 42
2.2 Definitions of ‘guidelines’ ...................................................................................... 43
  2.2.1 Related sources of design information .............................................................. 44
2.3 Effectiveness of guidelines ................................................................................... 44
2.4 Use of guidelines .................................................................................................... 46
  2.4.1 Investigating the use of guidelines ................................................................ 46
    2.4.1.1 Results ..................................................................................................... 46
    2.4.1.2 Mosier & Smith’s Conclusions ................................................................. 50
    2.4.1.3 Thoughts on Mosier & Smith’s study ......................................................... 50
2.5 Use of Guidelines in design ................................................................................... 51
  2.5.1 Use of a standard in menu design ................................................................... 51
    2.5.1.1 Results ..................................................................................................... 52
    2.5.1.2 de Souza & Bevan’s conclusions ............................................................... 54
  2.5.2 Use of style guidelines ...................................................................................... 55
    2.5.2.1 Results ..................................................................................................... 56
    2.5.2.2 Thoughts on Tetzlaff & Schwartz’s study ................................................. 58
  2.5.3 Use of a standard by student designers ............................................................ 59
    2.5.3.1 Results ..................................................................................................... 59
  2.5.4 Use of a standard by developers ...................................................................... 60
  2.5.5 Conformance of products to a standard ............................................................ 60
    2.5.5.1 Thovtrup & Nielsen’s conclusions .............................................................. 61
    2.5.5.2 Recommendations ................................................................................... 61
    2.5.5.3 Thoughts on Thovtrup & Nielsen’s studies .............................................. 62
  2.5.6 Comparison of ergonomic criteria with a standard .......................................... 62
    2.5.6.1 Results ..................................................................................................... 62
2.6 Use of Guidelines in evaluation ............................................................................ 63
  2.6.1 Use of a standard to test compliance ................................................................. 63
    2.6.1.1 Final thoughts on Thovtrup & Nielsen’s study ......................................... 63
  2.6.2 Validation of ergonomic criteria ....................................................................... 64
    2.6.2.1 Results ..................................................................................................... 65
    2.6.2.2 Bastien & Scapin’s conclusions and recommendations ........................... 65
  2.6.3 Usability of guidelines ...................................................................................... 66
    2.6.3.1 Development and assessment of the Criteria .......................................... 66
    2.6.3.2 Results ..................................................................................................... 67
    2.6.3.3 Bastien & Scapin’s conclusions ................................................................. 67
    2.6.3.4 Thoughts on Bastien & Scapin’s study ....................................................... 67
  2.6.4 Development of tools for working with Guidelines ........................................... 68
2.1 Introduction

This chapter reviews the literature on the use and effectiveness of guidelines (and other similar tools such as standards and style guides) for the design of user interfaces. Guidelines differ in terms of their specificity, derivation, number, purpose, and the way they are used. This review presents a number of studies in which guidelines have been evaluated for their usability and effectiveness. A study by Mosier & Smith (1986) found that guidelines are used in different ways by different groups of people. They found that guidelines are used at different stages of the development process, in particular the design and evaluation stages. These two stages are reflected in the other studies reviewed here. These studies have found that the users of guidelines experience several difficulties with using guidelines, such as selecting and implementing the guidelines. A common concern of many of the authors cited in this chapter is that guidelines need to be evaluated themselves in order for designers to be able to use them effectively. In other words, before guidelines are used in an iterative design process, they need to be subject to an iterative process themselves. A recommendation from two of the studies is that interactive computer-based tools should be developed in order to overcome the difficulties experienced by designers. The final sections of this review present the issues relating to the development of such tools.
2.2 Definitions of 'guidelines'

de Souza & Bevan (1990) believe that we need to create computer systems that serve the needs of users rather than expecting users to adapt to the ways in which computers function. The achievement of this goal requires system designers to have knowledge of human factors. There is therefore a need for mechanisms to transfer human factors knowledge to software designers. Collections of design guidelines are a common method for this transfer. de Souza & Bevan (1990) find that guidelines vary in terms of their level of specificity and their derivation: some state a small number of principles, others are derived from cognitive psychology or are based on empirical evidence.

Mosier & Smith define guidelines in the same way as de Souza & Bevan. They state that “guidelines represent an attempt to distill [sic] human engineering wisdom into a form useful to software designers” (Mosier & Smith, 1986, p39). Löwgren & Nordqvist (1992) also define guidelines as a document containing general human factors knowledge.

While guidelines are intended to transfer human factors knowledge from experts to software designers, de Souza and Bevan state that guidelines can be used in different ways. They can be used by designers as a direct source of information during the design process, they can be used by human factors experts to provide advice to designers, and be used as a resource in education and training.

Vandervonckt (1999) defines sets of guidelines as prescriptions for a range of user interfaces. Vandervonckt finds that sets of guidelines are presented in conference and journal papers and take the form of statements, perhaps with examples, explanations and comments. Vandervonckt states that individual guidelines have very specific uses. They define a guideline for the design of user interfaces (UIs) as

"a design and/or evaluation principle to be observed in order to get and/or guarantee the usability of a UI for a given interactive task to be carried out by a given user population in a given context" (Vandervonckt, 1999, p82).

In this thesis the term 'guidelines' is used to refer to a mechanism for transferring knowledge from user interface design experts to a particular audience. A set of
guidelines is made up of statements that guide the user in implementing a specific approach that is intended to produce a user interface that enables the end user to perform a task. The intended audience for guidelines may be designers, but they may also be used by software engineers, managers etc. Guidelines may be used differently by different people, and for different purposes, such as informing an overall design, or assisting in a specific design choice. This definition draws on the definitions given above by de Souza & Bevan (1990), Mosier & Smith (1986) and Vandervonckt (1999).

2.2.1 Related sources of design information

Vandervonckt (1999) makes a distinction between guidelines and other sources of design information. Vandervonckt provides definitions of these other sources, which are summarised below.

Design rules
- Set of functional and/or operational specifications for a particular user interface. Presented in a form that requires no further interpretation: as physical rules, screen formats, window templates etc.

Standards
- Set of functional and/or operational specifications for standardising user interface design. Define minimal thresholds for usability. Disseminated by national / international organisations (military, government, industrial). Aim at generalisation: advantage is generalisation across platforms.

Style guides
- Compendium of rules with recommendations for designing a user interface to a given specification: a high-level design document. Defines presentation and dialogue of a user interface. Ensures conformity with corporate specification and consistency. Large organisations often have their own custom style guide.

2.3 Effectiveness of guidelines

Several researchers have recognised that it is important that designers can use guidelines effectively.

"For a user interface standard to increase usability in the resulting products, two condition have to be met: The standard must specify a usable interface, and the standard must be usable by developers so that they can actually build the interface according to the specifications" (Thovtrup & Nielsen, 1991, p335).
Bastien & Scapin identify several techniques that can be used in the improvement of user interface design, including general design guides, sets of guidelines, checklists, standards, heuristics and criteria. The authors state “only a few of these documents have been evaluated in terms of their validity, thoroughness, reliability, effectiveness, and their ease of use by their potential users” (Bastien & Scapin, 1995, p343).

Mosier & Smith are concerned whether guidelines can be used effectively.

“Serious questions can be raised about how effectively design guidelines can be applied. Guidelines contain information which is potentially useful to the interface designer. But are guidelines useful in practice? Who uses guidelines, how do they use them, and with what success?” (Mosier & Smith, 1986, p39).

Several authors have found that designers have difficulties with using guidelines. Root & Uyeda (1993) find that designers have similar problems using styleguides as they do with guidelines. They state that there is “abundant evidence that styleguides are difficult for designers to use and rarely produce products with the intended characteristics … they are also notoriously difficult to develop and validate” (Root & Uyeda, 1993, p32). Bastien, Scapin & Leulier are of the same opinion:

“The evaluation performances obtained by … various approaches [to usability inspection] suggested that much remains to be done to transform these approaches into valid, sensitive, reliable and easy to use methods (by non-specialists as well as specialists) and to incorporate such methods into an iterative design process” (Bastien, Scapin, & Leulier, 1999: p301-2).

Several authors believe that sets of guidelines need to be evaluated in order for them to be more usable by designers. For example, de Souza & Bevan state that “with iterative development of … guidelines it should be possible to rectify the errors and difficulties of interpretation, identification and solution” (de Souza & Bevan, 1990, p439).

Thovtrup & Nielsen consider the use of standards.

“Given the potential future importance of usability standards, it seems reasonable to study the usability of … standards themselves to assess whether developers can actually apply the content of the documents. Not much research is available on this topic yet, but existing evidence does indicate the potential for “meta-usability problems” (usability problems in a usability document)” (Thovtrup & Nielsen, 1991, p335).


2.4 Use of guidelines

Several groups of researchers (including Mosier & Smith, de Souza & Bevan, Tetzlaff & Schwartz, Thovtrup & Nielsen, and Bastien, Scapin & Leulier) have studied how guidelines and similar sources of design information have been used by various groups, including designers, human factors specialists, students, and developers. This work is reviewed in the following sections.

2.4.1 Investigating the use of guidelines

Mosier & Smith (1986) investigated a number of issues concerning the use of guidelines. For example, how effectively can design guidelines be used? Are guidelines useful in practice? Who uses guidelines? And how do they use them, and with what success?

They evaluated a set of 580 guidelines for the design of user interfaces which were contained in a report by Smith & Aucella (1983). The report also contained material to supplement the guidelines. This included a table of contents, several pages of introductory text, a list of references, a glossary, and an index.

A questionnaire was sent to 400 people who had previously received the report containing the guidelines. The questionnaire asked about how the guidelines were read, how specific information was found, how and when the guidelines were applied, how useful they were perceived to be, and how they were used in conjunction with other sources of information. The questionnaire also asked how the guidelines could be improved. Responses from 130 people were analysed. Half the respondents were human factors specialists; others were system analysts, software designers, managers, teachers, researchers, and system engineers.

2.4.1.1 Results

2.4.1.1.1 Reading the guidelines

Mosier & Smith state that guidelines must be read and understood before they can be applied. They found that the respondents read the guidelines in different ways.
Some read the report cover to cover, others just read selected sections or selected guidelines. Half of the managers and designers only skim-read the report. Mosier & Smith believe that the way in which the report was read has implications for the way in which its content will be used. For example, it may be used to answer specific design questions, or more general questions.

2.4.1.1.2 Finding specific information
Mosier & Smith observe that for a guidelines document to be useful, readers need to be able to easily find information within it. The respondents were asked how often they were able to find specific information. The respondents reported variable success: only 57% always found what they were looking for.

2.4.1.1.3 Using and applying guidelines
The respondents were asked for which purpose(s) they used the guidelines. Most respondents reported more than one use of the guidelines. Some reported that they did not use all 580 guidelines, but used the document for reference to look for specific information. The most common uses were to establish requirements in advance of design, as an aid during design, and to evaluate a proposed design.

The respondents were asked whether they would recommend the guidelines for use in the design of future applications, and for what purpose. The most common recommended uses were the same as the actual uses: to establish requirements in advance of design; as an aid during design; and to evaluate a proposed design. However, the percentage of recommended use was higher than actual use in all except one category. Of the respondents who gave answers for both actual and recommended use, most said they would use the guidelines more extensively in the future.

2.4.1.1.4 Usefulness of the guidelines
When asked to rate the ‘usefulness’ of the guidelines on a scale of 0 – 10, respondents gave a mean rating of 7.1, although the rating differed for different groups. Software designers rated the guidelines to be less useful than the other
groups. They gave a mean rating of 6.2 compared to that of the human factors specialists (7.1) and system analysts (7.3).

The respondents were given categories that described possible opinions of the guidelines and how they might be used, and were asked what percentage of the guidelines fell into each category. The categories, and the percentages of the guidelines placed in each category were as follows:

- Guideline is confusing, could not understand it 2%
- Guideline seems to be wrong 2%
- Guideline was not relevant for this particular system 24%
- Guideline seems good, but could not be followed in this particular system for practical reasons 16%
- Guideline is too general to be useful 14%
- Guideline is too specific to be useful 3%
- Guideline was used with some rewording 18%
- Guideline was used directly, just as it is stated 21%

(After Mosier & Smith, 1986, p43.)

The guidelines placed in the first six categories were not actually used by the users of the guidelines: these total over 60%. Only the guidelines placed in the final two categories were used: which total less than 40%. These findings are in contrast to the answers to earlier questions that suggest the respondents used most of the guidelines and would use them in the future.

2.4.1.1.5 Other sources used

The respondents were asked if they used any other sources of information in conjunction with the guidelines report. Seventy-nine respondents cited 28 sources, which covered human factors texts and other sets of guidelines, including an IBM report, a military standard, and various in-house guidelines. This suggests that guidelines cannot necessarily provide all the information required by their readers, particularly if these readers are from different fields. Mosier & Smith do not say how the respondents used such sources, whether they compared the advice of the guidelines with the other source, or used the other source when they could not find what they wanted in the guidelines, or used the other source to find examples.
2.4.1.1.6 Improving the guidelines

The respondents were asked to discuss the problems they had encountered while reading or using the guidelines, and to suggest possible solutions.

2.4.1.1.6.1 Generality of guidelines

The most common problem reported was that the guidelines were too general to be applied. However, as Mosier & Smith point out, the more specific guidelines are, the less they are generally applicable. They do not present a solution for this difficulty.

2.4.1.1.6.2 Finding relevant guidelines

There were several suggestions for solving the difficulties in finding relevant guidelines. Some suggested the report could be re-organised, either by design stage, or by software application. Others thought the navigation within the document could be improved with the use of an expanded index, further indexes by design phase and task, and the use of cross-referencing by keyword.

Mosier & Smith suggest that many of the proposals are attempting to deal with the large size of the document and to minimise the reading of irrelevant material. The latter is required in view of the fact that 24% of the guidelines were found to be irrelevant to the current system. The way in which the guidelines could most usefully be organised would depend on the application.

Some respondents suggested that the guidelines could be made available as an online database, which would allow fast and accurate searching. The authors believe this could be a ‘powerful’ solution to the problems respondents experienced with finding guidelines.

2.4.1.1.6.3 Priorities

Users of guidelines need to decide which of the relevant guidelines they should use. Respondents reported making trade-offs among conflicting guidelines. They also reported difficulties in deciding which guidelines were required and which were ‘frills’. Some respondents suggested that the guidelines should be assigned
priorities, others suggest that each guideline should be assessed in terms of reliability and validity. Mosier & Smith (1986) recognise that the users of guidelines need assistance in choosing guidelines but believe that for different applications the priorities may be different, in which case the importance of each guideline would have to be determined for any particular application. The authors state that it would be possible to get human factors specialists to rate the general importance of each guideline but the variation of importance may be too great. They leave this as an open question.

2.4.1.2 Mosier & Smith’s Conclusions

Mosier & Smith conclude that the guidelines they evaluated were read, used, and recommended to others, and are likely to be used again. The guidelines were mostly applied in the early stages of a design, but were also used in evaluation. Software designers found the guidelines less useful than the human factors specialists did. A continuing problem is how to make the guidelines useful to different groups. Part of this problem lies in the generality of the guidelines, which was an issue raised by both of these groups.

Some of the suggestions made by respondents were incorporated into a new version of the report. The index was expanded, and more guidelines have been added (now 679 in total) providing new material on user interface functions. However, this new version of the report has not been further evaluated.

2.4.1.3 Thoughts on Mosier & Smith’s study

Mosier & Smith’s study investigates respondents’ perceptions of their use of guidelines, but does not consider the extent to which the interfaces designed using the guidelines are usable or conform to the guidelines. The respondents may have evaluated the usability of their designs but the authors did not ask about results of such evaluations. There are questions remaining to be answered such as whether the designs produced by using the guidelines were evaluated, and whether they were considered to be usable.
The authors do not mention whether the guidelines document contained examples that illustrated the implementation of the advice. If examples were present it is unfortunate that Mosier & Smith did not ask about their use since other studies (e.g. Tetzlaff and Schwartz (1991); Thovtrup and Nielsen (1991) – see sections 2.5.2 and 2.6.1 below) have shown the importance of examples in sets of guidelines.

2.5 Use of Guidelines in design

As Mosier & Smith found, guidelines are used by different groups of people, and for different purposes. In this section research on the use of standards and style guides in design is reviewed. In section 2.6, research investigating the use of standards, ergonomic criteria and ‘dimensions’ in evaluation is reviewed.

2.5.1 Use of a standard in menu design

de Souza & Bevan (1990) conducted a laboratory experiment to “evaluate the extent to which designers are able to use guidelines to design menu interfaces, and to suggest how the guidelines could be improved” (de Souza & Bevan, 1990, p435).

The guidelines used in the experiment came from a draft standard for menu interface design prepared for the International Standards Organisation (ISO). This document contained 87 guidelines. A subset of 45 guidelines, which referred to the design requirements for menu structure and presentation, was evaluated. The experiment involved three participants who were all interface designers with at least one year’s experience of designing interactive software.

The participants were given the guidelines one week prior to the experiment. They were asked to read the document carefully and to note anything they found to be unclear, and any other difficulties they had with the guidelines. When the participants attended for the experiment they were given a further 45 minutes to study the guidelines.

The participants’ task was to use the guidelines to improve a paper-based design of a menu-style interface. The design was purposefully poor. In the first part of the experiment the participants were asked to make the changes to the design, as they
thought necessary, based on the guidelines and their own experience. In the second part of the experiment the participants were asked to adapt their new design to conform to the guidelines. This was done by applying the guidelines one by one.

During both stages of the experiment the participants were asked to 'think aloud' and to justify all their design decisions. They were asked to describe what they were doing, why they were trying to do it, and which guideline they were currently using. The experimenter observed and video recorded the designers performing the task.

On completion of the task (or after 3 hours) the participants were interviewed. The aims of the interview were to acquire additional information in order to assess the cognitive and affective costs of using the guidelines, and to identify any difficulties or errors that may not have been observed by the experimenter.

The authors analysed the task of using the guidelines in terms of 'cognitive costs'. These were calculated from the errors made by the designers, and the difficulties they encountered. Errors are defined as design decisions made by the designer that contravened the guidelines. Such errors may or may not have been recognised by the designer. Difficulties are defined as design decisions that were consistent with the guidelines, but which increased the workload for the designer, or increased their feelings of frustration. The errors and difficulties were identified using a protocol analysis and through observations made by the experimenter. The sources of errors and difficulties were divided into four categories:

- Interpretation of the guidelines;
- Identification of the design problem;
- Solutions to the design problem;
- Implementation of the solutions.

### 2.5.1.1 Results

On average, each designer had errors or difficulties with 66% of the guidelines. Only 4 of the 45 guidelines did not cause any difficulties or errors. The authors calculated the percentage of guidelines that were associated with at least one error or difficulty:

- Interpretation: error or difficulty with 44% of guidelines (none of these were recognised by the designers);
- Identification: error or difficulty with 47% of guidelines (2% were recognised);
• Solution: error or difficulty with 31% of guidelines (18% were recognised);
• Implementation: no errors or difficulties (probably because they did not perform on-line implementation).

de Souza & Bevan describe examples of the errors and difficulties that were observed for each category. They explain the causes of these problems and make recommendations for improving the guidelines.

2.5.1.1.1 Interpretation of guidelines

The designers made errors in, and had difficulties with, the interpretation of the guidelines. These errors and difficulties were not recognised by the designers. The errors and difficulties were caused by several different aspects of the guidelines, such as “lack of clarity of the conditions under which the guideline should be applied” and “lack of clarity of the nature of the proposed design solution” (de Souza & Bevan, 1990, p438).

The authors recommend that the text of some guidelines be re-structured and re-worded to prevent the misinterpretation of the conditions, and specifying the conditions under which they should be used. The authors also recommend that some terms should be clarified, and some implicit conditions of use should be made explicit. For example, they recommend the use of phrases such as ‘in designing groups within a menu...’ and ‘if a menu panel has...’. They also recommend the use of terms that designers are more likely to be familiar with, such as the term ‘label’ instead of ‘designator’.

2.5.1.1.2 Identification of design problems

The designers made errors in the identification of the design problem, some of which they recognised and others they did not. These errors were caused by a “lack of information about the procedure to apply a guideline” (de Souza & Bevan, 1990, p438).

The authors recommend that information should be provided regarding procedures, for example the procedure to be used to assess the frequency of the use of specific
options in an interface. They also recommend that the scope of the guidelines should be clarified.

2.5.1.1.3 Solution of design problems

The designers made errors in the solving of the design problem, some of which they recognised and others they did not. These errors were caused by a lack of specific information on the use of the guidelines. For example, information about procedures, design goals, and the benefits of applying the guidelines.

The authors recommend that some procedures should be described, for example the procedure used to carry out empirical testing to identify the terminology used by potential users. They also suggest that design goals be made more explicit, for example that the distinctiveness between option names in an interface needs to be improved.

2.5.1.1.4 Evaluation of designs

An assessment of the extent to which the design conformed to the guidelines was performed by 'expert analysis'. It was found that on average only 11% of the guidelines were violated. The authors state that "in many cases it was possible [for the designers] to produce a satisfactory design despite the difficulties experienced in applying the guidelines" (de Souza & Bevan, 1990, p438). The authors do not describe how the expert analysis was performed, or by whom. Neither do they say what constitutes a 'satisfactory' design, or make any assessment of how serious the violations were.

2.5.1.2 de Souza & Bevan's conclusions

In order for guidelines to be effective and to be interpreted correctly it is essential that they contain clear information about:

- The design goals and benefits;
- The conditions under which the guideline would be applied;
- The precise nature of the proposed solution;
- Any procedure which must be followed to apply the guideline.

(de Souza & Bevan, 1990, p439)
It is important to clarify the design goals, and the procedures that should be used to achieve them. Otherwise “designers may apply what they understand to be the principles underlying the guidelines to design problems for which they were not intended” (de Souza & Bevan, 1990, p439). Such clarification will reduce the designer’s cognitive costs and decrease the likelihood of the misinterpretation of the guidelines.

de Souza & Bevan conclude that “with iterative development of ... guidelines it should be possible to rectify the errors and difficulties of interpretation, identification and solution” (de Souza & Bevan, 1990, p439). However, they recognise that “iterative refinement will maximise the potential usefulness of the guidelines, but it cannot itself guarantee that designers will actually use them”. A quote illustrates the reluctance of one participant to use the guidelines

“I do not need these guidelines because I know how to [design interfaces] already. I think that it is important to read [the guidelines] one or two times and once you have got the information from the back of your mind to the front of your mind, you do not need [the guidelines document] anymore, so you put it away ... possibly if you have a problem you might have a quick look at it” (de Souza & Bevan, 1990, p440).

2.5.2 Use of style guidelines

Tetzlaff & Schwartz (1991) conducted a study to investigate the role of style guidelines in the development of interface designs that conform to a specific style. The document contained 300 pages (250 of which formed a Reference section). The main document had three sections that described the style, its underlying principles and the interface design process. It prescribed styles for the appearance and behaviour of interface objects. The final Reference section described each available interface object via a textual description, a graphic illustration and list of conditions of use.

The style guide was being developed over the course of the study, with iterations from the results from previous participants. Each participant used the most recent version.
The nine participants were all experienced interface designers, but had varying amounts of experience of computer programming and had varying degrees of experience of using the guidelines. The authors of the guidelines also took part in the experiment as 'expert' participants in order to provide a baseline against which to compare the designs of the 'novice' designers.

Participants were supplied with a problem statement and the style guide document and were asked to create a paper-based design, while thinking aloud. The participants were video-recorded and a protocol was developed from the recording. After the task the participants were interviewed and completed a questionnaire.

### 2.5.2.1 Results

Tetzlaff & Schwartz state that the guidelines used in their study contained many concepts that were new to the participants. However, they do not say whether these concepts were new to some or all of the participants. These new concepts were problematic for the participants for several reasons:

- Some concepts were only partially understood.
- Some were difficult to implement.
- Some distinctions that were clear to the authors of the guidelines were not at all clear to the participants.
- Some concepts were perceived to contain potential usability problems.

Regarding the perception of usability problems, Tetzlaff & Schwartz suggest that if the use of guidelines results in unusable applications this will taint designers' opinions of both the individual guidelines and the whole set of guidelines.

#### 2.5.2.1.1 Examples

Tetzlaff & Schwartz describe the importance of the role of the graphical examples in the use of guidelines. The participants said they preferred learning from the examples to learning from the rest of the guide. They inspected the examples carefully, referred to them frequently, copied them in their design, and stated that they would like more to be included. A problem with the examples was that they were intended to illustrate the text of the guide, but the participants tended to treat the text as an annotation to the graphical examples. This meant that the participants missed important information within the text. For instance, one example that
illustrated the use of radio buttons happened to contain scroll bars. Some participants included the scroll bars in their design with the radio buttons, which was not intended. But in addition they did not consult the section on the use of scroll bars.

2.5.2.1.2 Strategies

Tetzlaff & Schwartz observed that the participants used two different strategies for incorporating the guidelines into their task. The majority began the task immediately and only consulted the guidelines when they felt it was relevant. Only a minority began by reading the introductory material. The majority found that they could not get very far without consulting the guidelines, but only read them until they found what they thought they required and then continued with the task. This meant that they missed information that the authors of the guidelines intended them to read.

Tetzlaff & Schwartz found that the alphabetical arrangement of the reference section "encouraged browsing and facilitated serendipitous contact with other entries" (Tetzlaff & Schwartz, 1991, p331). The authors also found that the designs alone may not have been a sensitive enough test of the effectiveness of the guidelines. The protocols revealed problems that were not apparent in the paper designs but may have been highlighted in the implementation of the designs:

"Problems were finessed by designers and evaluators alike. Designers copied by rote or invented where they did not understand, and evaluators compensated for designer error, treating serious conceptual flaws as minor labeling [sic] errors" (Tetzlaff & Schwartz, 1991, p332).

2.5.2.1.3 Evaluation of designs

The designs were judged by 'evaluators'. Despite the problems reported by the participants and those observed by the researchers, the participants produced designs that were judged to conform to the style guide. One of the evaluators was a co-author of the guidelines, but it is not clear who the other evaluators were. Nor is it clear what methods they used to assess conformance, or whether the assessment was performed under experimental conditions. It seems that the assessment did not involve any conformance criteria, or if it did, how such criteria related to the guidelines. Any such criteria did not appear to be available to the designers, but may have been useful.
2.5.2.1.4 Tetzlaff & Schwartz's conclusions

Tetzlaff & Schwartz conclude that when guidelines are used "much written material may go unread, misunderstood, and consequently unheeded. The longer and more complex the document the greater the risk" (Tetzlaff & Schwartz, 1991, p333). They believe that the most effective way of communicating style requirements is for the style to be illustrated in examples and interactive demos, which would be facilitated by toolkits and supported by iterative usability testing. The role of guidelines would be to complement these other techniques, to describe those principles that may be difficult to infer from the examples.

2.5.2.1.5 Recommendations

Despite the problems with the examples Tetzlaff & Schwartz emphasise the importance of examples and suggest that they should be made less ambiguous, and perhaps should be interactive. Another recommendation is that guidelines should explicitly state their scope and limitations and specify the conditions under which each component should be used.

In order to overcome many of the problems associated with guidelines the authors recommend that guidelines should be developed primarily to complement toolkits. However, a toolkit may be able to illustrate certain design choices, but the designer may not be able to infer from this the underlying principles of the choices. Therefore there are some design criteria that cannot be available from a toolkit which should be presented in guidelines.

2.5.2.2 Thoughts on Tetzlaff & Schwartz's study

There are a number of details missing from Tetzlaff & Schwartz's paper, which makes it difficult to assess some of their results. They mention that some of the participants had more experience of using the guidelines than others, but do not say whether the designs of these participants were any different from those of the others. It might be expected that their designs had higher levels of conformance. On the other hand the authors state that some participants who had previous experience did not read the sections of the guidelines that they recognised, which meant they missed
subtle changes that had been made. So it is possible that the designs of these participants had lower levels of conformance. A possible solution to this problem of those familiar with the guidelines not reading all sections might be to list the recent changes that may not be obvious and ensure that those who have read the guidelines before are made aware of the list.

Tetzlaff & Schwartz do not say whether there were any differences in the designs produced by the participants who had more experience of procedural or object oriented programming and those that had less.

The guidelines evaluated by Tetzlaff & Schwartz were modified in response to early participants in the study, but the authors do not state how the designs produced by later participants compared to those of earlier participants. It would have been interesting to know whether and how the modifications had influenced later designs.

2.5.3 Use of a standard by student designers

Thovtrup & Nielsen (1991) investigated the usability of a user interface standard in a series of studies. In one study, students on an interface design course were given a two-page standard to use in the design of an interface. They were also given a three-page specification of a sample system that complied with the standard. The task was to be assessed for the students' coursework, so it was expected that they would have a high degree of motivation to perform the task.

2.5.3.1 Results

Thovtrup & Nielsen compared the designs to a checklist of the elements specified in the standard. The interfaces designed by the students had an average of 71% compliance to the standard. The standard specified an unusual use of keys in the interface. It was this element of key allocation in which the students' designs had least compliance. The authors conclude, “real, live systems influenced many of the designers more than the standards document” (Thovtrup & Nielsen, 1991, p336).
It is not clear why the students were not provided with the checklist used by Thovtrup & Nielsen. Having a checklist may have encouraged the students to check their designs and discover any deviations from the standard.

### 2.5.4 Use of a standard by developers

In another study Thovtrup & Nielsen (1991) examined the use of an in-house user interface standard. Developers who regularly used the standard were asked their opinions of it. They mostly had positive opinions, but some negative points were made. For example, some stated that there were not sufficiently good programming tools to support the requirements in the standard.

The developers were asked whether their designs complied with the standard under all circumstances. Some (33%) said they would sometimes deviate from the standard. One of the reasons given for deviation from the standard was that the standard restricted the creativity if the designers and precluded new design ideas.

In relation to the usability of the standard, the developers were divided as to whether the rules of the standard were easy to remember and easy to apply. Of the developers 53% said that the rules were difficult to remember, compared to the 27% who found them easy to remember. In addition, 27% said that the standard was difficult to apply, compared to 40% who said it was easy to apply.

### 2.5.5 Conformance of products to a standard

In a further study, Thovtrup & Nielsen (1991) inspected and assessed the conformance of three products developed by one company to the standard. The authors developed a checklist that listed all the rules and divided them into three groups: mandatory rules; voluntary rules; and guidelines for the design process.

It was found that all three products broke between 7 and 12 of the 22 mandatory rules in the standard. However, Thovtrup & Nielsen judged most of the deviations to have a minor impact on the overall usability of the product (they mainly reduced the feeling of 'product family' across the products). Of the 6 voluntary rules, an average
of 3 had been broken. All of the design process guidelines had been complied with, except one concerning the use of a think aloud test.

The developers of the products were asked why the products deviated from the standard. Most claimed that they had chosen alternative design solutions because they found them to be better than the solution provided by the standard. Others said that they had not yet had time to implement compliance, or that their development tools did not support compliance with the standard. In no case did the developers state that they had implemented a deviation without realising because they had misinterpreted the standard.

2.5.5.1 Thovtrup & Nielsen's conclusions

Thovtrup & Nielsen conclude that standards are “very likely to be violated” (1991, p340). The participants in the three studies presented above were described as ‘highly motivated’, and the developers were described as having a “higher than average interest in usability” (1991, p337). Despite these factors, the participants were not able to apply the standard effectively, and the participants’ designs contained many violations of the standard. Thovtrup & Nielsen suggest that existing systems influence designers, perhaps more than the use of standards.

2.5.5.2 Recommendations

Thovtrup & Nielsen make three recommendations. Firstly, that tools should be developed to support the implementation of user interfaces that comply with standards. Secondly, that concrete examples should be used in standards, and that such examples must not contain any inadvertently misleading information. Finally, Thovtrup & Nielsen highlight the importance of the design of the standards document itself. They suggest that such documents should be designed using recognised principles and should include good ‘access mechanisms’ such as an index, a table of contents, word lists, glossaries, and checklists. They also suggest that a computerised access mechanism, such as hypertext, could be used.
2.5.5.3 Thoughts on Thovtrup & Nielsen’s studies

In the study of the use of a standard, Thovtrup & Nielsen do not appear to consider the possible relationship between those that found the rules difficult to apply and those who did not always comply with the standard.

The authors state that their experience from these studies “indicates that a checklist of specified design elements and rules helps tremendously during a conformance quality assurance review” (Thovtrup & Nielsen, 1991, p340). It seems strange that they did not recommend providing users of standards with a checklist that would assist them in checking their own work prior to any test of conformance.

2.5.6 Comparison of ergonomic criteria with a standard

Bastien, Scapin & Leulier (1999) conducted a study to compare a set of ergonomic criteria with a standard developed by the International Standards Organisation (ISO), in the evaluation of a database.

The 17 participants were divided into three groups. Six used the Criteria, five used the Standard and six formed a Control group. All the participants were students. They had studied psychology for four years and had had a few hours of lectures on ergonomics. The task was to find problems in an interface for a database containing music references. The interface was designed to contain usability problems. The authors do not explicitly define what they mean by the term ‘usability problem’. The researcher assumes they mean aspects of the interface which contravened the ergonomic criteria or the standard.

2.5.6.1 Results

The median number of problems found by the Criteria, Standard and Control groups respectively were 86.17, 62.2, and 61.83. Bastien et al. also compared the results from the Criteria group with results from a previous study. In the previous study human factors specialists using the same criteria found 18.5% more problems than human factors specialists who did not use the criteria. In contrast the current study involved novices but they found 51.5% more problems than the human factors specialists who did not use the criteria. Bastien et al. therefore conclude that novices
benefit more from the use of these criteria than specialists do. However, they believe there is still much work to be done to develop valid, reliable, and easy to use evaluation methods and to incorporate such methods into the design process.

2.6 Use of Guidelines in evaluation

Having examined the use of guidelines at the design stage of the development process, this review will now consider the use of guidelines at the evaluation stage.

2.6.1 Use of a standard to test compliance

A further study by Thovtrup & Nielsen (1991) examined the abilities of developers to apply an in-house standard in a conformance test (the same standard as was used in the studies described above in sections 2.5.4 and 2.5.5). The participants were 15 developers from the company to which the standard belonged. They were given four screendumps from an interface that had been designed intentionally to include 12 deviations from the standard. The task was to find the deviations in 15 minutes, using the standard.

On average the participants found only 4 of the 12 deviations, with a maximum of 7 found. Thovtrup & Nielsen described this performance as ‘surprisingly poor’. The participants judged some details of the interface to be deviations when they were not, with an average of 1.6 wrongly identified deviations per developer.

The authors analysed the developers’ use of different parts of the standard document. The participants were found to make comparatively less use of the main document that contained the actual specification and the rules, but relied on the examples and lists of approved keys and system terms.

2.6.1.1 Final thoughts on Thovtrup & Nielsen’s study

It is not clear why the developers were restricted to 15 minutes for the evaluation as this is less than four minutes per screendump. It is assumed that in non-experimental conditions they might take more time than this. This is particularly important considering the finding of the previous study by Thovtrup & Nielsen in the same
paper that 53% of those interviewed said they found the rules of the standard difficult to remember, and 27% found them difficult to apply. The authors do not appear to consider the relationship between the two findings.

2.6.2 Validation of ergonomic criteria

The overall goal of the work of Bastien & Scapin (1992) is to develop methods for the evaluation of user interfaces: methods that can be used by human factors specialists and non-specialists, such as interface designers. The aim of the study was to validate a set of ergonomic criteria that had been developed in a previous study. It was hypothesised that if the definitions, rationale and examples were sufficient, the criteria would be understandable, and participants would therefore be able to use them to identify user interface problems. To assess whether the validity of the criteria depended on the experience of the participant, two groups with different levels of experience were used.

The set of ergonomic criteria consisted of eight main criteria. Some of these main criteria were subdivided into one or two levels of 'subcriteria', referred to as 'elementary criteria', of which there were eighteen. Each elementary criterion was illustrated by a short definition, a rationale and some examples. The following is an extract from the list. The items shown in bold text are defined as elementary criteria because they are not subdivided.

- Guidance
- Prompting
- Grouping and distinguishing items
- Grouping and distinguishing items by location
- Grouping and distinguishing items by format
- Immediate feedback
- Clarity
  (After Bastien & Scapin, 1992, p187.)

The participants were 12 human factors specialists and 12 non-specialists. The participants were given a series of 36 usability problems (two problems per elementary criterion) in the form of screen dumps and textual explanations of the interaction context. The problems were either selected from an existing system by a group of experts, or created through modifications to the system. Participants were
presented with one problem at a time and asked to identify which criterion was broken by each usability problem.

Bastien & Scapin analysed the time spent reading the criteria and associated information, as well as the participants' responses. Both groups took less than 20 minutes to read the information and 55 – 57 minutes to complete the whole task.

2.6.2.1 Results

There was no significant difference between the groups in terms of the time spent reading the documentation, or the time taken to perform the task. The experienced participants correctly identified 63.7% of the violated criteria and the inexperienced participants identified 56%, but the difference between the groups was not significant.

2.6.2.2 Bastien & Scapin's conclusions and recommendations

Bastien & Scapin conclude that there was some confusion over the meanings of some of the criteria. The authors present a breakdown of the criteria mistakenly identified, and the other criteria they were systematically confused with. The authors suggest that the definitions of the criteria that were confused with only one other criterion may not have been sufficient to allow their discrimination. This may have meant that the experience of the human factors specialists interfered with their judgement. The authors suggest that some criteria were confused with more than one other criterion was because these criteria were too broadly defined.

Since there was no significant difference between the overall performances of the two groups the authors conclude that experienced and inexperienced people can use this type of list of criteria. However, 13 out of the 18 elementary criteria would need to be re-defined to ensure that experience did not interfere with their application. The criteria will need to be re-evaluated following the required modifications. Bastien & Scapin's final conclusion is that it is difficult to provide clear and non-ambiguous definitions of criteria.
2.6.3 *Usability of guidelines*

In a further study the same authors, Bastien & Scapin (1995), use the term ‘dimensions’ to refer to usability principles, criteria, heuristics and standards. (It is assumed that they would also place guidelines in this group.) Bastien & Scapin find that dimensions vary on different levels, such as the number of individual rules, and their level of precision. The differences stem from the use of different design strategies. Some recommendations have been extracted from knowledge about cognitive processes and organised into high level dimensions. Other approaches have been to organise published dimensions into a common structure in which personal experiences are coupled with existing principles, and databases of usability problems compared to sets of usability heuristics.

Bastien & Scapin wanted to establish the impact of the available ‘dimensions’ and answer questions regarding their internal validity, independence, explicitness, consistency, and external validity.

2.6.3.1 *Development and assessment of the Criteria*

Bastien & Scapin (1995) used an empirical approach to design a ‘valid’ dimension. They took the results of research and a large set of guidelines and translated them into rules. These rules were then grouped into sets based on the rationale for their use. After several iterations of this process a set of criteria were produced. The authors concluded that the criteria could be judged to be ‘valid’ and ‘complete’ and that they synthesised most or all of the available guidelines. In a previous study (Bastien & Scapin, 1992), the set of criteria was evaluated, and was found “to allow an accurate description and classification of the usability problems found by experts in an evaluation task” (Bastien & Scapin, 1995, p345). Having established the validity and completeness of the criteria, Bastien & Scapin also wanted to assess the reliability and effectiveness of the criteria. The reliability was determined by whether the criteria could be used to classify design flaws consistently. The effectiveness was determined in terms of the role of criteria as an aid to evaluation.

The study involved two groups of human factors specialists. Their task was to find design flaws in an interface that purposefully included 500 such flaws. The task was
performed in two phases. During the first 'exploration-diagnosis' phase both groups explored the interface and made general comments on particular states of the interface. During the second phase the same interface states were presented in a replay of the exploration phase and evaluated by the participants. One group performed the evaluation using the criteria, while the other group did not use the criteria.

2.6.3.2 Results

The results showed that in the first phase both groups found a similar number of design flaws, and in the second phase the use of the criteria increased the number of problems found and improved the diagnosis of problems. The use of the criteria also increased the completeness of the evaluation and decreased the number of evaluators required.

2.6.3.3 Bastien & Scapin's conclusions

The authors conclude that their criteria are reliable, usable and improve evaluation. However, they also recognise that the criteria still have certain limitations. The criteria do not specify how evaluators should systematically explore an interface, or how each criterion should be applied. The criteria should therefore be expanded to provide more detail. For example they should be prioritised, and define pre-requisites (such as evaluator experience) as well as define how to systematically examine all states and elements of the interface.

The authors intend to assess future iterations of the set of criteria with designers and end-users, perform a cost / benefit analysis, and compare their method with other evaluation methods. They state that the criteria method is not intended to replace other methods, but to provide a means of ensuring compliance with guidelines, and to uncover design flaws prior to user testing.

2.6.3.4 Thoughts on Bastien & Scapin's study

Unfortunately, Bastien & Scapin do not say how many people participated in the study, nor what proportion of the 500 design flaws were found. It is therefore difficult to compare their results with those of other studies.
What is particularly useful from this study is the emphasis on the importance of the wider context in which the guidelines are used for evaluation, for example, considering how the user interface should be systematically explored, and the amount of experience the evaluator has. In other words, it is important to consider the situation in which the guidelines are used, as well as what they are used for, and how.

2.6.4 Development of tools for working with Guidelines

Previous research has found that “even though user interface guidelines and style guides contain much useful knowledge, they are hard for user interface designers to use” (Löwgren & Nordqvist, 1992, p181). Some researchers have concluded that a better approach would be the development of tools to support the use of guidelines in order to minimise dependence on guidelines for example Tetzlaff & Schwartz (1991) and Thovtrup & Nielsen (1991).

Löwgren & Nordqvist (1992) were motivated by such conclusions and have developed a knowledge-based tool that contains design knowledge drawn from general guideline documents and some toolkit specific style guides. The tool can evaluate a user interface design that is produced by a user interface management system.

2.6.4.1 Problems with guidelines

Löwgren & Nordqvist define guidelines as a document containing general design rules. They found that guidelines were a popular medium for the propagation of human factors knowledge and the knowledge was often validated through controlled experiments or experience. However, they also found in the literature that there are problems with the application of guidelines, such as guidelines being too general and difficulty in interpreting guidelines.

Löwgren & Nordqvist believe that “guidelines and style guide documents are inefficient ways of communicating human factors knowledge to the designer” (Löwgren & Nordqvist, 1992, p181). Their solution was to augment the design and
implementation environment of a user interface management system (UIMS). They added a knowledge base containing human factors knowledge to the UIMS. The tool evaluates a design at the designer’s request in order to provide feedback for subsequent iterations of the design.

2.6.4.2 Issues relating to development of tools
Löwgren & Nordqvist raise several issues relating to the development of knowledge-based tools. First of all the knowledge has to be acquired. The authors did this by collating publicly available guideline sources, such as Smith & Mosier (1984). A user interface expert then interpreted the guidelines and an independent expert then validated the interpretation. Löwgren & Nordqvist emphasise that this is not a trivial or mechanical process.

2.6.4.3 Evaluation of design
The next issue is of how the tool should produce comments on the designer’s design. There are two possible approaches. 1) The design can be analysed in respect of its potential flaws: analytical critiquing. 2) The tool can generate its own solution to the design problem and compare this to the designer’s solution: differential approach. Löwgren & Nordqvist find that the second approach is not appropriate for user interface design because the problems are not well defined and there can be more than one valid solution to a problem.

2.6.4.3.1 When to evaluate
Löwgren & Nordqvist raise the issue of the level of evaluation performed by a tool. The authors’ tool only evaluated the design at the request of the designer, but they report that in other studies designers left it too late to request an evaluation. A possible solution would be for the tool to alert the designer as soon as it detected a problem. However, Löwgren & Nordqvist believe this concept requires empirical testing before being taken any further.
2.6.4.3.2  Level of evaluation

Löwgren & Nordqvist find that the tool cannot make some judgements and these have to be left to the designer. For example, one instance of a rule contained in a tool is “applications should provide accelerators for frequently used menu items” (Löwgren & Nordqvist, 1992, p185). There is no way for a tool to determine whether an item will be used frequently. Löwgren & Nordqvist present three different solutions to this problem: the designer can make the decision; or the tool can incorporate specifications about user tasks and domain semantics. A further solution, preferred by Löwgren & Nordqvist, involves logs of user testing with the interface under development being collected and analysed. The data from such testing can be incorporated into the tool. For example, the tool could check the frequency that a particular item was chosen from a menu, and then check that the designer had assigned accelerators.

2.6.4.4  Löwgren & Nordqvist’s conclusion

Löwgren & Nordqvist have used a tool to apply design knowledge and guidelines to a user interface design produced in a UIMS. This is a valuable step towards bringing human factors knowledge closer to the design process. However, much design knowledge can only be applied with consideration on the situation of use. Data collected during tests of the prototype can be used to bring use-related design knowledge to bear. Empirical testing of the tool was expected to begin soon after the paper was published.

2.6.4.5  Thoughts on Löwgren & Nordqvist’s study

It would be interesting to read the results of the intended testing of the tool to see whether designers find it easier to use the knowledge contained in a tool than to use a set of guidelines. However, the researcher has not been able to find any further work by these authors.

2.6.5  Purposes of tools

Vandervonckt (1999) has edited a special issue of the journal 'Interacting with Computers', on the development of tools for working with guidelines. Vandervonckt
sees the purpose of such tools as supporting the underlying activities of working with guidelines, such as searching for guidelines, and applying them. Vandervonckt believes that tools could also support other activities such as collecting, writing, propagating, and verifying guidelines, and even teaching and learning about guidelines.

2.6.5.1 Issues regarding tools

Vandervonckt considers many of the issues associated with the use of tools. In particular, who could use such tools and how would they support the user, and how tools might change the role of the designer.

“If we consider a person who recently learned software ergonomics and using such tools, is this person able to assume the same roles of a person who acquired extensive experience in human factors, but not using such tools? For instance, could a human factors expert be replaced by an inexperienced designer assisted by software tools? Could an expert be supported by tools?” (Vandervonckt, 1999, p83).

2.6.5.2 Research agenda for computer-based tools

Vandervonckt presents an agenda for future research on the development of tools. Below is a summary of Vandervonckt’s suggestions.

- Collection, gathering, merging, and compilation of guidelines;
- Sorting, classifying and integrating these initial guidelines into organisational structure;
- Incorporation of organised guidelines into a methodology;
- Automation and digital representation of the guidelines;
- Refinement of the methodology to optimise the benefits.

(After Vandervonckt, 1999.)

2.6.5.3 Use of examples in guidelines

One of the many problems Vandervonckt found to be associated with the use of guidelines was they were not sufficiently illustrated. The author suggests the provision of more examples could be a possible remedy. However, this suggestion does not address the issues identified in the literature relating to the use of examples. For example, the case of designers following examples to the exclusion of the text, and designers implementing ambiguous details contained in examples, as reported by Tetzlaff & Schwartz (1991). These problems need to be resolved for any set of guidelines, whether they are paper-based, on-line, or incorporated into a tool.
2.7 Development of guidelines from empirical evidence

In sections 2.5 and 2.6 research on the use of guidelines (or similar tools) in design and evaluation has been reviewed. In this section a small study is reviewed in which guidelines have been developed from empirical evidence. This study is of interest because it raises the question of where the information contained in guidelines comes from and whether the use of guidelines actually serves the needs of end-users of the system. The study suggests a methodology for involving the end-users in the development of guidelines that will be used to develop systems or interfaces.

In 1997 Leonard Kasday conducted a small survey regarding the preferences of visually impaired people for different types of alternative text for images. The results were reported to the Browser and Screen Reader discussion list (BASR-L) on 17 Feb 1997. (This discussion list is not archived.) There were 11 responses to three questions. A 1-10 scale used to measure the responses. Three points along the scale were identified as 0 = very bad, 5 = neutral, 10 = very good.

The first question asked whether alternative text should be provided for all images, without exceptions. The mean response was 7.2. One respondent said they wanted to be able to decide for themselves whether an image is relevant. In contrast, another said they wanted to be able to get useful information as efficiently as possible.

The second question asked whether pictures should be described literally. For example, should a picture of a customer representative have the alternative text 'smiling customer representative', or 'customer service'? The former describes the actual image, whereas the latter describes its function. The mean response was 6.8. One respondent liked the former because it was informative, but not too wordy. In contrast, another respondent thought alternative text should be as short as possible.

The third question asked about alternative text for images used for bullet points in lists. Respondents were asked to rate several different examples. These are listed below with their mean ratings.

'Tiny blue sphere': 4.0.
Some respondents did not want to hear the phrase ‘tiny blue sphere’ repeated for each list item, and this outweighed the principle of wanting to know what an image looks like. Other respondents liked the examples that only contained punctuation marks because they could set their screenreaders to ignore punctuation, and would then have the choice of having it repeated for each list item.

Kasday proposes a guideline for the use of alternative text based on these results. All images should have alternative text, which should give an indication of the content, or ‘message’ of the image. Even images that may appear to have no relevance should have alternative text so that a blind user can determine the relevance for themselves.

Kasday also proposes a convention for bullet points. Lists should be preceded by text such as ‘list with 22 items marked by tiny blue spheres’, and each bullet image should have ‘*’ as their alternative text. This convention would cater for those who want to know what the image looks like, and those who do not want a descriptive phrase repeated for each list item.

The studies presented above examine the use of guidelines and similar tools at the design stage of the development process. It has been seen that designers experience many difficulties in the use of such tools, such as finding a relevant guideline, and translating examples for the current design problem. The importance of examples has been highlighted. An important factor in the evaluation of the effectiveness of guidelines is the evaluation of the product of using the guidelines. Interestingly, of the three studies that evaluated a design that incorporated guidelines, two found that the designs ‘conformed to the styleguide’ (Tetzlaff & Schwartz, 1991) or was ‘satisfactory’ (de Souza & Bevan, 1990). This was despite the problems the designers experienced in using the guidelines. In contrast, one of the three studies found that designers achieved only 71% compliance to a standard.
2.8 Conclusions on the use of guidelines

This review has considered several studies that evaluated the use of guidelines (or similar tools) in the design and evaluation of user interfaces. The study by Mosier & Smith (1986) found that user interface guidelines are used by different groups, in different ways, for different purposes. Therefore these groups have different requirements for the guidelines and had different perceptions of their ‘usefulness’. Several studies found that participants had difficulties with finding, choosing, interpreting and implementing guidelines.

de Souza & Bevan (1990) found that designers made errors that contravened the guidelines - sometimes on purpose, but mostly without realising. Thovtrup & Nielsen (1991) found that designers’ experience had more influence than the use of a standards document on their current designs and that developers and students both had relatively poor performance in complying with the standard. Tetzlaff & Schwartz (1991) found that designers had difficulty in interpreting guidelines and relied on pictorial examples to the exclusion of the text. The findings of these studies emphasise the need for the evaluation of guidelines in order to facilitate their effective use by designers.

A number of researchers have suggested the use of computer-based tools to overcome some of these problems, for example Tetzlaff & Schwartz (1991) and Thovtrup & Nielsen (1991). Only one tool is known to have been developed for implementing the knowledge contained in standards and guidelines, but it has not yet been evaluated. It is essential that future tools are evaluated for usability and for the usability of the interface designs that are produced with the tool.

Several studies highlighted the importance of examples: how participants used examples to the exclusion of the text, and found some to be ambiguous. In two cases, the examples and participants’ experience of other systems were found to have more influence than the guidelines or standards themselves.

The evaluation of designs has provided further insights into the ways in which the guidelines were interpreted. However, as Tetzlaff and Schwartz point out, verbal
protocols are important for identifying problems with the guidelines that may not be
apparent in the designs, particularly in paper designs that have not been
implemented.

The literature reviewed in this chapter has highlighted a number of factors that
should be considered in the development and use of guidelines, which are listed
below. The researcher hesitates in labelling these factors as ‘guidelines’ because of
the inherent problems that have been identified in this chapter. However, it seems
useful to highlight the issues and recommendations gathered from the literature.

First issues regarding the content and presentation of guidelines are presented. These
are followed by issues relating to the development of guidelines. Finally, issues
relating to the development and use of tools are presented.

2.8.1 Issues relating to the content and presentation of guidelines

- Support navigation (e.g. for finding guidelines)
- Index, table of contents, checklist etc. (Mosier & Smith, Thovtrup & Nielsen)
- Provide priorities (for choosing guidelines) (although these may be different for
different contexts) (Mosier & Smith, Bastien & Scapin 1995)
- Provide clarifications of terms and make conditions of use explicit (to prevent
misinterpretation) (de Souza & Bevan)
- Provide information on the procedure for applying guidelines (de Souza &
Bevan, Bastien & Scapin 1995)
- Clarify the scope of the guidelines (de Souza & Bevan)
- Make the design goals of the individual guidelines explicit (de Souza & Bevan)
- Explain new concepts (otherwise they will be difficult to implement) (Tetzlaff &
Schwartz)
- Provide sufficient number of examples (Vandervonckt)
- Ensure examples are concrete do not contain ambiguous information (Tetzlaff &
Schwartz, Thovtrup & Nielsen)
- Alphabetical arrangement of guidelines may facilitate browsing and
serendipitous encounters with important information (Tetzlaff & Schwartz)
- For evaluation with guidelines, specify how the interface should be
systematically examined (Bastien & Scapin 1995)

2.8.2 Issues relating to the development of guidelines

- Guidelines need to be relevant to designer’s design current problem (Mosier &
Smith)
- It cannot be guaranteed that the guidelines will be followed (de Souza & Bevan)
- Sets of guidelines should be developed iteratively and involve evaluations with
experts and users (de Souza & Bevan)
• When evaluating, examine verbal protocols as well as designs (the former may contain information not contained in the latter) (Tetzlaff & Schwartz)
• Bear in mind that examples may be followed to the exclusion of the text (and therefore important information may be missed) (Tetzlaff & Schwartz, Thovtrup & Nielsen)
• Guidelines may only be consulted when there is a perceived need (Tetzlaff & Schwartz)
• Guidelines may not be read in full, in particular the introduction and background information may be missed (Tetzlaff & Schwartz)
• The designer’s experience of other systems may have an impact on their designs, as well as the guidelines (Thovtrup & Nielsen)
• Designers may implement their own solutions because they are perceived to be better than those presented in the guidelines (Thovtrup & Nielsen)
• Consider using hypertext to facilitate navigation (Thovtrup & Nielsen)
• Novices may benefit more from a set of criteria than experts (Bastien, Scapin & Leulier)
• It is difficult to develop criteria that are clear and non-ambiguous (Bastien & Scapin 1992)

2.8.3 Issues relating to the development and use of tools

• Tools may be useful for implementing interfaces that comply with guidelines or standards (Thovtrup & Nielsen)
• Knowledge needs to be acquired, then interpreted into guidelines, then these must be validated (Löwgren & Nordqvist)
• How will the tool evaluate designs? Either critique the design or produce a new design for comparison (Löwgren & Nordqvist)
• When should the tool evaluate? (Löwgren & Nordqvist)
• Consider that the tool cannot make all judgements – human judgement is also required (Löwgren & Nordqvist)
• The appropriateness of a tool depends on the particular design area (Löwgren & Nordqvist).
# Evaluation of Guidelines in Development of Web Pages

## 3.1 Introduction

## 3.2 Method

- **3.2.1 Design**
- **3.2.2 Participants**
  - **3.2.2.1 Number of participants**
- **3.2.3 Equipment**
  - **3.2.3.1 Hardware and software**
  - **3.2.3.2 Video recordings**
- **3.2.4 Think aloud method**
- **3.2.5 Materials**

## 3.3 Procedure

- **3.3.1 The task**
- **3.3.2 Allocation of the task**
- **3.3.3 Pilot of the procedure**
- **3.3.4 Final procedure**

## 3.4 Data analysis

## 3.5 Results

- **3.5.1 Overview of participants’ actions**
- **3.5.2 Decisions with potential negative impact on accessibility**
- **3.5.3 Problems reported by participants, and solutions suggested**
  - **3.5.3.1 New tags for accessibility**
  - **3.5.3.2 Testing pages**
  - **3.5.3.3 Examples provided in Guidelines**
  - **3.5.3.4 Navigation in and between documents**
  - **3.5.3.5 Overall structure of Guidelines**
  - **3.5.3.6 Cost vs. benefit**
  - **3.5.3.7 Intended audience**
  - **3.5.3.8 Further information required**
- **3.5.4 Positive comments on Guidelines**
- **3.5.5 Discussion with WAI on results**
  - **3.5.5.1 General response to the reports**

## 3.6 Discussion on the usability of the Guidelines

- **3.6.1 New tags**
- **3.6.2 Priorities**
- **3.6.3 Examples**
- **3.6.4 Navigation**
- **3.6.5 Audience**
3.1 Introduction
This chapter presents the first in a series of two studies. These studies examine the use and the content of the Web Content Accessibility (WCA) Guidelines\(^{42}\) developed by the Web Accessibility Initiative\(^ {43}\) (WAI). In this chapter and the next the WCA Guidelines will be referred to here as ‘the Guidelines’. The Guidelines document and its associated documents are described in Chapter 1, Section 1.3.3.4.1. A number of the HTML tags that are mentioned in this and further chapters are described in Appendix 1.

The aim of this first study was to investigate how the Guidelines would be used by a group of page authors. The need to evaluate guidelines has been established by several researchers, one of whom wrote:

"Guidelines have been the traditional means of transferring HF [Human Factors] knowledge to designers (and especially to designers not trained in HF). Considerable effort and resources have been expended in developing such guidelines. However, although they exhibit potential for transferring HF knowledge … there has been little research validating their effectiveness" (de Souza, Long, & Bevan, 1990, p340).

In previous studies, described in Chapter 2, guidelines and standards have been evaluated by designers and the designs that resulted from the use of the guidelines or


\(^{43}\) Web Accessibility Initiative home page available at: http://www.w3.org/WAI/
standards have been tested for conformance to the guidelines or standards. However, none of these studies involved user testing of the designs. To fully evaluate the WCA Guidelines the researcher believed it was essential that blind users should test the pages designed to the Guidelines. Therefore the aim of the second study in this series was to evaluate the accessibility of the pages that were the product of this first study.

As described in the introduction of this thesis, Web accessibility for visually impaired people is a many-layered issue. The design of Web pages, browsers and screenreaders contribute to this accessibility. The root of many accessibility problems has been found by the World Wide Web Consortium\textsuperscript{44} (W3C) to be the way in which key HTML elements are implemented in Web pages. The W3C provides a Factsheet\textsuperscript{45} on the WCA Guidelines, which gives an overview of the elements that can present barriers for people with disabilities:

- images and imagmaps without alternative text;
- misleading use of structural elements;
- video without description or captions;
- lack of alternatives to frames or scripts;
- tables that are difficult to decipher when read in a linear fashion;
- poor contrast between foreground and background.

The researcher would add two other items to this list: forms without clear labels; and link (anchor) text that is not self-explanatory. The scope of this study was limited to the five key elements, imamaps, images, forms, frames and tables. It is hoped that other elements, such as scripts, video, audio, etc. will be the subject of future research.

3.2 Method

3.2.1 Design

In this study the usability of the WCA Guidelines was evaluated. This is the first usability study involving these guidelines. In the light of previous research, it was expected that the participants would encounter difficulties with various aspects of using the Guidelines. The design of the study was a ‘between subjects’ design.

\textsuperscript{44} World Wide Web Consortium home page available at: http://www.w3.org/ (checked Nov 2000).

\textsuperscript{45} Fact sheet for “Web Content Accessibility Guidelines 1.0” available at: http://www.w3.org/1999/05/WCAG-REC-fact.html (checked Nov 2000).
Participants used the Guidelines to adapt Web pages containing a range of different elements. The dependent variables were the problems encountered by the participants and the comments they made. A simplified version of a think aloud protocol was used and the sessions were video recorded for future analysis. The procedure was piloted with the first participant and then improved for the other participants.

It was not feasible to ask all participants to deal with the full range of elements of interest because they did not all have prior experience of working with all the elements. Therefore a procedure was developed in which the elements were allocated to those participants who had prior experience of implementing them.

3.2.2 Participants
The participants were recruited via posters within the university. The posters requested people who had experience of creating Web pages to contact the researcher. No specific level of experience was requested. A number of people responded to the poster who did not have the required skills in creating Web pages. These people were thanked for their interest and informed that they were not able to take part in the study. This left 12 volunteers who met the requirements for the study, who went on to participate. The participants were all university students (9 under-graduates and 2 post-graduates), except one who was a school pupil. The students were from a range of disciplines: six from Computer Science; three from Psychology; one from Humanities; and one from Engineering. The ages of the participants ranged from 14 to 35 years. The mean age was 23 years (median = 21).

The participants' experience of creating Web pages ranged from having a few personal pages to having large personal sites, and some having been employed to create larger sites for companies. The number of pages they had previously created ranged from 4 to over 100, with a mean of 37 (median = 40).

The mother of this participant contacted the researcher to say that her daughter would be interested in taking part. When the researcher spoke to the young woman it was clear that she had a good knowledge of HTML. It was therefore decided that she should be allowed to participate because her contribution would be comparable to the other participants.
The participants were paid £5 per hour for the time spent in advance preparing for the study and the time spent in the laboratory performing the task.

3.2.2.1 Number of participants
The number of participants involved in the current study is comparable to the numbers involved in previous usability studies conducted by Wright & Monk (1991), Virzi (1992) and Nielsen (1994).

Wright & Monk (1991) compared the number of usability problems found by 13 teams of two student evaluators with the number found by expert evaluators. The student teams and experts evaluated the same interface using a 'relaxed' think aloud method with one participant. This method is described briefly in Section 3.2.4 below. Wright & Monk found that the 13 teams identified 90% of the problems identified by the experts, that is, with 13 participants.

Virzi (1992) conducted an evaluation of the manual of an interface with 12 users. Virzi used the advanced version of the think aloud method, which is described briefly in Section 3.2.4. Virzi found that the first few users uncovered most of the problems: 65% of the problems were found by three users and 80% were found by five users.

Nielsen (1994) conducted a study in which two groups of students ran evaluations of two interfaces. They used a 'simplified' version of the 'think aloud' method (described briefly in Section 3.2.4). Both groups found that one user found an average of 29% of the usability problems. Overall, 86% of the problems were found by 6 users. Nielsen recommends that 4 ± 1 users are sufficient for finding a large number of problems. He concludes that the identification 86% of problems is sufficient in 'discount' usability engineering, particularly in an iterative design process, in which a design is likely to be evaluated again in the future.

In the light of these studies it was expected that a large number of the problems in the WCA Guidelines would be identified with 12 participants.
3.2.3 Equipment

3.2.3.1 Hardware and software

All the participants used a PC computer (rather than a Mac or a mainframe). The participants used a variety of software to perform their task. They were all provided with the software that they had indicated as their preference in the initial questionnaire. Seven participants used a text editor, such as Wordpad, Notepad or VI to edit the HTML code. Four participants used an authoring tool, either FrontPage or DidaPro. One of the participants who worked on imagemaps used Mapedit to create the imagemap. The participants used a variety of browsers to view their pages: Netscape (NS), Internet Explorer (IE), and Lynx (see Table 3.1). Tables 3.1, 3.2 and 3.3 show the number of participants who used different combinations of software, under different platforms.

Table 3.1: The authoring tools, editors and browsers used by the participants.

<table>
<thead>
<tr>
<th>Authoring tool</th>
<th>Editor</th>
<th>Browser</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>DidaPro</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>FrontPage</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Mapedit</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Notepad</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Wordpad</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>VI</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Netscape</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Internet Explorer</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lynx</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.2: The number of participants who used each combination of editor or tool and browser.

<table>
<thead>
<tr>
<th>Editor / tool and browser combination</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>FrontPage + IE</td>
<td>3</td>
</tr>
<tr>
<td>Wordpad + NS</td>
<td>3</td>
</tr>
<tr>
<td>Notepad + NS</td>
<td>2</td>
</tr>
<tr>
<td>Notepad + IE</td>
<td>1</td>
</tr>
<tr>
<td>Mapedit + IE</td>
<td>1</td>
</tr>
<tr>
<td>DidaPro + NS</td>
<td>1</td>
</tr>
<tr>
<td>VI + Lynx</td>
<td>1</td>
</tr>
</tbody>
</table>
3.2.3.2 Video recordings

Two video recordings were made of the participants during the study; one of the computer screen, the other of the participant. The participant was recorded using a standard video camera. This recorded what the participants said, what the researcher said, when the participant referred to the Guidelines, and when they made a note of something. In order to record the computer screen the computer was linked to a ‘mediator’ device which, in turn, was linked to a video recorder so that everything that occurred on screen while the participant was working was recorded. This created a useful record of everything the participant did that might have otherwise been lost. For example, if a participant implemented an example from the Guidelines but then changed it, there was a record of the first version. Also, if a participant implemented a particular tag, and then for some reason removed it, it was possible to view the first version.

In order to synchronise the two recordings during analysis, the recording of the screen included an audio recording of the ambient sound. This was of poor quality, but it was good enough to synchronise the two recordings.

3.2.4 Think aloud method

The method used was a ‘simplified’ or ‘relaxed’ version of the think aloud method, as used by Nielsen (1994) and Wright & Monk (1991) respectively. The usual, or ‘advanced’, think aloud method involves participants thinking aloud while performing a task. The researcher does not say anything except to remind the participant to ‘keep talking’. The utterances of the participant are recorded, transcribed and then subjected to a formal protocol analysis (Nielsen, 1994). The simplified version of this method involves the participant thinking aloud while performing a task and the researcher identifying misconceptions and problems from the participant’s utterances (Nielsen, 1994). The key difference between the
advanced and simplified versions is that the former involves a formal protocol analysis.

As part of Nielsen’s ‘discount usability engineering’ philosophy he simplified the think aloud method so that it would be easier and cheaper to use, and so that it could be used by experimenters who were computer scientists, but not usability specialists. The current researcher is not a ‘usability specialist’ but has a computer science background, and therefore expected to be able to use this technique effectively.

Wright & Monk (1991) used a ‘relaxed approach’ to the think aloud method. This allows the experimenter to ask questions of the participant, which maximises the possibility of identifying and correctly interpreting the problems experienced by the participant. Wright & Monk found this to be a superior method because it is more natural, which makes it easier for the experimenter to learn and use.

The use of a simplified, relaxed version of the think aloud method in the current study meant that participants were not asked to describe everything they did, but more to describe their thoughts of whatever was happening. The intention was not to investigate the actual cognitive processes of the participants, but to ascertain their opinions and identify the problems they experienced. For example, a participant would not be expected to announce that they were going to switch from viewing their file with the browser to viewing it with the editor, because it was obvious to the researcher that this was happening, and it was being recorded on video. Instead, participants were encouraged to talk about what was happening, for example, to explain that they had tried one approach to solving a problem and they were now going to try another. There were occasions on which the researcher needed to ask for clarification of what the participant was doing or thinking. On such occasions she would ask questions such as, "what are you doing now?" or "why did you do that?". Participants were invited to ask questions on the task procedure at any time.

3.2.5 Materials

Each participant was given a file that contained the element(s) they were to work with. The file for the images and tables was a condensed version of a paper the
researcher had written. The file included headings, paragraphs, and links as well as a table and three images (see Appendix 2). The files for the frames, forms and imagemaps were also created by the researcher.

Participants were also provided with a printed copy of the file they were working with, as it would be displayed in Netscape, so that they could refer to it when it was not visible on the screen, and could make notes on it. Participants were also provided with paper and a pen should they wish to make any other notes.

A selection of HTML reference books was placed beside the computer and participants were informed that they could refer to them if they wished. A printed copy of the Guidelines was provided for those participants who preferred to use it. The participants were invited to bring any reference materials (or addresses of online materials) to the laboratory that they might normally refer to when making Web pages.

3.3 Procedure

3.3.1 The task

The task (for all the participants except the first) was to adapt an existing Web page, making it accessible by following the advice contained in the Guidelines. The participants were shown the files to be adapted, both in a browser (to see what the page looked like) and in a text editor (to view the HTML code). It was explained that the researcher intended to test the pages they adapted with blind people later in the study. The participants were told that they were not under examination and that it was the Guidelines that were being examined, also that there were no 'right' answers. The participants were asked to proceed once the researcher had explained the think aloud method and the participants had asked any questions.

3.3.2 Allocation of the task

In advance of attending the laboratory, the participants were asked which HTML elements they had experience of, what software they would prefer to use, and whether they would prefer to use a Mac or a PC computer. (The questions are
provided in Appendix 3.) This information was used to allocate the HTML elements to participants who had experience of them. The information was also used to provide the participants with the hardware and software they preferred.

In order to give each participant a task of a reasonable size one element was allocated to each participant; either a form, frames, an imagemap or a table. Making images accessible is a comparatively easy task and so images were allocated along with tables. The four different tasks were tables plus images; frames; forms; and imagemaps. Each element was allocated to a participant who had prior experience of implementing that element.

Images and tables are among the most common causes of accessibility problems on Web sites, as identified by the W3C in the Factsheet47 on the WCA Guidelines. Therefore half (6) of the participants performed the tables-plus-images task, and the other 6 worked on the other elements (two with forms, two with frames, and two with imagemaps).

3.3.3 Pilot of the procedure
The first participant was allocated the tables and images task. They were given a text file that contained the information to be included in the Web page and asked to insert HTML tags to make it into a Web page. However, this task was found to be too much work for one session. Inserting the basic tags was too time-consuming and did not contribute to the investigation of the Guidelines.

The first participant did not see the Guidelines in advance of attending the laboratory. This participant appeared to find it difficult to work with the Guidelines having never seen them before and needed some time to familiarise themselves with the relevant documents.

---

3.3.4 Final procedure

On the basis of the problems identified with the procedure for the first participant the procedure was changed in two key ways for the subsequent participants. Firstly, these participants were given an HTML file containing the elements they were to work with, rather than a text file. This meant that the participants could focus on improving the accessibility of the page, rather than on its construction.

Secondly, these participants were given access to the Guidelines and related documents in advance of attending the laboratory. This gave the participants time to read through the documents and familiarise themselves with their structure, rather than having to work with documents they had never seen before. They were given the URLs of the documents, but told they could collect printed versions from the researcher if they preferred (two did so). The participants were asked to familiarise themselves with the content and the structure of the documents and to make notes of anything they did not understand, or had some other difficulty with. (de Souza & Bevan (1990) also gave participants the guidelines they were to evaluate seven days in advance.)

When the participant appeared to have finished the task, the researcher asked whether they were happy with what they had done, and what they would normally do at this stage. This enabled the participant to describe any further action they might take if this was their own page, such as testing it with another browser.

At the end of the session the researcher asked a series of questions in a semi-structured interview (see Appendix 4 for the questions). Participants were asked their course and age, their opinions of the Guidelines, and whether they thought they would refer to them again in the future.

Participants spent an average of just over 3 hours performing the whole task. This is made up of an average of just over 1 hour reading the Guidelines in advance (time reported by the participants), and just under 2 hours in the laboratory (time measured by the researcher). Those that performed the task involving tables and images spent more time in the laboratory than those that performed the other tasks.
3.4 Data analysis

In order to analyse the participants' work, the researcher viewed both video recordings. From the recording of the scene notes were made on the problems the participant had, comments they made about what they were doing, and errors they made which were recognised by the researcher but not the participant. Notes were also made of details to be obtained from the recording of the screen, such as any HTML code that was created and then discarded. From the recording of the screen notes were made on discarded code, which section of which document the participant was reading when they encountered a difficulty, etc.

A full transcription was not made of the sessions. While a full protocol analysis is important for examining the cognitive processes of a participant, in this case such processes did not form the central issue. It was more important to collect qualitative data on the way in which the participants used the Guidelines, the problems they experienced, and the end product of the task.

3.5 Results

This section presents the results of the study: what the participants did; and on what basis they made design decisions. The data from the first (pilot) participant are included in the results because the procedure they followed was the same as that followed by later participants, except that the first participant was asked to produce an HTML page from scratch, rather than being provided with an HTML page. Also the data from this participant were found to be comparable with that of the subsequent in that this participant had experienced similar difficulties with using the guidelines. The problems the participants encountered are presented, along with suggestions they made to improve the Guidelines. The researcher reported many of the problems and suggestions to the WAI Web Content Working Group. The ensuing discussions are summarised in Appendix 5 and discussed at the end of this results section.
3.5.1 Overview of participants’ actions

The participants were observed to act on the advice of the Guidelines and several commented that they had acted on the advice rather than following their usual practices. Participants also took actions that they said were based on their own experience and knowledge. These were actions that they would have taken outside of the laboratory. The following provides an overview of the actions that the participants took on the basis of the Guidelines. Actions taken on advice of the Guidelines are indicated by (G), and those based on existing knowledge are indicated by (K).

Tables
- Three participants implemented Headers and IDs. (G)
- One re-worded headers in order to remove sub-headers.
- One split the table to make it easier for all users. (K)
- Two provided a Summary. (G)
- One provided a Caption. (G)
- One participant used table headers rather than using bold text. (G)

Frames
- Two participants implemented no-frames tag. (1 = G, 1 = K)
- One changed layout from vertical to horizontal.
- One added Titles to Contents and Papers frames. (G)
- One used Arial and Fontsize.
- One tried Accesskey for link in Table of Contents. (G)
- One placed square brackets around text links to separate them. (K)

Forms
- One participant grouped controls using Fieldset and Legend. (G)
- One used Legend. (G)
- One re-grouped the controls. (G)
- One added default text in address field. (G)
- One used Label for some controls. (G)
- One used Label on all controls. (G)
- One tried Accesskey. (G)
- One used Tabindex. (G)

Imagemaps
- Two participants implemented alternative text links. (K)
- One separated links with vertical bars. (K)
- Two gave each section ALT text. (K)
- One gave whole image ALT text. (K)
Headings
- Two participants checked that the headings were correctly nested, for example H1 followed by H2. (G)
- One checked that the headings used the Heading tags, rather than bold text. (G)
- One used Heading tags for headings instead of the Font and Fontsize tags. (G)

Images
- Six participants gave images ALT text, except on one figure. (G)
- One implemented Longdesc but replaced it with D-link when found it was not supported. (G)
- One added a caption. (G)

Structure
- One participant created links throughout the page to the top of the page, divided the page into sections to reduce the need for scrolling, provided links to the next section in each, and placed a table of contents at the beginning of the page. (K)
- One changed the layout of the authors names and addresses in the header, removed the superscript numbers, made the email addresses mailto links, and added the words 'Tel' and 'Fax' before the numbers. (K)

Testing
- One participant checked their page with Weblint and HENSA validators. (K)
- One checked their page with Bobby before starting work. (K)

Other
- One participant created mailto links for email addresses. (K)
- One added 3-point list to abstract. (K)
- One provided additional punctuation. (K)
- One changed text of links for references. (K)

3.5.2 Decisions with potential negative impact on accessibility
None of the decisions described above were likely to have a negative impact on the usability or accessibility of the page. However, some other decisions made by the participants were more likely to have a negative impact. These decisions are outlined below. In some cases the participants purposefully did not follow the Guidelines. These are indicated by ‘*’.

Tables
- Two participants decided it would be too much work to translate the table into a linear format, or to implement any of the other recommendations. *

Forms
- One added a note indicating that the form worked best with IE.
Images
- One participant decided that figures 1 and 3 did not require ALT text because the purpose of the images was described in the text. *

Imagemaps
- One participant removed vertical bars between links because they were not sure how a screenreader would read them.

Headings
- One participant did not use H1 for the first heading because they believe that as long as the headings are in the correct order, i.e. H3, H4, H5, then they do not need to start with H1. *

Structure
- One participant used a table to align text.

Links
- One participant used italics in links to make them more visible, particularly for partially sighted people.

Other
- One participant removed an Accesskey because it did not work in their browser. *
- One placed all text within the Head element.
- Two checked their pages by visually scanning it for anomalies or other aspects that they thought may cause accessibility problems, but may have missed problems that were not visual.

3.5.3 Problems reported by participants, and solutions suggested

3.5.3.1 New tags for accessibility

Background
The Guidelines recommend the use of certain tags that were introduced in HTML Version 4.0, whose purpose is to improve accessibility. These tags include Accesskey, Headers and IDs, and Longdesc. The tags are new to HTML and new to page authors, who may be knowledgeable about HTML but know little about accessibility. Page authors do not necessarily know how the assistive technology used by people with disabilities works and therefore do not know how users will interact with the new tags. This can make it difficult to understand how to implement the tags. Some participants did implement new tags because their aim was to improve accessibility. They carefully followed the examples and then tried to test the tags. This testing mostly involved seeing how their usual browser would display the tag. In many cases the pages did not look any different to them, which

* It could be argued that this participant made the right decision not to repeat existing information. However, the decision is reported here because it contravened the Guidelines.
confused the participants. They did not know why the new tags were not visible. Either the example contained incorrect syntax, or they had implemented it incorrectly, or their browser did not support the tag. Although the participants could have tested the tag in another browser, or run the page through a validator, few of them tried this.

The researcher suggested to the participants that they may not be able to see the results of the new tags because their browser may not support them. Two participants removed new tags because they did not want their page to contain tags that did not appear to work. Some of the participants were frustrated that the Guidelines recommend tags that are not yet supported by their browsers. They said they would have liked to have been informed that this was the case. They felt they then would have had the choice as to whether to implement the tags.

The outcomes of this confusing situation were that participants did not know if they had solved the accessibility problem, or whether any assistive technology supports the tags. From this overview it is clear that the issues of new tags (whether they are supported, how the user might interact with them, and how they can be tested) are closely related. These issues are presented in the following sections covering the implementation of new tags, unsupported tags and the priorities of unsupported tags. It is difficult to make clear distinctions between the issues because there is much overlap. The following sections are presented in as logical order as possible, but the reader may find some repetition of points.

3.5.3.1.1 Implementing new tags

3.5.3.1.1.1 Transferring existing knowledge

The problem
Participants did not seem confident in transferring their existing knowledge of related tags to these new tags. For example, the participants who worked with images were all familiar with the use of the ALT tag. They knew that the result of using this tag would not be visible in their browser at the same time as the images it was associated with, i.e. they knew the text was an alternative. However, the same participants who
experimented with the use of the Longdesc tag were not sure whether it should be visible in their browser, i.e. they were not sure if it was an alternative.

These difficulties may have the same cause as those found by Tetzlaff & Schwartz (1991). Tetzlaff & Schwartz state that the guidelines used in their study contained many concepts that were new to the participants. These new concepts were problematic for the participants for a number of reasons:

- Some concepts were only partially understood.
- Some were difficult to implement.
- Some distinctions that were clear to the authors of the guidelines were not at all clear to the participants.
- Some concepts were perceived to contain potential usability problems.  
  (After Tetzlaff & Schwartz, 1991.)

3.5.3.1.1.2 Support in implementation

The problem

Some participants stated that they needed additional support in implementing some of the tags recommended by the Guidelines, especially for new tags, such as Accesskey. Five of the twelve participants consulted one or more of the HTML reference books provided, looking for further details not provided by the Guidelines. One participant was searching for specific details about how to implement Accesskey, but found it was not mentioned in one of the reference books (Zakour, Foust, & Kerven, 1997). In this case the participant was more dependent on the Guidelines and Techniques documents than they might have been for other, more commonly used tags.

Discussion

It would not be possible for the Guidelines and Techniques documents to provide full support for the implementation of all the tags that are recommended. However, it would be possible for the WAI to identify which tags are not described in popular HTML reference books, and provide additional support for such tags. Clearly, it is a difficult balance to strike between providing comprehensive information about a tag and keeping the sizes of the documents to manageable levels. The HTML 4.0 defines all the tags recommended by the Guidelines.
A provisional solution might be to make the links from the Guidelines and Techniques documents to the HTML 4.0 specification more clear. Several such links exist, but their text does not indicate that they link to the specification. This would be a provisional solution because the specification itself is not a particularly easy-to-read document.

3.5.3.1.2 Browser support of new tags

The problem

The participants found that the browsers they were using, Internet Explorer (IE) and Netscape (NS) did not display the new tags they implemented, such as Accesskey and Longdesc. These browsers did not support Accesskey, Longdesc, or Headers and IDs (and NS did not support Fieldset and Legend). This lack of support of the new tags by the main browsers caused several difficulties for the participants.

Participants stated that they were not sure whether the reason the tag was not displayed in their browser was because they had implemented the tag incorrectly, or because it was not supported, or even because it was not meant to be presented ‘visually’ in the browser. They did not know whether they had increased the accessibility of the element they were working with. Some participants believed they may have made a mistake and carefully compared their work with the example, although they did not find any errors. Worse still, in terms of accessibility, some participants removed a tag because it did not appear to work. On a large scale it could be disastrous for accessibility if many authors were to try a tag that improves accessibility and then remove it because they believe it does not work. When one participant spent a lot of time implementing the Headers and IDs for a table and then found that these tags were not supported they said “the Guidelines lead you on a wild goose chase” (YS).

Participants seemed to be aware that the main browsers do not support the same set of tags as each other, and that there are some tags that are only supported by one main browser. However, they expected the tags recommended by the Guidelines to be supported by the main browsers. While participants seemed to accept that some
tags may only be of use to people with disabilities, they still expected their browser to support the tags.

Following these difficulties the researcher informed the participants that many new tags were not currently supported by the main browsers or by special browsers. Some participants stated that they were frustrated when they found that browsers used by visually impaired people did not support the tags they had implemented because implementing them had not solved the immediate accessibility problem. Some even felt they had wasted time implementing such tags if they are not currently supported. All the participants who implemented tags that are not yet supported said that they would like the Guidelines to inform them that this is the case.

3.5.3.1.3 Priorities of new tags

Background

The priority ratings of the Guidelines make the distinction between Checkpoints that 'must' be followed (priority 1), those that 'should' be followed (priority 2), and those that 'may' be followed (priority 3). The current priorities make no distinction between tags that are currently supported by the main browsers and those that are not yet. Previous versions of the Guidelines recommended the use of interim techniques to be used until browsers supported tags such as Headers and IDs. This involved the use of paragraph breaks between the cells of a table. However, this technique did not appear in the versions of the Guidelines used by the participants, and is not present in more recent versions. This means that a page author 'must' implement some tags that are not yet supported, for example Headers and IDs for tables.

The problem

To one participant it seemed unreasonable to state that authors must use tags that are not yet supported.

Solution suggested by participant

The participant suggested that the Guidelines should be less prescriptive and acknowledge that there are always exceptions to rules.
3.5.3.1.4 Testing new tags

The problem

Several participants used their usual browser to test whether the tag was correctly implemented. This is problematic because the major browsers are 'forgiving' in that they often ignore invalid code. This may lead authors to assume that their code is correct or valid when it is not. The WAI believes that browsers should not be used in place of HTML validators to check that code is correct: the best way to check that code is valid is to use a validator such as that developed by the W3C\textsuperscript{49}. However, in previous versions of the Guidelines this was not made explicit. The result was that only two of the participants used a validator while the others all relied on their usual browser to show any errors. This may not be the best practice, but it was nonetheless common practice.

The participants who implemented new tags, such as Accesskey and Longdesc, and found that the tags were not supported by their browsers, did not seem to be aware of any alternative methods of testing whether the tag was implemented correctly. The Guidelines recommend the use of an HTML validator, but few of the participants read the relevant section, and seemed to be unaware of any other method of validating their HTML code.

Discussion

The ideal way for these participants to overcome these difficulties with testing to see if they had correctly implemented the tags would be to use a validator for HTML 4.0, rather than testing in their browser. The participants' lack of awareness of testing methods may be due to their lack of professional experience in managing a Web site. Some of them may never have had reason to test their pages in any way other than perhaps viewing them in a different browser. But even among the few participants who had been employed to develop Web sites there was only one who said that they would normally use a validator to test their pages.

\textsuperscript{49}W3C Validator available at: http://validator.w3.org/ (checked Nov 2000).
3.5.3.1.5 Using a validator

The problems

One participant read the recommendation of the use of a validator and tried to use that developed by the W3C. However, the validator required information that the participant was not able to provide, due to lack of experience and knowledge. The participant continued the validation process but was surprised when the validator rejected many of the ‘ordinary’ tags used in the file. The output from the validator was difficult for the participant to understand and they eventually gave up with it, which meant that they were still not sure whether their code was valid.

Another participant also had difficulties with a validator. When they tried to validate their page they found that the validators they normally use did not accept the new tag Longdesc. This was because Longdesc is part of HTML 4.0 and the validators only validated HTML 3.2.

3.5.3.2 Testing pages

Background

The Techniques document suggests that page authors should check their pages with alternative browsers, such as Lynx and WebSpeak. It also suggests that pages should be checked in GUI browsers such as IE and NS, particularly with different settings, including with images and sounds switched off, and with older versions. The document acknowledges that it is not possible to run more than one version of the same browser on a computer.

The problems

One participant pointed out that it can be difficult to use different GUI browsers if one’s computer does not have the capacity for multiple browsers, or if downloading software from the Internet is not permitted onto a networked computer. An alternative approach might be to ask users of other browsers to check pages, but this participant said that it was not always easy to find users of other browsers, particularly older browsers.
Another participant found that when they tried to view their page with different settings in the same browser they could not change the way in which the page was presented. They were trying to view the page without the images being loaded, but when they switched this option off in IE the images kept loading. The researcher suspects this may be due to the images being saved in the browser’s cache. The Guidelines could acknowledge this potential difficulty and perhaps suggest that page authors clear their browser’s cache when attempting this type of testing.

3.5.3.3 Examples provided in Guidelines

3.5.3.3.1 Visual and Speech rendering

Background

Examples of the implementation of recommended tags are provided in the Techniques document. There were several aspects of the examples that caused problems for participants. Some participants made suggestions for improvements to the examples.

Some HTML tags are intended to improve accessibility for visually impaired people. Such tags are likely to be presented by the user’s software in synthesised speech, or less commonly, on a braille display. The Techniques document briefly described the potential problems associated with the presentation of some tags in synthesised speech and provided examples of HTML code, but did not provide examples of how such examples might be rendered, either visually, or in synthesised speech.

The problem

The participants either did not find the description of the problems with synthesised speech at all, or they did not find the description sufficient. The participants did not know how the new tags might be rendered by a special browser, or in synthesised speech, or how a user might interact with them. This made it difficult for the participants to implement these tags. An example is the use of the Accesskey tag. The Techniques document provides an example of how to implement Accesskey and explains its purpose, but does not explain how it would be presented to a user or how they would interact with it.
Solutions suggested by participants

Participants suggested that the Guidelines could show how each example would be rendered visually, in their browser and in other browsers. For example, the example illustrating the use of Longdesc could show that the result of using this tag would not be displayed in a graphical user interface (GUI) browser, such as NS. The example illustrating the use of Fieldset and Legend could show how these tags might be displayed in browsers such as NS. A participant suggested that screen-shots could be used to illustrate the way different browsers rendered different tags. They had seen these used in HTML reference books and found them to be useful. Participants also suggested that the Techniques document could illustrate how the ‘special’ tags would be rendered in synthesised speech.

Discussion

In theory it is a good idea for the examples provided in the Techniques document to show how different tags might be rendered by different browsers. However, some tags are displayed differently in the common GUI browsers, such as NS and IE. For example, in IE, the Fieldset tag is rendered as a box around the set of form controls and Legend is rendered as a label for the box. However, neither tag is supported by NS and is therefore not rendered at all. These and other tags are probably rendered differently by other, less common, browsers. Even very common tags, such as the ALT tag, are displayed differently by different browsers: Lynx displays all ALT text because it cannot display images, whereas NS and IE only display ALT text when images are not loaded. It would therefore be difficult for the Techniques document to show how such tags are rendered in all browsers.

3.5.3.3.2 CSS rendering

Background

The issues discussed above regarding the rendering of tags visually and in synthesised speech also apply to the rendering of Cascading Style Sheets (CSS). CSS may provide solutions to many of the current accessibility problems, but many browsers, particularly special browsers, do not yet support them.
**The problem**

Some participants found that the examples of CSS provided in the Techniques document assumed knowledge and experience of CSS, which they did not have.

**Solution suggested by participants**

Participants stated that would like the Techniques document to provide examples of how CSS might be rendered in different browsers, and presented in synthesised speech.

### 3.5.3.3.3 Cut & paste

**The problem**

Many of the participants who followed the examples provided in the Techniques document did so by re-typing the example in their file.

**Solution suggested by participant**

A participant suggested that the examples could be provided in such a way that enabled a page author to cut-&-paste them from the Techniques document to the page author's file. The WAI have since provided examples that can be used in this way.

### 3.5.3.3.4 Order of examples

**The problem**

Some participants were observed to implement the first example they came to, rather than reading through all the examples in order to find the most appropriate. It is difficult to identify the underlying reason for this. It may be that the participants were not supported in finding or choosing the most appropriate solution. For example, if the most appropriate example was the last example listed, the participants would have to read all the other examples in order to find it.
3.5.3.3.5 More examples needed

Background

Previous versions of the Techniques document provided examples illustrating different approaches to making certain HTML elements more accessible. In the case of tables, the document presented the different approaches that can be used to make tables more accessible. It also illustrated the difference between a 'simple' table and a 'complex' table.

The problem

In many cases the Techniques document did not indicate the practical differences between different approaches, or which approach should be used under which circumstances. These limitations of the examples provided meant that some participants had difficulty in applying the examples to their task. For example, two participants found that the examples did not provide enough information. One participant wanted to associate one header cell with another header cell in a table, the other wanted to associate headers horizontally as well as vertically, but the example did not indicate whether either of these approaches were possible.

Solution suggested by researcher

The researcher believes that the provision of further examples could solve this problem. Additional examples could show the tags being used in a variety of contexts, which may improve page authors’ ability to apply the examples to their own work.

3.5.3.3.6 Variables

Background

Some of the examples provided in the Techniques document involve the use of variables. For example, the allocation of keys for the Accesskey tag, and the allocation of IDs for table Headers. In some cases it is not clear whether the choice of variable in an example is important.
The problem

Some participants were observed to be unsure whether they should copy the variables in an example verbatim, or whether they could choose their own variables. For example, one of the examples of IDs and Headers for tables used variables ‘a2’-’a8’, but did not use ‘a1’. It is not clear whether the absence of ‘a1’ is deliberate, or if it only applies in that specific case. It is also unclear whether the variables have to begin with ‘a’, or whether the allocation of the variables is down to personal choice.

3.5.3.4 Navigation in and between documents

Background

When using the Guidelines online, navigation involves reading text and following links to other sections (within either the Guidelines document or the Techniques document). For example, reading the table of contents (which is a list of links), then following a link to a particular section, reading that section, and then perhaps following another link to further supplementary information.

Participants needed to 'navigate' in the Guidelines and Techniques documents in order to find information regarding a particular HTML element, such as images or frames. Participants also needed to navigate when they required supplementary information, for example, once they had found a relevant guideline, they needed to find information about when to use it and how to implement it. Finally, participants also needed to navigate when they wished to find information that they had previously found again, for example if they remembered seeing a guideline regarding the presentation of links and needed to find it again.

While looking for information in the Guidelines, participants were observed to perform two different 'finding' tasks: firstly finding information on a particular topic for the first time, and secondly finding information again later. In performing these tasks participants were observed to have difficulties that the researcher believes may have been caused by 'navigational' problems.
3.5.3.4.1 Finding information in the table of contents

The problems
Participants were observed to have difficulties with finding information on a particular topic for the first time. The source of this problem may have been located in the design of the table of contents. Most participants began the task of finding information by reading the table of contents. The table of contents was a list of links the text of which gave each Guideline in full. However, the text of the links in the table of contents did not necessarily mention the HTML element concerned, such as images or tables.

Two participants were observed to use the ‘Find’ command to search the Guidelines document because the table of contents did not list what they were looking for. One found the relevant section by searching for ‘image’ and ‘imagemap’. However, this method was not successful for the other participant. They could not find the section that mentioned Accesskey because they did not recall the exact spelling and searched for ‘access key’ (including a space). Another participant commented that it was “really, really hard to find what you need” (JK).

Apart from the difficulties participants experienced in finding the information they required, the Guidelines themselves state that the text of links should be short phrases, rather than whole sentences. One participant observed that the document was breaking its own Guideline.

Solutions suggested by participants
Participants stated that they wanted the table of contents to list all the elements that they may be working on, such as tables, frames etc, rather than just providing a list of the Guidelines themselves.

One participant pointed out that the Guidelines recommend the use of a site map because such a device may help navigation within and between the documents. They suggested this could be provided for the Guidelines and Techniques documents.
Discussion

This latter suggestion is an interesting one. Site maps are often visual in nature and if a site map was included for the Guidelines and Techniques documents it would have to be accessible. The concept of an accessible site map has not been addressed by the Guidelines.

3.5.3.4.2 Finding again

The problems

Participants were also observed to have difficulties in 'finding information again'. Many spent much time following links from one place to another and expressed frustration at not being able to find information that they had found before. For example, one participant recalled that the Guidelines mentioned the text that should be used on buttons, particularly the Submit / Reset buttons on forms. Although they kept looking for the relevant section, they could not find it again.

There seemed to be several related causes to these problems. Firstly, participants often appeared to be unsure about where the links in the Guidelines and Techniques documents went to, or what to expect at the other end of a link. Secondly, they often appeared to be unsure about which document they were currently viewing, and thirdly, they seemed to be confused about where the current section was located in relation to other sections they had read. These three issues are discussed in the following sections.

3.5.3.4.2.1 Link text

The text of the links does not always inform the user where the link goes, or what to expect at the other end. Links go to several different types of destination, such as examples, related topics, additional information, or definitions. Each of these types of destination is located in different places, such as in another section of the Guidelines, in an appendix, in the Techniques document, or in the HTML 4.0 specification. However, in many cases neither the type of destination nor the location of the destination is indicated in the link text.
The researcher also found that some links that had the same text did not go to the same place. An example from an earlier version of the Guidelines is that in Checkpoint 1.3 the link "provide a text equivalent" went to examples of imagemaps, whereas the link "provide a text equivalent" in Checkpoint 2.2 went to examples for images. This was actually a contravention of the Guideline that states that links with the same text should go to the same destination.

3.5.3.4.2.2 Destination of link

One effect of participants not knowing which document a link went to, was that they often lost track of which document they were currently viewing. Many were observed to follow links, then find they had not arrived where they expected, and try to scroll back to the link they followed. If the link had taken them to another document they would not be able to scroll to the link. It was almost as if the transition between the Guidelines and Techniques documents was too seamless.

3.5.3.4.2.3 URLs

For one participant navigational problems were caused by the fact that the URLs were not fully visible in the status line the browser window (the window was too small, even though it was maximised). When the participant moved between documents it was the second part of the URL that changed, but the participant could not see this. They commented that this was interfering with their usual strategy for keeping track of where they were. The URLs of the documents are currently very complex. They specify several directories, few of which have recognisable names, and include non-alphanumeric characters. In 1999 the URLs of the Guidelines and Techniques documents were:

http://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505
http://www.w3.org/TR/1999/WAI-WEBCONTENT-TECHS-19990505

These specify that they are Technical Recommendations of the W3C and the dates they were created. Recently these URLs have been shortened: the year and the date have been removed, which makes them easier to read.

http://www.w3.org/TR/WAI-WEBCONTENT/
http://www.w3.org/TR/WAI-WEBCONTENT-TECHS/
3.5.3.4.3 *Grouping of information*

*The problem*

The information about any one particular HTML element is spread amongst various sections of the Guidelines and Techniques documents. Two participants expressed a desire to have all information on one topic in the same place.

*Discussion*

This perceived requirement to have all the information on one topic in the same place may be a symptom of the navigational problems described above. If participants were having difficulty finding information it is not surprising that they expressed this need. Because it is difficult to identify the reason for this need it is also difficult to identify the most appropriate solution.

The researcher believes that it would not be appropriate to re-organise the Guidelines documents on the basis of the opinions of two people without further evaluation. However, such a re-organisation may not be necessary if the navigation of the documents is improved, as described above, such as the re-location and re-wording of the link text.

3.5.3.4.4 *Sections of the Guidelines*

*The problem*

Participants found that it was not always easy to distinguish between different sections of the Guidelines and Techniques documents. This seemed to make navigation within and between the documents more difficult.

*Solution suggested by participant*

One participant suggested that colour could be used to help distinguish between sections. However, a problem with this approach is that it needs to be made accessible itself: any mark-up used to create visual cues needs to be replicated non-visually, for blind people and users of browsers that do not display colours. This suggestion is definitely worth further investigation and evaluation.
Discussion

The current version of the Guidelines includes coloured backgrounds behind the text of each guideline and the structure is also indicated by the use of section separators. The researcher believes that further improvements could be made with additional use of 'white space'.

3.5.3.4.5 Section Numbering

Background

The Guidelines and Techniques documents have numbered sections. The Guidelines have three main sections: A, B, and C, which are then divided into numbered sections and sub-sections: 1, 1.2 etc. The Techniques have five main sections numbered 1, 2 ... 5, with numbered sub-sections.

The problem

The section numbers in each documents do not correspond. For example, section A, 1.1 of the Guidelines refers to text equivalents of all images, whereas section 1, 1.1 of the Techniques refers to Accessibility themes for HTML. One participant found this confusing because they expected the sections in each document to correspond.

Solution suggested by participant

The participant suggested that the section numbering could be standardised across all the documents.

3.5.3.4.6 Missed information

The problems

An important implication of page authors not being able to navigate effectively and efficiently between the Guidelines and Techniques documents, and their various sections, is that important information may be missed. There were several important items in the Guidelines and Techniques documents that many participants were observed to miss while performing their task. For example,
• The Guidelines contain many links to the Techniques document but only one has the text 'techniques', the others have different text. One participant was searching for a specific link to the Techniques but could not find one.
• Two participants missed the statement in the Guidelines that Longdesc is intended for use with 'important' images, and used this tag with images that probably did not require Longdesc.
• One participant provided alternative text links for an imagemap and separated them with vertical bars. (This is considered to be good practice and is recommended in the Guidelines.) However, the participant was concerned that the vertical bars may not be read by a screenreader, and decided to remove them. They had not read the relevant section on screenreaders in the Guidelines and therefore did not realise that vertical bars are recommended.

This problem of missing information is also evident from the fact that some of the participants suggested improvements to the Guidelines and Techniques documents that already exist. For example, one participant suggested that an explanation of the way in which screenreaders function should be provided, although a brief description is provided.

The problems described above probably have various specific causes, but this type of problem may be reduced if the navigation and clarity of the whole document is improved.

3.5.3.5 Overall structure of Guidelines

The problems
Another possible reason that some participants missed important information could be that the way in which the Guidelines were presented meant that the participants found it difficult to read the documents. Many were observed to skim read the Guidelines and Techniques documents, rather than read them in detail. (Although they had time to read the documents prior to attending for the experiment, they needed to re-read much of what they had read before.) For example, bold text was used to highlight various parts of the text, such as to indicate the groups of users affected by each guideline, and to emphasise particular points. The bold text and links embedded in the paragraphs made it very difficult to read a whole paragraph.
Solutions suggested by participants

3.5.3.5.1 One large document
One participant suggested that the Guidelines, Techniques, and Checklist could be combined into one document, with the Checklist at the beginning. This would mean that a page author could read the whole document from start to finish, beginning with the Checklist, which provides an overview of the issues. If the page author then required more details they could read the Guidelines section. If they then required further information on how to implement the Guidelines this could be found in the Techniques section.

3.5.3.5.2 Smaller documents
In contrast to the previous suggestion, another participant suggested that each guideline should be in its own file with links to the previous and next guidelines. The participant believed that this would enable a page author to concentrate on one issue at a time, and the pages would have the added benefit of loading faster than the whole Guidelines or Techniques documents.

3.5.3.5.3 Bullet points
Another suggestion was that the Guidelines and Techniques documents might be easier to read if they were presented in shorter phrases with more bullet points.

This is a difficult balance to maintain: on the one hand the documents need to be easy to read, but on the other hand they need to provide detailed information that could not necessarily be presented in bullet points.

3.5.3.5.4 Printable documents
A further suggestion was that a ‘printable’ version of the Guidelines and Techniques documents could be provided so that they did not have to be read on-line.

This seems to be a reasonable suggestion, especially since some research has found that on-line documents can be difficult to read (Dillon, 1994). It is an interesting issue because the Guidelines and Techniques were designed to be a hypertext system, with links between related sections. The documents can be printed from the Web and are available in zipped, postscript and PDF formats. However, the relationships
between sections, i.e. the links, are lost in such formats. In order to create a printable version that could be used effectively many changes would have to be made, including the replacement of links with references.

3.5.3.5.5 Summary of Guidelines
It was suggested that the Guidelines could be summarised in a very short document, which would be aimed at more experienced page authors. The rationale for this was that an experienced author might read in such a document that, for example, ALT text is required for images, and know how to implement this without reading the Guidelines.

The Education and Outreach group of the WAI has since developed a 'quick tips' card\textsuperscript{50} that summarises the main points of the Guidelines.

3.5.3.6 Cost vs. benefit
Four participants stated that were trying to compare the amount of work it would take to make a page more accessible compared with the small numbers of people who would benefit from the work.

One participant was reluctant to cater for the minority of people who do not use either IE or NS. The other three were concerned about making the table more accessible. One did not know whether they should focus on designing the table to improve access for blind people, or whether they should also be considering the usability for sighted people. The other two decided that it would be too much work to use any of the suggested approaches for tables. One of these participants compromised on redesigning the headers of the table to make them easier for sighted people to read, and hoped that they would also be more accessible to blind people.

\textsuperscript{50} WAI Quicktips card available at: http://www.w3.org/WAI/References/QuickTips/ (checked Nov 2000).
3.5.3.7 Intended audience

Background
The Guidelines currently appear to be aimed at page authors with much HTML knowledge and experience. The language is therefore quite technical. The document is quite detailed which may assist less experienced page authors, but hinder more experienced authors. The language is technical and perhaps more suitable for experienced authors, but there is perhaps insufficient support for using HTML syntax for those less experienced.

The problem
The less experienced participants thought that the language was too technical, although they appeared to be able to follow it and implement the advice. A more experienced participant thought that the Techniques document was too basic and repetitive. They thought that experienced authors would find that the Techniques document provided too much detail, was too long, and that the Guidelines alone provided enough information.

Solution suggested by participant
There was a suggestion that the Guidelines document should specify who its intended audience is.

3.5.3.7.1 Users of authoring tools
One section of the Guidelines' potential audience who may not be currently catered for are those page authors who use an authoring tool rather than a text editor to create Web pages. Four of the twelve participants used an authoring tool (three used FrontPage, the other used NS and DidaPro). These tools do not yet support page authors in implementing accessibility tags, such as Accesskey or tags for use with tables. These participants used the tools' HTML editors to incorporate the accessibility tags into their pages, but appeared to find this task more difficult than those participants who used text editors. They were less certain of the structure and syntax of the tags.
The WAI has a group working on another set of Guidelines for the developers of authoring tools who wish to support the creation of accessible pages. In the meantime the WAI could consider how the Web Content Guidelines can support users of tools, who may not be as familiar with the structure and syntax of HTML tags as page authors who produce their own code.

3.5.3.8 Further information required

From the observations of the participants and their comments, several areas were identified in which further information was required, including information on disability and assistive technology, and information to support decision-making.

3.5.3.8.1 Disability and assistive technology

The problem

The Guidelines currently provide some information about disability and assistive technology, but it may not be sufficient. From the observations of the participants it is clear that some required further information about the way in which people with disabilities might interact with a page. For example, one participant asked whether a blind user would be making notes on the contents of a table as they were listening to it in synthesised speech. The researcher said that it would be possible for a blind person to make notes, but that it was unlikely unless the information was of great interest. This participant also thought that a screenreader would automatically indicate to the user that certain sections of text were displayed differently, for example it would distinguish between the main headers of a table and the sub-headers. A screenreader might allow the user to set it up to make such distinctions, but it is unlikely that a screenreader do this by default. The participant justified the fact that they had used italics for the navigation links by saying that they thought the links would be more visible to a partially sighted person. (It is unlikely that partially sighted users would find such text easier to read than non-italicised text.)
Solutions suggested by participants

Several participants stated that they would like further information regarding disability and assistive technology. The topics they identified included the following.

- How blind people use the Web.
- The percentage of the population that is disabled.
- How screenreaders present certain elements, such as Frames, text in columns, or a D-link.
- How screenreaders would present new tags, such as Accesskey, Headers and IDs in tables, and Longdesc.
- If Longdesc creates a link to a file, how would a user follow the link if it were not visible.
- How a user would select a link if they were not using a mouse.
- Whether there is a braille format for tables.
- How would a user know which keys were to be used to use an Accesskey.

Participants believed that if they had more information about the way in which users of assistive technology would interact with tags, it would help them implement the tags in an appropriate manner.

3.5.3.8.2 Explanation of new concepts

The problem

Participants expressed a need for further definitions of concepts mentioned in the Guidelines. For example, one participant did not know what was meant by ‘putting a table into a ‘linear format’, or how a browser would ‘filter’ table ‘categories’. Two participants required clarification of whether an Accesskey is used to jump to a link or to follow a link, and of how default text placed in a text-entry box can be helpful.

Discussion

The WAI should perhaps consider how to explain these concepts to people for whom they are new concepts. It is possible that the participants did not understand these concepts because they were relatively inexperienced page authors. Future research involving professional page authors should discover whether such participants have the same difficulties with these concepts.
3.5.3.8.3 Support in making decisions

The problem

In several different contexts the participants were observed to have difficulty in deciding which of the available options they should implement. For example, those working with a table needed to decide whether the table was ‘complex’ or ‘simple’, and whether they should use Scope, Headers or Categories for tables. The Guidelines provide an example of a ‘complex’ table, but do not state why it is defined as ‘complex’. Those working with forms needed to decide which keys should be allocated for use as Accesskeys, and how long an option list should before Option Group should be used. The Guidelines do not provide information about the allocation of keys, or define how long a ‘long’ list might be.

Discussion

In some cases the information they needed was available in the Guidelines or Techniques documents, but the participants had not found it. In other cases the information is not available, but could be provided, perhaps in the Glossary, or in the examples.

3.5.4 Positive comments on Guidelines

While the participants made lots of criticisms of the Guidelines, they also made some positive comments. Several participants said that the Guidelines were useful for raising awareness of the issues regarding the accessibility of the Web. Some found they had learned about new tags and found the examples useful. One participant particularly liked the suggestion to think about how a page would be presented via the telephone in order to simulate how a screenreader might present it.

3.5.5 Discussion with WAI on results

The researcher reported the findings of the current study to the WCA Guidelines working group. The findings were reported in two stages. The initial findings were reported in November 1998, followed by more detailed reports in March 1999. On both occasions the researcher began new ‘threads’ on the discussion list for each
The discussions are archived on the WAI Web site\textsuperscript{51}. The reports sent to the discussion list contained brief descriptions of the difficulties the participants had experienced, and suggestions made by the participants or by the researcher. In some cases the researcher did not agree with participants’ suggestions, or saw the difficulties with implementing them. Despite this the suggestions were reported to the working group because members of the group may have had different responses to those of the researcher.

An overview of the issues raised with the WCA Guidelines Working Group, the suggestions made by the researcher, and summaries of the discussion of each issue within the group can be found in Appendix 5.

### 3.5.5.1 General response to the reports

The general response to the reports sent to the WCAG Working Group was that the results of the study were “very helpful” and the reports contained “valuable observations” and “great comments” (Guidelines editors).

One respondent made the following comments on the results of the study.

- Acceptance of the comments on the organisation and layout of the document should be tempered in view of the fact that the participants spent an average of one hour reading the Guidelines and Techniques documents: this was too brief a reading. For page authors who believe that accessibility is just one of the technical requirements of a page, a quick look at the Checklist would be sufficient. On the other hand page authors who have learnt ‘presentational tricks’ to create pages that look fine on a particular browser will need to ‘unlearn’ much of what they have learnt in order to incorporate accessibility issues in their designs. This ‘re-learning’ could not be done in three hours. The situation is not helped by the fact that there are very few authoring tools that support accessibility.

The participants spent as much time as they needed to read the documents in advance and to perform the task: they were not given any limit. The researcher agrees that it would take more than three hours to identify and correct all the accessibility problems in a Web site. However, in this case the participants were just focusing on one accessibility problem each, and took as much time as they needed. It is believed

\textsuperscript{51} Archive of WAI Web Content Working Group discussion list available at: http://lists.w3.org/Archives/Public/w3c-wai-gl/ (checked Nov 2000)
that many of the difficulties experienced by the participants reflect the difficulties that professional Web designers might experience in using the Guidelines.

A number of other issues were raised by the researcher concerning minor editorial problems. Many of the researcher’s suggestions made were incorporated into later versions of the Guidelines and Techniques documents.

### 3.6 Discussion on the usability of the Guidelines

The participants in this study encountered and reported difficulties in using the Guidelines. The main issues raised are the use of new tags, the examples provided, navigation within and between the documents, the intended audience of the Guidelines, and further information required in order to implement the Guidelines effectively. The researcher raised these issues with the WCA Guidelines Working Group and got various responses. These issues and the Working Group’s response are discussed further in the following sections. This is followed by a discussion of some limitations to the method used in this study, which have been identified by the researcher.

#### 3.6.1 New tags

There are several problems associated with the new tags recommended in the Guidelines. Firstly, page authors need further support in implementing the correct syntax because the examples in the Techniques document are not sufficient. Also some of the new tags are not described in mainstream books on HTML 4.0.

Secondly, some of the tags are not supported by browsers such as IE and NS. This means that page authors do not know how a user would interact with the tags, and cannot test them. Although the Guidelines suggest that the most appropriate way to test pages is to use a validator, this point is not prominent. Of the two participants who tried to validate their work, one found the W3C validator difficult to use, the other found their usual validator did not recognise the new tags.

The researcher believes that an important issue is that page authors may need support in making their pages more accessible and the Guidelines need to remove the potential for any difficulties authors may experience. While some new devices may
support the new tags, these devices are not necessarily being used by people with
disabilities, in particular blind people. The Guidelines are intended to improve
accessibility for people with disabilities and page authors are encouraged to use the
relevant tags because they are also of benefit to users of other devices, such as
mobile or in-car devices. In the short term, the fact that these latter devices support
the new tags should not distract authors from the fact that the primary purpose of
implementing the advice of the Guidelines is to make sites more accessible to people
with disabilities.

The researcher is not convinced that page authors should be encouraged to
implement tags that are not supported without informing them this is the case. It
does not seem to be a good idea to shield page authors from the fact that some tags
are not supported. All of the participants who used tags and later found that they
were not supported said that the Guidelines should have informed them that this was
the case. The Guidelines need to be clear that they are recommending the use of tags
that are not yet supported, but which are hoped to improve accessibility in the future.
Page authors would then have a choice of whether to implement them for future use
or to delay implementation until they are supported. Despite the fact that the WAI
are developing guidelines for developers of browsers and other user agents, there is
currently no guarantee that future browsers and assistive technologies will support
these tags.

The Guidelines or Techniques documents could state which tags are not currently
supported by the main browsers and providing a statement such as the following.
When testing your implementation of this tag, bear in mind that it may not be
supported by, or available to, your browser and may only be supported by, or
available to, ‘special’ devices that are currently under development. Use an HTML
4.0 validator, such as that of the W3C, to check the implementation.

Since the discussions with the WAI, some information about browser support of tags
has been included in more recent versions of the Guidelines. However, this is not
provided for all tags that are not yet supported. The researcher believes that it is
necessary to provide authors with this information to prevent them from removing
potentially useful tags, and to prevent them from believing they have made a mistake.

The researcher believes that it would not necessarily be appropriate for the Techniques document to illustrate how CSS might be rendered. CSS is a whole topic in its own right and it would probably be most appropriate for the Techniques document to provide links to sources of information external to the Guidelines on CSS and how it is rendered. If this information does not already exist, it could be provided by the WAI, but not within the Guidelines or Techniques documents.

3.6.2 Priorities

The WCA Working Group who developed that WCA Guidelines have attempted to overcome the problem identified by Mosier & Smith (1986) that designers need assistance in choosing between guidelines. In the WCA Guidelines priorities are assigned to each checkpoint, and clear rationales are provided. The WCA Working Group has ongoing discussions about whether the priorities apply in all possible circumstances in which individual guidelines might be implemented. However, the results of this current study suggest that the priorities do not necessarily overcome the problem entirely.

The high priority given to new tags that are not yet supported was an important issue for one participant. They believed it was unreasonable to require a page author to use an unsupported tag in order to achieve the highest conformance to the Guidelines. The researcher believes there is a key difference between, a) tags being supported by browsers used by people with disabilities although not supported by mainstream browsers, and b) tags not currently being supported by any browser. The Guidelines should acknowledge that some tags that are priority 1 are not currently supported by most browsers. A page author could spend time implementing a tag and then find that no-one can currently benefit from its presence. It could be argued that if a page author implements such a tag now, it will benefit users in the future, but the researcher believes that page authors should have a choice about this.
The WAI implemented a partial solution to this problem. Several Checkpoints now contain the phrase ‘until user agents’. For example,

Checkpoint 1.3: “Until user agents can automatically read aloud the text equivalent of a visual track, provide an auditory description of the important information of the visual track of a multimedia presentation. [Priority 1]” (WAI).

Checkpoints that include this phrase contain links to the following entry in the Glossary.

“Until user agents...
In most of the checkpoints, content developers are asked to ensure the accessibility of their pages and sites. However, there are accessibility needs that would be more appropriately met by user agents (including assistive technologies). As of the publication of this document, not all agents or assistive technologies provide the accessibility control users require...
Checkpoints that contain the phrase "until user agents..." require developers to provide additional support for accessibility until most user agents readily available to their audience include the accessibility features.” (WAI, op. cit.).

Currently the phrase ‘until user agents’ is used in approximately twelve checkpoints. The researcher believes the phrase should be used in all the checkpoints that refer to tags that are not yet widely supported.

The Guidelines document does not advise page authors on how they might find out which user agents support a given tag. However, it does provide links to a ‘User Agent Support’ page53, and the following note.

“The W3C WAI Web site provides information about user agent support for accessibility features. Content developers are encouraged to consult this page regularly for updated information.” (WAI, op. cit.).

An entry in the Glossary regarding the ‘User Agent Support’ page states that “This page documents known support by user agents (including assistive technologies) of some accessibility features listed in this document” (WAI, op. cit.). However, at the time of writing the User Agent Support page contained only a couple of entries. This should be extended and become a definitive resource on tag support by mainstream and special browsers.

52 Web Content Accessibility Guidelines 1.0 available at: http://www.w3.org/TR/WAI-WEBCONTENT/ (checked Dec 2000).
The priority of unsupported tags is obviously an extension of the general discussion on the recommendation of unsupported tags. From the WAI’s perspective the priority of tags such as Headers and IDs for tables cannot be lowered because potential future accessibility would be lost. The researcher believes that page authors should be informed that some priority 1 tags are not yet supported by any browser, and will not necessarily be supported in the future. Also, interim measures, such as the use of paragraph breaks in tables should be re-introduced into the Guidelines. Otherwise, some elements, such as tables, will still be inaccessible even if new tags such as Headers and IDs are implemented because assistive technology does not support them.

3.6.3 Examples

Previous studies have found that the examples provided in guidelines or standards, or that accompany such documents, play a key role in the use of guidelines and standards. Tetzlaff & Schwartz (1991) found that designers preferred to learn from examples than from text. They also found that some designers used examples to the exclusion of the guidelines text, which meant that they missed important information. One of Tetzlaff & Schwartz’s conclusions was that examples must be unambiguous, otherwise designers may implement aspects of examples that are not relevant to the current design problem. Thovtrup & Nielsen (1991) found that examples had as much impact on designers as a standard.

The participants in the current study also carefully read the examples provided in the Techniques document and often copied the syntax exactly. They were also observed to re-read the examples several times, but in this case it is thought that they did this because the presentation of the examples made them difficult to assimilate on the first reading. Like the participants in Tetzlaff & Schwartz’s study they also said that more examples would be useful and could illustrate more clearly how the HTML elements might be rendered. Some suggested that further examples could take the form of screendumps, which would be more illustrative.

In particular participants wanted the examples to show how the new tags would be rendered; visually, in synthesised speech, and special devices. They wanted to know
how a user would interact with the tags. It would be difficult to provide illustrations of how all tags would be rendered by all browsers because even IE and NS display some tags differently, and tags such as ALT text are presented in different ways. The researcher believes that there is an argument for providing illustrations of the use of some tags. This particularly applies to those tags that might be regarded as 'special', such as Longdesc, Accesskey, and Headers and IDs. This would, at least, alert page authors to expect tags to be rendered in different ways. If some examples are provided in text, a page author could read how a tag would be presented in synthesised speech. It is important to recognise that page authors who know nothing of accessibility and the different devices used by people with disabilities will be on a steep learning curve and need to start with the basics. It is also important to indicate to page authors the circumstances under which each example should be applied, and to make the variables as transparent as possible. The fact that page authors may follow examples without reading the associated Guidelines and Techniques, or may follow the first example they come to need to be considered.

From the WAI's perspective it is not always possible to provide examples of how tags are rendered in different browsers because it is not yet known how some new tags will be presented in the future. The researcher was surprised to learn this because it was assumed that it was known how tags will be rendered.

The Guidelines editors did take up the suggestion regarding variables in examples, and liked the suggestion of listing examples to assist page authors in selecting the most appropriate example. At the time of writing this latter suggested had not been implemented.

3.6.4 Navigation

It is interesting that Thovtrup & Nielsen (1991) recommended the use of hypertext for the presentation of a guidelines document. The researcher believes that the results of this study suggest that the process of creating guidelines in hypertext is not a trivial matter, and that enabling the reader to jump between different sections does not necessarily support them in finding what they need.
Participants needed to navigate the documents in different ways, depending on the current task. They needed to find information initially, follow links to different sections, or find a section they had read before. Several aspects of the documents made this navigation difficult. The text of the links in the table of contents did not include the names of the elements the participants were working with, which meant some had to resort to using the ‘Find on page’ command. It was also difficult to distinguish between the Guidelines and Techniques documents and between the sections of these documents. This problem was compounded by the text of links not indicating their destination, and even some links with the same text going to different destinations. That some participants suggested structural changes to the documents is perhaps evidence of these navigational problems. Further evidence is that several participants missed important information on the interpretation and implementation of the Guidelines in the documents. The researcher believes that the best solution to the problem of the participants not being able to find the topic they need is to rephrase the table of contents and to rephrase the text of many links.

The links embedded in the paragraphs of each Checkpoint could be moved to the end of the paragraph, so that the Checkpoint can be read without the distraction of links. The text of the links could include the type of destination, such as an example or a definition, and could also include the location, such as in an appendix, or in the Techniques document. For example,

Paragraph on images “use ALT text, or use Longdesc for ‘important’ images”
... (containing no links)
<Link> Appendix: example of use of ALT.
<Link> Techniques: definition of ‘important’.
<Link> Techniques: example of use of Longdesc.
<Link> Related guideline: imagemaps.

The length and complexity of the URLs of the documents meant that it was not possible to use the display of the URLs in the status bar of the participants’ browsers to keep track of the current document. The complexity of the URLs is inherited by the WAI from the W3C and is required to include several elements. However, the researcher believes it is worth considering an alternative. If the URLs were less complex it might be easier for page authors to keep track of the document they are currently reading. More suitable URLs might be:
These URLs provide the necessary information that the Guidelines belong to the WAI, which in turn is part of the W3C, that these are the ‘Content’ Guidelines and Techniques, as opposed to the guidelines for User Agents or Authoring Tools. The proposed URLs do not include the date the documents were created because this is stated at the top of each document.

The Working Group responded positively to the researcher’s suggestions regarding navigation, such as the wording and location of links. Also the links between the Guidelines and Techniques documents are now relative, so the URLs displayed in the status bar of browsers should be shorter.

3.6.5 Audience

Some participants found the Guidelines too detailed or too technical, whereas others found the Guidelines did not provide sufficient information, or were too repetitive. Some of the participants who used an authoring tool found these tools did not support the implementation of new tags. The participants were not confident in using the tools’ HTML editors to implement the tags. There was no response from the Working Group to the suggestion that the Guidelines could specify the intended audience, although the W3C fact sheet\textsuperscript{54} on the Guidelines states that the Guidelines are aimed at a variety of groups.

Mosier & Smith (1986) found that the guidelines they evaluated were read and used by a variety of groups. These groups differed in their perception of the usefulness of the guidelines. Mosier & Smith conclude that it is important to make guidelines of use to different groups, but do not suggest how this might be done.

The WAI have produced several different documents regarding accessibility. These include the Guidelines, factsheets and Quicktips cards, which are each aimed at

\textsuperscript{54} W3C Fact sheet on WCA Guidelines available at: http://www.w3.org/1999/05/WCAG-REC-fact.html (checked Nov 2000).
different audiences. The WAI Web site could state for which groups these
documents are intended.

Another issue relating to the intended audience of the Guidelines was raised in the
discussion with the Working Group regarding the readability of the Guidelines. A
respondent stated that those who do not read the entire Guidelines document should
not be catered for. Mosier & Smith (1986) and Tetzlaff & Schwartz (1991) found
that users of guidelines did not read the guidelines in the manner expected by the
guidelines authors, in particular some did not read the guidelines in full. The
researcher believes that it would be preferable to support page authors in using the
Guidelines in the ways that this study and previous studies have found, rather than
expecting the Guidelines to be read in a certain way and not taking responsibility
when page authors do not do this.

3.6.6 Further information required

Participants said they would like more information on, for example, disability and
assistive technology, to help them understand how people with disabilities use the
Web. They were quite specific about the details they required, such as the
percentage of the population that is disabled, and how a user would select a link
without using a mouse. Some of the Working Group respondents believe it is not
important to provide information, such as statistics, because the Guidelines are
intended to inform page authors how to design for accessibility. The W3C’s fact
sheet\(^{55}\) on the WCA Guidelines contains information regarding the numbers of
people with disabilities across the world, although it does not reference its sources.
The researcher believes it would be sufficient to provide a reference in the
Guidelines to the factsheet.

Participants were also found to need further explanation of new concepts and further
support in making decisions. In some cases the relevant information was contained
in the Guidelines or Techniques, but the participants did not find it. The
implementation of the researcher’s suggestions for improving navigation may

\(^{55}\) W3C Fact sheet on WCA Guidelines available at: http://www.w3.org/1999/05/WCAG-REC-
fact.html (checked Nov 2000).
support page authors in finding the information they require and knowing it is available. However, the Working Group could also consider how to provide further explanations and support for decision-making.

3.6.7 Positive comments

On a more positive note, some participants said they had learnt much about accessibility by using the Guidelines and that some ideas of how to think about accessibility were particularly useful. The Guidelines prompted some participants to consider the costs and benefits of designing accessible Web pages.

3.6.8 Accessibility of the Guidelines

Some of the issues highlighted by the participants raise the issue of the accessibility of the Guidelines and Techniques document themselves, for example, consideration of how examples should be presented, and whether background colours should be used. The WAI want to produce documents that assist designers as much as possible, but they would not want these documents to be inaccessible. The researcher believes this is an issue that will have to be resolved in the future.

3.6.9 Limitations of method

With hindsight there are a number of aspects of the method that should be changed if this study were to be repeated. This includes aspects specific to the HTML elements, as well as more general aspects.

3.6.9.1 Participants

The participants of this study were relatively inexperienced Web designers, although some had been employed to design Web sites. These participants identified some fundamental problems with the Guidelines, including problems with navigation and limitations of the examples. If the study is repeated with professional Web designers they may identify the same problems, and further problems.
3.6.9.2 Topic

The content of the page containing a table and images given to the participants was an academic paper written by the author, who put it on the Web. A disadvantage of the paper being on a topic with which the participants were not familiar meant that it was sometimes difficult for them to make appropriate decisions. For example, it was difficult to provide ALT text that was meaningful, or to provide an appropriate caption for the table. It also made the table difficult to understand.

On the other hand it is likely that if these participants were to work as professional Web designers they would be asked to design pages containing information that they were not necessarily familiar with. Therefore this may have been quite a realistic situation.

3.6.9.3 Tables

The table provided may have been too complex. Participants had to deal with two levels of headers, which may have been too complicated, and made the examples in the Techniques document difficult to apply.

The table given to the participants had a label associated with it. It may have been preferable if the label was not provided. This would have meant that those who considered the use of a caption or summary would not have to consider how those tags related to the label.

3.6.9.4 Images

The images provided were all functional and already had labels. It would have been better if the purpose of the images was not described in the text and if they did not have labels. Several participants said that they did not know what to include in the ALT text because they did not want to repeat the information in the text or the labels. It would also have been better if some of the participants had been given decorative images or images that contained information, such charts.
3.6.9.5 Technical issues

Technical problems prevented some of the participants performing the task in the way that they preferred. For example, problems with WebSpeak prevented two participants from viewing their pages with this browser. Network problems prevented some participants from viewing their pages with the Lynx browser, and from using the Bobby checker. The one participant who used VI on Unix over the network found it to run very slowly. Early problems with the computer meant that it crashed when NS was run, but these problems were solved for later participants.

3.7 Conclusions on the usability of the guidelines

This study has identified many aspects of the design and content of the Guidelines and Techniques documents that caused difficulties for the participants. The researcher believes that if this study was repeated with professional page authors those participants would experience similar difficulties in navigation and following examples of tags they were not familiar with.

Some of the problems found are associated with the hypertext nature of the documents concerned. Mosier & Smith (1986) and Thovtrup & Nielsen (1991) concluded that an on-line database or a hypertext system would improve designers' ability to use guidelines or standards. The results of the current study suggest that converting guidelines to an electronic format does not necessarily solve the problems, and may introduce its own inherent problems.

Tetzlaff & Schwartz conclude that “much written material may go unread, misunderstood, and consequently unheeded. The longer and more complex the document the greater the risk” (1991, p333). They believe that the most effective way of communicating style requirements is for the style to be illustrated in examples and interactive demos, which would be facilitated by toolkits and supported by iterative usability testing. The role of guidelines would be to complement these other techniques, to describe those principles that may be difficult to infer from the examples.
The WCA Guidelines also form a long and complex document, particularly if the Techniques and Checklist documents are included. The Working Group have said that they expect people to read the entire documents, but perhaps need to heed Tetzlaff & Schwartz’s concern regarding the length and complexity of a guidelines document.

Tetzlaff & Schwartz also conclude that guidelines should explicitly state their scope and limitations and specify the conditions under which each component should be used. The WCA Guidelines state their scope, but do not perhaps address their limitations. They do state when and how each technique should be used, but this was not as clear as it might have been to some of the participants.

3.7.1 Conclusions on WAI discussion

The fact that the WAI is a program of the W3C certainly gives the program credence, which it may not have had if it were an academic or commercial organisation. However, this relationship also places restrictions on the WAI, which in turn have an impact on the content and the organisation of the Guidelines. For example, the system of URLs inherited from the W3C, and the fact that the documents have to be ‘timeless’ and cannot take account of changing support of tags by browsers.

A theme that runs through the researcher’s discussions with the WCAG Working Group is that when the researcher reported the experiences of the participants of the study, the discussion that followed focused on whether the participants used the Guidelines correctly, and whether they misunderstood them, rather than on how such difficulties can be overcome. A responsibility seems to be placed with page authors to understand and interpret the Guidelines, rather than the authors of the Guidelines having a responsibility to ensure that, as far as possible, page authors are supported in this process.

The literature suggests that designers may not read guidelines in the way that is expected. Tetzlaff & Schwartz (1991) observed that the participants in their study used two different strategies for incorporating the guidelines into their task. A minority began by reading the introductory material but the majority began the task
immediately and only consulted the guidelines when they felt it was relevant. This latter group found that they could not get very far without consulting the guidelines, but only read them until they found what they thought they required. Again this meant that they missed information that the authors of the guidelines intended them to read. Tetzlaff & Schwartz appear to support the researcher’s opinion that people may not read guidelines in the way they are intended to be read. The WAI do not seem to have considered this issue with regard to the WCA Guidelines.

3.7.2 Summary of changes made in response to results

The WCA Guidelines Working Group made a number of changes to the Guidelines, in particular the way in which the checkpoints and examples are presented, based on suggestions from the researcher, which were in turn based on the results of this study. These changes are listed below.

- Ensured that all links with the same text went to the same place.
- Reworded many links to make link-text more meaningful.
- Moved many links to end of checkpoint paragraphs, instead of being embedded.
- Included the words ‘Guidelines’ and ‘Techniques’ in links that go to these documents, and the word ‘Checkpoint’ in links that go to Checkpoints.
- Made the variables in the examples of Headers and IDs more transparent.
- Added a note regarding the problems with images being loaded from the cache when trying to test pages without loading images.

3.7.3 Summary of remaining issues

The researcher finds that there are still some issues that have not be addressed by the WAI regarding the content of the Guidelines. These are listed below.

- List of users affected to be placed at end of guideline, not mentioned within the paragraph.
- Use of screen shots to illustrate examples.
- Example illustrating concept of tables presented in a 'linear fashion'.
- Indication to page authors if recommended tags are currently supported by only one browser.
- Indication to page authors of whether recommended tags will be 'visible' in main browsers.
- Explanation and illustration of how 'special' tags will be rendered, i.e. the purpose of the tag and how the user might interact with it.
- Provision of guidance for using the W3C validator, e.g. what the page author will be expected to provide (this may be more relevant for the validator page).
3.7.4 Recommendations from literature for WCAG

There are several findings in the literature on the use of guidelines that are relevant to the WCAG Guidelines. These are listed below. There are some ideas that have already been addressed by the WAI, either in the Guidelines or in related documents, and these are marked with a ‘*’.

From Tetzlaff & Schwartz (1991):

- Any changes made to the Guidelines in the future, particularly those that are more subtle, could be indicated to people who have previously used the Guidelines. This may avoid the problem of those who are familiar with the document missing subtle changes.
- The importance of the role of examples should be considered, and the existing examples could be further improved to remove any ambiguity, particularly in the use of variables.
- The impact of the length and complexity of the Guidelines and their associated documents should be considered when expecting designers to read all the documents in full.
- People will not necessarily read the Guidelines document and the associated documents in full, despite the intentions of the Guidelines’ authors.
- The fact that designers may find new concepts difficult to understand and implement should be considered.

From Mosier & Smith (1986):

- It should be recognised that different types of readers (designers, managers, students) will read the guidelines in different ways, and need to be able to use them in different ways. *
- Designers require guidelines to be given priorities so that they can prioritise their implementation. *
- It can be difficult to apply general guidelines to specific applications, and to generalise from very specific guidelines for a particular situation.

From Thovtrup & Nielsen (1991):

- The use of hypertext may be a useful way of enabling readers to navigate through a guidelines document. *
- The provision of a checklist supports designers in producing compliant designs. *
4 CHAPTER 4 – EVALUATION OF WEB PAGES BY BLIND AND PARTIALLY SIGHTED PEOPLE

4.1 Introduction ...........................................................................................................133
  4.1.1 Key to abbreviations used in this chapter.......................................................133

4.2 Method ......................................................................................................................134
  4.2.1 Design ...............................................................................................................134
  4.2.2 Participants.........................................................................................................134
    4.2.2.1 Recruitment ...............................................................................................135
    4.2.2.2 Sight status ...............................................................................................135
    4.2.2.3 Experience of using the Web ..................................................................135
    4.2.2.4 Monthly Web usage .................................................................................136
    4.2.2.5 Browsers ..................................................................................................136
    4.2.2.6 Screenreaders and speech synthesers ....................................................136
    4.2.2.7 Braille displays ........................................................................................137
    4.2.2.8 Operating system .....................................................................................137
    4.2.2.9 Common combinations of software .......................................................138
  4.2.3 Materials ..........................................................................................................138
    4.2.3.1 The Web site ............................................................................................138

4.3 Procedure ..................................................................................................................141
  4.3.1 The set of tasks ...............................................................................................141
  4.3.2 Tasks performed by individual participants ................................................142
  4.3.3 Questionnaire ..................................................................................................142

4.4 Tables .......................................................................................................................143
  4.4.1 Overview of differences between examples of tables ..................................143
  4.4.2 Results on Tables ............................................................................................145
    4.4.2.1 Accessibility of Tables .............................................................................145
    4.4.2.2 Answers to the set question ....................................................................147
    4.4.2.3 Uncertainty ...............................................................................................149
    4.4.2.4 Incorrect answers .....................................................................................149
    4.4.2.5 General comments on tables ................................................................150
    4.4.2.6 Comments on specific aspects of the tables ...........................................152
    4.4.2.7 Access to tables for partially sighted participants ................................158
    4.4.2.8 Suggestions for improvements ...............................................................158

4.5 Forms .........................................................................................................................158
  4.5.1 Overview of differences between forms ..........................................................158
  4.5.2 Results on Forms .............................................................................................159
    4.5.2.1 Accessibility of Forms .............................................................................159
    4.5.2.2 Factors affecting accessibility .................................................................160
    4.5.2.3 Comments on the accessibility of form controls ....................................160
4.5.2.4 Comments on accessibility features .............................................162
4.5.2.5 General comments .................................................................163
4.5.2.6 Access to forms for partially sighted participants ......................163

4.6 Frames ......................................................................................164
4.6.1 Overview of differences between frames .....................................164
4.6.2 Results on frames .................................................................165
4.6.2.1 Accessibility of Frames .......................................................165
4.6.2.2 Comments on specific aspects of frames ...............................166
4.6.2.3 General comments on frames ..............................................167
4.6.2.4 Access for partially sighted participants to frames ................169

4.7 Imagemaps ..............................................................................170
4.7.1 Results on Imagemaps .............................................................171
4.7.1.1 Accessibility of imagemaps .................................................171
4.7.1.2 Comments on imagemaps ..................................................172
4.7.1.3 Suggestion .................................................................173

4.8 Headers .....................................................................................173
4.8.1 Overview of differences between headers ..................................174
4.8.2 Results on Headers ...............................................................174
4.8.2.1 Accessibility of Headers ....................................................174
4.8.2.2 Comments on headers ......................................................175
4.8.2.3 Comments from partially sighted participants .....................177

4.9 References ...............................................................................178
4.9.1 Overview of differences between references ............................178
4.9.2 Results on References ............................................................178
4.9.2.1 Accessibility of References ................................................178
4.9.2.2 General comments on references ......................................179

4.10 Images .....................................................................................180
4.10.1 Overview of differences between images ...............................180
4.10.2 Results on Images ................................................................181
4.10.2.1 Accessibility of Images .................................................181
4.10.2.2 General comments on images .........................................182
4.10.2.3 Comments from partially sighted participants ..................184

4.11 Discussion on accessibility of examples ......................................184
4.11.1 Discussion on Tables .............................................................184
4.11.2 Discussion on Forms ............................................................186
4.11.3 Discussion on Frames ...........................................................187
4.11.4 Discussion on Imagemaps .....................................................189
4.11.5 Discussion on Headers ........................................................189
4.11.6 Discussion on References .....................................................190
4.11.7 Discussion on Images ...........................................................190
4.11.8 Page authors' design decisions ..............................................191
4.11.9 Inaccessibility of the examples ........................................... 192
4.11.10 Assessment of examples by Bobby .................................... 192
  4.11.10.1 Agreement between Bobby and participants .................. 196
  4.11.10.2 Disagreement between Bobby and participants ............... 196
  4.11.10.3 Limitations of Bobby .............................................. 198
4.11.11 Assessment of examples by W3C Validator ......................... 198
4.11.12 Discussion of method .................................................. 199
  4.11.12.1 Difficulties with recruitment ................................... 199
  4.11.12.2 Limitations of method ............................................ 201
  4.11.12.3 Decline in participation ......................................... 202
  4.11.12.4 Negative choices .................................................. 202

4.12 Conclusions on the accessibility of the examples ....................... 203
  4.12.1 Definition of ‘accessibility’ ....................................... 208
    4.12.1.1 Usability ...................................................... 209
  4.12.2 Alternative approach to improving accessibility .................. 209
  4.12.3 Summary of conclusions ............................................ 211

4.1 Introduction

This chapter presents the second of two related studies. The first of these two studies investigated the use of the Web Content Accessibility Guidelines by page authors and provided some interesting insights into how the Guidelines were used and how the page authors interpreted the advice. The other aspect of the Guidelines that the researcher wanted to evaluate was the accessibility of the pages that had been designed according to the Guidelines. It was therefore decided that people with visual impairments should evaluate the pages which had been produced in the first study. The pages were collated into a Web site that was made available on the Web for visually impaired people to evaluate the accessibility of the pages.

4.1.1 Key to abbreviations used in this chapter

In this chapter the browsers and screenreaders used by the participants are frequently referred to. In order to save space and to make reading easier a key has been developed that will be used to refer to the browsers and screenreaders. For example, IE401 + JFW indicates that the participant was using Internet Explorer 4.01 with the Jaws for Windows screenreader. Details of this key are shown in Table 4.1, below.

133
Table 4.1: Key to abbreviations used for browsers and screenreaders.

<table>
<thead>
<tr>
<th>Browser</th>
<th>Abbreviation</th>
<th>Screenreader</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>WebSpeak</td>
<td>WS</td>
<td>Jaws for Windows</td>
<td>JFW</td>
</tr>
<tr>
<td>Internet Explorer 3.0</td>
<td>IE3</td>
<td>WindowEyes</td>
<td>WE</td>
</tr>
<tr>
<td>Internet Explorer 4.01</td>
<td>IE401</td>
<td>WinBridge</td>
<td>WB</td>
</tr>
<tr>
<td>Internet Explorer 5.0</td>
<td>IE5</td>
<td>VocalEyes</td>
<td>VE</td>
</tr>
<tr>
<td>Braillesurf</td>
<td>BS</td>
<td>Draculanet</td>
<td>DN</td>
</tr>
<tr>
<td>Lynx</td>
<td>LX</td>
<td>SR/2</td>
<td>SR2</td>
</tr>
<tr>
<td>Netscape</td>
<td>NS</td>
<td>Hal95</td>
<td>HL</td>
</tr>
<tr>
<td>VIP</td>
<td>VIP</td>
<td>WWW</td>
<td>WWW</td>
</tr>
<tr>
<td>Internet Explorer 5 set for</td>
<td>IE5 font +</td>
<td></td>
<td></td>
</tr>
<tr>
<td>large font and high contrast</td>
<td>contrast</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet Explorer 4 set for</td>
<td>IE4 font +</td>
<td></td>
<td></td>
</tr>
<tr>
<td>large font and high contrast</td>
<td>contrast</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2 Method

4.2.1 Design

This study was an investigation into the extent to which a set of Web pages was accessible to blind and partially sighted Web users. The Web pages had been adapted using the WCA Guidelines. As part of the evaluation of the Guidelines it was important to test the accessibility with blind and partially sighted Web users. Visually impaired people were invited to evaluate the accessibility of several examples of Web pages and indicate which example they found to be the most accessible, and why. The Web site was evaluated remotely and participants sent their responses via email.

4.2.2 Participants

Twenty participants took part in the study. Two of these used two different browsers. It is common practice for blind people to use more than one browser because different browsers can present the same information in different ways. These two participants purposely performed the series of tasks twice, using a different browser each time. Both of these participants used the IE401 and WS browsers. One participant used IE401 with JFW; the other used IE401 with both

---

56 VIP = Visual Image Processing (http://www.jbliss.com/).

57 WWW is a browsing tool developed by the W3C (http://www.w3.org/LineMode/)
JFW and Hal. This latter participant switched between the two screenreaders when they felt it was necessary. The data from each of these two participants have been treated as if they were provided by two participants. In order to make the following discussion easier to read the data is presented as if 22 participants provided it. However, the reader should bear in mind that only twenty participants took part in the study, two of who performed the series of tasks with two different browsers.

An overview of the hardware and software used by each participant is provided in Appendix 6.

4.2.2.1 Recruitment

Some participants were recruited via email lists that discuss issues related to computer access for blind people. The email message sent to the lists can be found in Appendix 7. Other participants had taken part in previous studies run by the researcher. Several participants were recruited at a course at which the researcher assisted. Forty-one people were sent personal invitations to take part because they were known by the researcher or her colleagues.

4.2.2.2 Sight status

Of the 20 participants, 18 used synthesised speech or braille output to access the computer. These participants are defined as ‘blind’. Two participants had some useful sight and set their browsers to display text in ‘enlarged font’ and ‘high contrast’. These participants are defined as ‘partially sighted’.

4.2.2.3 Experience of using the Web

The participants had different levels of experience of using the Web. It ranged from "very little" to 13 years using the Internet, with an average of 2.7 years.

---

38 The researcher had unexpected difficulties recruiting sufficient numbers of participants. This is discussed further in Section 4.11.12.1 of this chapter.
39 The figure of 13 years is taken to represent approximately 10 years using the Web, which was formed in 1989 (Levine, Baroudi, & Levine Young, 1998).
4.2.2.4 Monthly Web usage

The participants reported spending different amounts of time using the Web in a month. It ranged from “very little” to 150 times per month with an average of 33 times per month.

4.2.2.5 Browsers

The participants used ten different browsers, or combinations of browsers (see Table 4.2). IE401 was the most common browser. One participant alternated between using IE401 and LX, another alternated between BS and IE3. Of the browsers listed in Table 4.2, WS, BS, and VIP are designed for visually impaired users (as described in Chapter 1) whereas the others are ‘mainstream’ browsers.

<table>
<thead>
<tr>
<th>Browser and version</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>IE 4.01</td>
<td>8</td>
</tr>
<tr>
<td>WebSpeak 2.5</td>
<td>2</td>
</tr>
<tr>
<td>IE 3.0</td>
<td>2</td>
</tr>
<tr>
<td>IE 5.0</td>
<td>2</td>
</tr>
<tr>
<td>WebSpeak 2.25</td>
<td>1</td>
</tr>
<tr>
<td>WebSpeak x</td>
<td>1</td>
</tr>
<tr>
<td>IE 4.01 and Lynx 2.8</td>
<td>1</td>
</tr>
<tr>
<td>Braillesurf and IE 3.0</td>
<td>1</td>
</tr>
<tr>
<td>Lynx 2.8.2-pre</td>
<td>1</td>
</tr>
<tr>
<td>Netscape 4.05</td>
<td>1</td>
</tr>
<tr>
<td>VIP 2.1</td>
<td>1</td>
</tr>
<tr>
<td>WWW</td>
<td>1</td>
</tr>
</tbody>
</table>

NB. The figures in all tables include the data from the two ‘extra’ participants.

4.2.2.6 Screenreaders and speech synthesisers

The participants used eight different screenreaders (see Table 4.3). JFW was the most common. Eleven different speech synthesisers were used in total (see Table 4.4). Apollo and Softvoice were the most common.
Table 4.3: Screenreaders used.

<table>
<thead>
<tr>
<th>Screenreaders</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jaws</td>
<td>8</td>
</tr>
<tr>
<td>WindowEyes</td>
<td>2</td>
</tr>
<tr>
<td>WinBridge</td>
<td>1</td>
</tr>
<tr>
<td>VocalEyes</td>
<td>1</td>
</tr>
<tr>
<td>Draculanet</td>
<td>1</td>
</tr>
<tr>
<td>SR/2</td>
<td>1</td>
</tr>
<tr>
<td>Jaws / Hal95</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.4: Speech synthesisers used.

<table>
<thead>
<tr>
<th>Speech synthesisers</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apollo</td>
<td>4</td>
</tr>
<tr>
<td>Softvoice</td>
<td>4</td>
</tr>
<tr>
<td>DecTalk</td>
<td>3</td>
</tr>
<tr>
<td>Eloquence</td>
<td>2</td>
</tr>
<tr>
<td>Proverbe</td>
<td>2</td>
</tr>
<tr>
<td>Flextalk</td>
<td>1</td>
</tr>
<tr>
<td>Keynote</td>
<td>1</td>
</tr>
<tr>
<td>Doubletalk</td>
<td>1</td>
</tr>
<tr>
<td>Multivoice</td>
<td>1</td>
</tr>
<tr>
<td>Sounding board</td>
<td>1</td>
</tr>
<tr>
<td>Prose 2020</td>
<td>1</td>
</tr>
</tbody>
</table>

4.2.2.7 Braille displays

Five participants used a braille display (see Table 4.5). The most common was Alva.

Table 4.5: The braille displays used.

<table>
<thead>
<tr>
<th>Braille display</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alva</td>
<td>3</td>
</tr>
<tr>
<td>Braille Window</td>
<td>1</td>
</tr>
<tr>
<td>Clionetbraille</td>
<td>1</td>
</tr>
</tbody>
</table>

4.2.2.8 Operating system

Five different operating systems were used (see Table 4.6). Most participants used Windows 95.
Table 4.6: The operating systems used.

<table>
<thead>
<tr>
<th>Operating system</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows '95</td>
<td>16</td>
</tr>
<tr>
<td>Windows '98</td>
<td>4</td>
</tr>
<tr>
<td>OS/2</td>
<td>1</td>
</tr>
<tr>
<td>Windows NT</td>
<td>1</td>
</tr>
<tr>
<td>DOS</td>
<td>1</td>
</tr>
</tbody>
</table>

4.2.2.9 Common combinations of software

Seven of the participants who used JFW 3.2x used it in conjunction with IE401, probably because JFW 3.2x has been developed to work with IE401. The four participants who used WS also used the Softvoice speech synthesiser, probably because this synthesiser is supplied with WS.

4.2.3 Materials

4.2.3.1 The Web site

The pages adapted in the first of the two studies were collated into a Web site60. It might have been possible to include two different types of "control" Web pages. One such condition would have been an 'ideal' set of Web pages with the pages adapted by the page authors could have been compared. However, it is not possible to define the 'ideal' way of interpreting the WCA Guidelines: while the WAI provide examples of how to implement the Guidelines these are not definitive. For example, the results of the previous chapter show that the Guidelines relating to tables were interpreted differently by different participants. Therefore it would not have been possible to create 'ideal' pages. Another control condition might have been a set of pages with no adaptations. However, these would have been difficult for blind people to access, and thus to evaluate. Therefore this approach could not be considered.

It was thought that the pages that contained the tables and images contained too much text for the participants to read through, particularly because the same information would be read in several different examples. The pages were therefore

*The test site is available at: http://phoenix.herts.ac.uk/sdru/test/
condensed so that they contained only the sections that the participants in the previous study had adapted. For the Tables and Images the preceding paragraph from the original page was included to provide a context the elements.

Some of the pages adapted in the previous study contained more than one element or type of presentation, other pages contained just one of these. For example, some of the pages contained tables, images, headers and references, and other pages contained one imagemap, frameset or form.

All the examples of each element or type of presentation were presented together. For instance, the examples of tables were presented as a group, as were the headers of papers. This enabled the participants to compare the elements and presentations with one another, rather than considering one page containing different elements.

The Web site containing the examples had an introductory page that welcomed potential participants and provided instructions for taking part in the study. A ‘zipped’ version of the site was made available to allow participants to evaluate the site offline. The welcome page included navigation links to all the groups of elements and a link to the questions. Links were also provided to enable participants to navigate between the groups of elements and different presentations, and between the different examples.

A blind computer expert was asked to view the site before it was made public. He was sent a copy of the instructions so that he could work through the series of tasks. The expert provided feedback regarding the instructions. He identified one potential problem: the instructions stated that participants could do as much of the series of tasks as they liked, but did not state which parts of the task they should begin with, in case they did not have time to complete it all.

On the basis of these comments, an additional instruction was added which stated that the most important parts of the overall task were tables, forms and frames, and that these parts should be completed first if the participant thought they may not be
able to complete the whole task. These elements are the most important because they are known to cause the most access difficulties for visually impaired people.\(^1\)

Although participants were guided to complete the tasks in an appropriate order, it was not possible for the researcher to counterbalance the order in which the examples were evaluated because the participants were accessing a web site from remote locations. This remote aspect of the study means that it was also not possible to control for practice or fatigue effects.

An overview of all the examples evaluated is provided in the following sections. Screenshots of all the examples can be found in Appendix 8. The HTML code of the examples can be found in Appendix 9.

Five different examples of tables were evaluated. The examples differed in terms of visual presentation and in the underlying HTML code.

Two examples of forms were evaluated. They were similar to each other in that they included the same controls, but differed in a few key ways which are outlined below.

There were two examples of frames. Both examples offered the same functionality, but differed in their design. The function of the frames was to provide a frame with a table of contents containing links to a set of papers. When a link is selected the relevant paper is displayed in another frame.

There were two examples of imagemaps. The imagemaps are divided into five different areas which represent blindness, deafness, mobility impairments, low vision and learning impairments.

Four examples of headers were evaluated. Each example presents the header section of a paper in a different way.

---

\(^1\) See W3C fact sheet on the WCA Guidelines available at: http://www.w3.org/1999/05/WCAG-REC-fact.html (checked Nov 2000).
There were three examples of references which each presented a paragraph in different ways. The paragraph contained a reference to an item in a list of references.

Three different images were evaluated, which are referred to as figures 1, 2 and 3. There were six examples of figure 1 and three examples each of figures 2 and 3. All the examples of each figure were presented on one page. The examples were separated by headings, 'Example 1', 'Example 2' etc.

4.3 Procedure

The participants visited the Web site, tested the examples, and decided which they found to be the most accessible. They then indicated their preference in a questionnaire (described in Section 4.3.3 below and presented in Appendix 7) along with details of their reasons for their preference. The questionnaire was then emailed to the researcher.

4.3.1 The set of tasks

The series of tasks performed by the participants involved the evaluation of the accessibility of several different HTML elements and different presentations of common information. The elements were tables, forms, frames, imagemaps, and images. The information presented in different ways were references and headers. The number of examples in each 'set of examples' is shown in Table 4.7 below.

One of the participants did not take part in the study from a remote location, but attended the laboratory. This participant does not have an Internet connection at home and said they would not have been able to get permission from their employer to perform the series of tasks during work hours. In order for the participant to use their usual screenreader and browser the researcher installed a demonstration version of JFW, and IE401 on the computer in the laboratory. This participant only provided

---

62 In order to make the discussion easier to read the term 'sets of examples' will be used to mean a set of examples of an HTML element, such as frames, or a set of examples of the different presentations of information, such as the presentation of references. So a 'set of examples' could be the examples of frames, or the examples of references.
data on the tables because they only had time to evaluate these examples before they had to leave.

Table 4.7: Number of examples in each set of examples.

<table>
<thead>
<tr>
<th>HTML Elements</th>
<th>Number of examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tables</td>
<td>5</td>
</tr>
<tr>
<td>Frames</td>
<td>2</td>
</tr>
<tr>
<td>Forms</td>
<td>2</td>
</tr>
<tr>
<td>Imagemaps</td>
<td>2</td>
</tr>
<tr>
<td>Image: figure 1</td>
<td>6</td>
</tr>
<tr>
<td>Image: figure 2</td>
<td>3</td>
</tr>
<tr>
<td>Image: figure 3</td>
<td>3</td>
</tr>
<tr>
<td>Presentations</td>
<td></td>
</tr>
<tr>
<td>Headers</td>
<td>4</td>
</tr>
<tr>
<td>References</td>
<td>3</td>
</tr>
</tbody>
</table>

4.3.2 Tasks performed by individual participants

The number of participants who evaluated the nine sets of examples is shown in Table 4.8, below. This table presents sets of examples in the order in which they were presented in the Web site. It can be seen that there was a gradual tail-off from 20 and 21 participants evaluating the first two sets, to 15 participants evaluating the final sets.

Table 4.8: Number of participants who evaluated each set of examples, and order in which elements were presented.

<table>
<thead>
<tr>
<th>Set of examples</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tables</td>
<td>20</td>
</tr>
<tr>
<td>Forms</td>
<td>21</td>
</tr>
<tr>
<td>Frames</td>
<td>18</td>
</tr>
<tr>
<td>Imagemaps</td>
<td>19</td>
</tr>
<tr>
<td>Headers</td>
<td>18</td>
</tr>
<tr>
<td>References</td>
<td>18</td>
</tr>
<tr>
<td>Images</td>
<td>15</td>
</tr>
<tr>
<td>All 9 sets of examples</td>
<td>11</td>
</tr>
</tbody>
</table>

4.3.3 Questionnaire

The email message sent to participants included a questionnaire (see Appendix 7). This asked participants which example of each element they preferred, and to
describe any particular problems they had with the element and any features of the element they particularly liked. Regarding the tables, participants were asked to state which example was the most accessible, and to find the answer to a particular question, “What size was the 2 cm sphere perceived to be with the outside presentation?”. This meant that the tables had to be interrogated in order to find the answer. This question was added because it was thought that the participants might not have been able to make an objective decision about the accessibility of a table, unless they had to navigate and interrogate it.

Participants emailed their answers to the questionnaire to the researcher. An alternative method might have been the use of an online form, which participants could have completed online. The advantage of this method would have been that the participants would not have had to switch from their Web browser to their email package to complete the questionnaire. However, a disadvantage would have been that the form might not have been accessible to all participants.

4.4 Tables

Five examples of tables were evaluated by 22 participants.

4.4.1 Overview of differences between examples of tables

The relevant rows and columns from examples 1-5 are presented in Appendix 8. In examples 1, 2 and 5 there were two rows of headers and two of the column headers were merged across two columns. Example 4 avoided merged columns by changing the headers so that there was one header per column. All examples included the phrase “See note x”. In examples 1, 2, 4 and 5 ‘note 1’ and ‘note 2’ were links to notes at the end of the table. Headers and IDs were provided in four of the five examples, in four different ways. Summaries were provided in examples 4 and 5. A caption was provided in the two tables in example 3. The features of each example are summarised below, including whether Summaries, Captions, links and association were used.

In example 1:
- No Summary or Caption was provided.
• Links were used for Notes 1 and 2.
• All headers were given IDs, T1-T10, which included the 'Cube' and 'Sphere' headers.
• All data cells were associated to headers.
• Data cells in the 'Actual size' column were associated to that header.
• 'Mean' and 'Over/under estimation' headers were associated with 'Perceived size' header.
• Data cells in the 'Mean' and 'Over/under estimation' columns were associated with those headers. (T5 and T6 were associated with T3, T7 and T8 were associated with T4.)
• No cells were associated with headers T1, T9, or T10.

In example 2:
• No Summary or Caption was provided.
• Links were used for notes 1 and 2.
• No additional mark-up was used: IDs were not given to headers, so there was no association of cells with headers.

Example 3 was different from the other examples in three ways. Firstly, the data was split into two tables, one each for cubes and spheres. Secondly, the layout was changed so that the main headers were row headers, rather than column headers. Thirdly, the text 'Mean (cm)' and 'Over/underestimation' were not marked-up as headers but were marked-up with the Strong tag, which made the text bold. Finally, all cells were given widths in percentages. In the two tables:
• Captions were provided for both tables: 'Table of experiment results for object type CUBE' and 'Table of experiment results for object type SPHERE'.
• Links to the notes were not provided.
• Summaries were not provided.
• The headers 'Inside presentation' and 'Outside presentation' were given IDs, 1 and 2.
• The 'Mean' and 'Over/under estimation' 'headers' were associated with IDs 1 and 2.
• Association was only created between headers: no data cells were associated with any headers.
In example 4:

- A Summary was provided, 'This table shows the relationship between the actual size of objects and their perceived sizes'.
- Links were used for notes 1 and 2.
- All headers were given IDs, T1-T6.
- All cells in each column were associated with the column header.

In example 5:

- A Summary was provided, 'This table charts the errors in perception of several different objects'.
- All headers were given IDs, T1-T10.
- The cells in the 'Actual size' column were associated with three headers: 'Actual size', 'Object type' and 'Cube / Sphere'.
- The cells in the 'Mean' column were associated with the 'Mean' and 'Perceived size' headers.
- The cells in the 'Over/under estimation' column were also associated with the 'Perceived size' header.
- The first data cell of each row was associated with T1, T2 and T9 or T10. The other data cells on the row were only associated with the relevant column headers.

4.4.2 Results on Tables

4.4.2.1 Accessibility of Tables

Table 4.9 (below) shows the number of participants who found each example to be the most accessible. One of the two partially sighted participants found all the tables accessible and scored 5/5. However, the other scored 0/5.
Table 4.9: Number of participants who found each example of tables the most accessible.

<table>
<thead>
<tr>
<th>Example</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>2</td>
</tr>
<tr>
<td>Example 2</td>
<td>1</td>
</tr>
<tr>
<td>Example 3</td>
<td>8</td>
</tr>
<tr>
<td>Example 4</td>
<td>3</td>
</tr>
<tr>
<td>Example 5</td>
<td>1</td>
</tr>
<tr>
<td>All examples</td>
<td></td>
</tr>
<tr>
<td>equally accessible</td>
<td>2</td>
</tr>
<tr>
<td>No examples</td>
<td></td>
</tr>
<tr>
<td>accessible</td>
<td>3</td>
</tr>
</tbody>
</table>

In general, those that found example 3 the most accessible did so because it was split into two separate tables and the headers were on rows, rather than on columns. Those that found example 4 the most accessible did so because none of the column headers were merged over two columns, so it was more clear which header related to which column. Those that found none of the examples accessible said that all the cells ran into each other so that they could not be distinguished from each other.

The preferences of the participants who used text-based browsers were evenly distributed. However, there was a trend towards significance for the users of GUI browsers to prefer example 3 (Chi square = 16.0, df=6, p<0.025). It is thought that this preference for example 3 by those who used a GUI browser is due to the different layout being more usable with a GUI browser.

For analysis the participants were divided into those who had up to and including 2 years’ experience of using the Web and those that had more than 2 years’ experience. A Chi square showed that those in the less experienced group were significantly more likely to find none of the examples accessible, whereas those with more experience were significantly more likely to find all the examples equally accessible. Chi square likelihood ratio = 13.86, df = 6, p=0.03.

The average length of experience of using the Web for those who found none of the examples accessible was 1.1 years, in comparison to the average experience of those
that chose example 3 was 3.3 years (median = 2.5), and 4 years for example 4 (median = 4.).

Four of the five participants who used braille displays found example 3 the most accessible (the other found none of the examples to be accessible). These four may have found example 3 the most accessible because the rows were presented in a useful way by the braille display. Unfortunately, these participants did not comment about this directly.

The two partially sighted participants each chose different examples. One chose example 1. Although they do not say why, they did say they found example 3 the most difficult to read. This opinion is the opposite to that of the majority of the blind participants, who found example 3 easier to read. The other partially sighted participant chose example 5 because the columns were found to be of equal width, and the table was narrower than the other examples, which meant they did not have to scroll so much. This participant also said that the headers in examples 1 and 2 were confusing because the cells were not directly beneath each other. Of example 3 this participant said they had to scroll a lot and that they generally prefer vertical tables to horizontal tables.

4.4.2.2 Answers to the set question

Table 4.10 shows the number of answers to the set question provided by each participant, and the answers given for each example. Table 4.11 shows the number of correct answers given for each example. Three of the participants found none of the examples to be accessible and were not able to provide any answers to the questions. The participants who evaluated the tables did not all provide answers to all the set questions.

The scores were not significantly related to the amount of experience the participants had of using the Web. The average experience for those who scored 5/5 is 3 years (median 3.5 years) compared to the average experience of those who scored 0/5 of 3.6 years (median 2 years). The median experience was greater for those with a higher score than that for those that had low scores.
Table 4.10: Answers given by each participant for each example and the percentage correct.

<table>
<thead>
<tr>
<th>Example</th>
<th>Correct answer</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Number Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1.8</td>
<td>1.5</td>
<td>1.6</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>P2</td>
<td>1.8</td>
<td>1.5</td>
<td>1.6</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>P3</td>
<td>1.8</td>
<td>1.5</td>
<td>1.6</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>P4</td>
<td>1.8</td>
<td>1.5</td>
<td>1.6</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>P5</td>
<td>1.8</td>
<td>1.5</td>
<td>1.6</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>P6</td>
<td>1.8</td>
<td>1.5</td>
<td>1.6</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>P7</td>
<td>1.8</td>
<td>1.5</td>
<td>1.6</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>P8</td>
<td>1.8</td>
<td>1.5</td>
<td>1.6</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>P9</td>
<td>1.8</td>
<td>1.5</td>
<td>2.0 x</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>4/5 (80%)</td>
</tr>
<tr>
<td>P10</td>
<td>1.8</td>
<td>2.5 x</td>
<td>2.3 x</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>3/5 (60%)</td>
</tr>
<tr>
<td>P11</td>
<td>2.3 x</td>
<td>2.3 x</td>
<td>2.3 x</td>
<td>2.3 x</td>
<td>2.3 x</td>
<td></td>
<td>0/5 (0%)</td>
</tr>
<tr>
<td>P12</td>
<td>10% x</td>
<td>25% x</td>
<td>8% x</td>
<td>&quot;no size given&quot; x</td>
<td>2.00% x</td>
<td></td>
<td>0/5 (0%)</td>
</tr>
<tr>
<td>P13</td>
<td>2.3 x</td>
<td>2.3 x</td>
<td>2.3 x</td>
<td>2.3 x</td>
<td>2.3 x</td>
<td></td>
<td>0/5 (0%)</td>
</tr>
<tr>
<td>P14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.9</td>
<td>2.0</td>
<td></td>
<td>2/2 (100%)</td>
</tr>
<tr>
<td>P15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.8 x</td>
<td>-</td>
<td></td>
<td>0/1 (0%)</td>
</tr>
<tr>
<td>P16</td>
<td>-</td>
<td>-</td>
<td>2.0 x</td>
<td>-</td>
<td>-</td>
<td></td>
<td>0/1 (0%)</td>
</tr>
</tbody>
</table>

Key: ‘x’ = incorrect answer, ‘-‘ = no answer given.

Table 4.11: Number of correct answers given for each example.

<table>
<thead>
<tr>
<th>Example</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of correct answers given and percentage</td>
<td>10/13 (77%)</td>
<td>9/13 (69%)</td>
<td>8/14 (57%)</td>
<td>11/15 (73%)</td>
<td>11/14 (78%)</td>
</tr>
</tbody>
</table>

The scores of participants were not significantly related to the browser and screenreader they used, whether the browser was special or mainstream, or whether the browser was text-based or GUI. Five participants scored 0/5 (4 x IE + JFW, 1 x BS + DN). Eight scored 5/5 (3 x IE + JFW, 1 x IE + WE). Analysis of this data using a Chi Square found that there were no significant differences between the number of correct answers given for each example (Chi square = 1.1, df = 4, n.s).
4.4.2.3 Uncertainty

Two participants expressed uncertainty about the answers they had given. Both said they were not sure that they were extracting the correct information from the tables. In fact, these participants scored 5/5 and 4/5. The participants who got lower scores did not comment on their confidence in their answers.

4.4.2.4 Incorrect answers

As well as examining the participants' scores, it is also useful to look at the incorrect answers given. The pattern of the incorrect answers provides some insight into which column or row the participant was reading. Of the 21 wrong answers given, 11 were '2.3'; two participants gave '2.3' as the answer for all examples (despite being informed that the answer was different for all examples).

It is assumed that the participants who gave the answer '2.3' were probably reading the column for inside presentation, which was two columns to the left on the same row, in examples 1, 2, 4 and 5. Or they may have been reading the row for the 2.5 cm sphere, which was in the same column but in the next row down in examples 1, 2, 4 and 5. These assumptions may apply particularly to the two participants who gave '2.3' as the same answer for all the examples. One participant gave all their answers in percentages but it is not clear why.

The incorrect answer of '2.3' was given more often (3 times) for example 3 than for the other examples (compared to twice for all the other examples). In example 3 this figure was in the same column but two rows up, and also in the same row but in the next column to the right.

It is also useful to look at the comments given by the 3 participants who scored 0/5. One said they found in examples 1, 2, 4 and 5 it was difficult to distinguish between the rows and columns, and the data was read in a constant stream, rather than in useful chunks (IE401 + JFW). This participant found example 3 was also read in a stream, but found it easier to identify the columns and rows. Another participant reported that they used the JFW cursor to find the data but found this to be time-consuming (IE5 + JFW).
One of the partially sighted participants gave 5 incorrect answers (IE5 font + contrast). They reported that they found the headers confusing because the cells were not all the same width and were not directly underneath each other.

4.4.2.5 General comments on tables

Participants made the following comments regarding the difficulties they experienced with the examples of tables.

- At first, the word 'see' in 'see note x' was mistaken for 'c'.
- Access was very difficult and only possible with much effort and concentration.
- Cells that are blank can make the table more difficult to read.
- It was more difficult to assimilate information because the columns are complicated and formatted in a 'fussy' way.
- Access was impossible because all the cells ran into each other and it was impossible to differentiate between the columns and rows. All were equally bad.
- The data in the cells could not be related to the headers.
- The headers could not be related to the data cells.

One participant (WWW under OS/2) speculated that they might have found navigation of the tables easier if they had been using a GUI browser and a screenreader because these would support Microsoft Active Accessibility (MSAA) and this may have made the tables more accessible.

4.4.2.5.1 Methods employed

One participant described the method they used to navigate the tables. This involved first acquiring a general idea of the layout of the table, and then deciding which of the columns contained the answer to the question. They then followed this column down a certain number of rows until they got to the row that they believed to contain the answer. This participant scored 3/5 which suggests that their method was not

---

63 Microsoft Active Accessibility (MSAA) provides 'hooks' in applications developed by Microsoft that screenreaders (and other assistive technology) can utilise in order to improve the accessibility of the application. For further information see http://www.microsoft.com/enable/msaa.
fail-safe. This method is similar to that used by the participant who performed the task in the laboratory.

4.4.2.5.2 Switching modes

Two participants who described the methods they used mentioned that they had to switch between two different screenreader 'modes'\(^{64}\). One said they had to use the 'Jaws cursor' to distinguish between the columns and rows. The other participant said they had to switch modes in order to scroll down the page. The way in which these participants describe the action of switching modes suggests that it is not an easy transition; it seems as though the switching interferes with their assimilation of the page.

4.4.2.5.3 Braille displays

Three participants mentioned the use of braille displays in relation to the tables. One participant who used two different browsers, used one with a braille display, and said that provided better access to the tables. Another participant speculated that the tables would be easier to navigate using a braille display. A third participant said they believed the tables would be impossible to navigate without a braille display unless they had "extraordinary memory" (BO).

It is assumed that braille displays provide more information about the spatial layout of a table and therefore make it more accessible. However, the scores of the five participants who used braille displays vary greatly. Two found none of the examples to be accessible; two scored 5/5; the fifth scored 4/5.

---

\(^{64}\) Screenreaders often have two 'modes' which each read the contents of the screen in a different way. Each mode involves either the application cursor, which is the one used to interact with the application, or the screenreader cursor which can be used to explore the screen without moving the application cursor. One mode reads as the application cursor is moved, the other reads as the screenreader cursor moves.
4.4.2.6  Comments on specific aspects of the tables

4.4.2.6.1 Table headers

Examples 1, 2, 4 and 5 were similar, particularly in terms of their headers. Five participants reported a number of difficulties associated with reading the table headers (4 x IE401 + JFW, 1 x WWW + SR2). The difficulties reported for examples 1, 2, 4 and 5 are similar. These five participants describe problems with the following:

- reading the column headers;
- reading one column header at a time;
- following the headers;
- reading the headers in the correct order;
- knowing which headers relate to what data;
- distinguishing between the column headers and the title of the table.

One of the partially sighted participants found the headers confusing because the cells were not all the same width and were not always placed directly underneath each other (IE5 font + contrast).

4.4.2.6.2 Association between cells

None of the participants commented on the association between headers and cells in examples 1, 3, 4, and 5. This is probably because none of the browsers or screenreaders used by the participants support the IDs and Headers tags. However, the use of these tags does not seem to have interfered with the participants’ navigation of the tables.

4.4.2.6.3 Table headers: wrapped text

In all the examples, except example 3, the text of the header cells wraps onto more than one line. The reports of the participants suggest that screenreaders do not present the whole header cell as one unit, but read across the first line of each header cell, then across the second line, etc. This was confirmed when the researcher observed one of the participants performing the task. The first two rows of examples
1, 2, 4 and 5 are reproduced in Figure 4.1 below. Below this, the way in which it was typically read by a screenreader in a typical system is presented.

<table>
<thead>
<tr>
<th>Object type</th>
<th>Actual size</th>
<th>Perceived size: inside presentation</th>
<th>Perceived size: outside presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (cm)</td>
<td>Over/under estimation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean (cm)</td>
<td>Over/under estimation</td>
</tr>
</tbody>
</table>

**Figure 4.1: The header rows from example 1.**

The above rows would typically be read as:

Object, Actual, Perceived size: inside, Perceived size: outside, type, size, presentation, presentation, Mean (cm), Over/under, Mean (cm), Over/under, estimation, estimation.

In contrast to these difficulties, one participant who used WS reported that the way in which the rows containing the headers was presented informed them of the number of columns that were present.

### 4.4.2.6.4 Vertically centred headers

In examples 1, 2, 4, and 5 the headers in the ‘Object type’ column (i.e. ‘Cube’ and ‘Sphere’) were centred vertically in their cells (for example see Figure 4.2 below). Three participants reported that this meant that the header was not read with the first row of the data for that header (2 x IE401 + JFW, 1 x IE401 + WE). The header information was only read with the second or third row of data. The first four rows of data from examples 1, 2, 4 and 5 are reproduced in Figure 4.2. Below this, the way in which these rows were typically read is presented.

<table>
<thead>
<tr>
<th>Cube</th>
<th>1</th>
<th>1.8</th>
<th>+80%</th>
<th>See Note 1</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.5</td>
<td>1.7</td>
<td>+13%</td>
<td>1.6</td>
<td>+7%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.4</td>
<td>+20%</td>
<td>2.0</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td></td>
<td></td>
<td>2.4</td>
<td>-7%</td>
</tr>
</tbody>
</table>

**Figure 4.2: The first four rows of data from examples 1, 2, 4, and 5.**

The above data was typically read as:
One of the participants who used two different browsers (WS and IE401 + JFW) said that WS read the 'object type' column in a more logical position than JFW. Again, this difference between the way in which the table is read by a screenreader and a special browser is probably because the screenreader reads line by line and the special browser reads element by element.

4.4.2.6.5 Table headers: example 3

Example 3 had a different layout to the other examples. Three participants liked this layout (2 x IE401 + JFW, 1 x NS + JFW) for the following reasons:

- the headers were row headers, rather than column headers, which made it easier to find data;
- there was less need to remember what the headers were;
- the headers were more explicit and therefore it was easier to understand the table;
- it was more obvious whether the data cells related to spheres or cubes.

While several participants found example 3 to be the most accessible, one participant reported that they had the same difficulties with example 3 as they had with examples 1 and 2 in following the column headers (WWW + SR2).

4.4.2.6.6 Table headers: example 4

One participant reported difficulties with understanding the layout and meanings of the headers in example 4. In contrast, two other participants found example 4 easier to use. One said that it was easier to count the number of columns than in the other examples; the other that they found the column headers to be more explicit.

Interestingly, these three participants all used the same browser and screenreader (IE401 + JFW).
4.4.2.6.7 Table headers: memorising

Three participants mentioned that they memorised the column headers in order to identify which header was associated with each cell. One said that they thought they could only find the answer because the tables were small and they could memorise the headers. Another said that they found all the examples difficult to read because they had to remember the headers. The third participant said that they found example 3 easier because there was less need to remember the headers.

The participant who attended the laboratory was observed to memorise the headers by counting the column headers from left to right. He then counted along the rows and associated each cell with its appropriate header. This was a long process because the headers were only clear as separate entities after the participant had read them repeatedly. Only once the headers were clear could the participant go on to read the cells.

4.4.2.6.8 Discussion on headers

From the participants’ reports regarding headers it seems that, in general, those who used mainstream browsers with screenreaders (the majority of participants) had more difficulties with reading the headers than those using special browsers. This is probably because the screenreaders read the tables row by row, either in one string, or column by column, but without giving the user any indication of the relationships between the different pieces of text. In contrast, the special browsers probably read one cell at a time, making the cells easier to distinguish.

4.4.2.6.9 Rows

One participant reported a difficulty with their screenreader not stopping at the end of a row (IE401 + JFW and HL). This made it difficult to distinguish between one row and the next. They reported that Hal would stop reading at the end of a row, whereas JFW sometimes ‘strayed’ onto the next row. This problem was only reported for example 1 but not for the others. It is possible that this problem was caused by the underlying HTML of example 1, rather than the design of the screenreaders. There is a difference in the mark-up of the data cells between
example 1 and examples 2, 4, and 5. In these other examples the data cells are marked-up as separate paragraphs, whereas the cells in example 1 are not.  

4.4.2.6.10 Columns

The JFW screenreader has a facility to reformat a Web page that is presented in columns. This function strips out the columns and presents the information as if in one column. One participant who used this function found that in the reformatted version it was still difficult to identify the different columns and what each represented. It is possible that this function is less useful if the information needs to be presented in a particular order.

4.4.2.6.11 Layout of Example 3

Example 3 was split into two separate tables, one each for cubes and spheres. Eight of the participants said they found this example to be the most accessible. Five of these participants gave the following reasons for this preference:

- It was more simple;
- It was not necessary to switch screenreader modes to find the answers;
- It was more understandable;
- It was broken down into manageable chunks of information.

Overall, the scores were the lowest for example 3 57% compared to the other examples: compared to 69% (example 1), 77% (example 2), 73% (example 4), and 78% (example 5). This is surprising because eight participants found example 3 to be the most accessible.

In contrast, two participants said that they found example 3 more difficult to navigate. One found this example the most confusing, another said it was more 'fiddly' because they had to scroll down (which involved switching screenreader modes).

65 In examples 2, 4 and 5 the cells are marked-up: `<TD><P ALIGN="CENTER">1.5</TD>`>, whereas in example 1 they are marked-up differently: `<TD><CENTER>1.5</CENTER></TD>`
The participants that found example 3 to be the most accessible were all blind. Neither of the two partially sighted participants preferred example 3. One partially sighted participant said that in this example it was “the most difficult to figure out where to look for the data” (IS), and it was difficult to navigate because they needed to scroll more.

Two participants reported difficulties with the cells that contained the phrase ‘See note x’ in example 3. Both said that the text wrapped onto three lines because the column was too narrow, which made it difficult to read.

One participant speculated about whether the approach taken in example 3 would be suitable for larger tables that contained more data. They suspected that it might be just as problematic as the other layouts if the table were larger.

4.4.2.6.12 Preference vs. performance
The fact that some of the participants found example 3 the most accessible but got lower scores with that example raises the issue of ‘preference vs. performance’. This is an issue discussed by Nielsen & Levy (1994) who performed a meta-analysis of studies which compared objective measures of users’ performance and their subjective preferences across two or more interfaces. It was found that while there was a strong positive association between users’ task performance and their subjective preference there were still many cases in which users preferred systems with which they performed less well. In the current context it is important to consider that a Web page that is perceived by a user to be accessible, may not allow them to find the information they require. While on the other hand, a Web page may not be perceived to be accessible even if it enables information to be accessed.

4.4.2.6.13 Summary / Caption
None of the participants commented on the Summaries or Captions that were included in the examples, perhaps because the browsers and screenreaders used do not support these tags.
4.4.2.7 Access to tables for partially sighted participants

One of the partially sighted participants found that in all the examples, except example 5, they had to scroll the window more than they would have liked. This was because they had the default font of their browser enlarged and therefore could see less of the content at any one time. Whether or not they had to scroll was the deciding factor for this participant in terms of which example they preferred. This participant found example 5 the most accessible because they did not have to scroll, and because all the columns appeared to be the same width.

One of the partially sighted participants described the attributes that an ideal table would have, for them: all columns would be the same width and the overall the table would not be too wide. (They do not say what constitutes ‘too’ wide.)

4.4.2.8 Suggestions for improvements

Some participants made suggestions regarding the design of the tables. One participant made a suggestion regarding tables that contain empty header cells. They said it would be easier to relate a data cell to the appropriate header if the empty cells in the header row contained a symbol to indicate that they were empty.

One participant said they liked the links for ‘see note x’ but suggested it would be better if there was also a way of getting back to the link in the table in order to continue reading. This participant could have used the ‘back’ command in their browser to get back to the link, but perhaps they did not know this.

4.5 Forms

Two examples of forms were evaluated by 21 participants (see Appendix 8).

4.5.1 Overview of differences between forms

In example 1:

- The Fieldset element and its Legend attribute were used to divide the controls into two groups, ‘Personal details’, and ‘General options’. Four controls at the end of the form were not in a group.
- The instructions were longer, often beginning with the word ‘please’.
• Default text was used in the Address field: ‘Please enter your address’.
• The Label attribute was used on two controls, Gender and Browser.
• The Accesskey and Tabindex attributes were not used.

The order of the controls was: Name, Address, Gender, Browser, Computer, Password, Browse, Submit, Reset.

In example 2:
• A heading was provided above the form: ‘For easy navigation MS Internet
  Explorer is better’.
• The Fieldset element was not used.
• One occurrence of Legend was used at the top of the form: ‘Use this form to
  provide your personal details’.
• The Label attribute was used on three controls: Name, Address, Browse button.
• The Accesskey attribute was used on the Browse control.
• The instructions were shorter, such as ‘Name’, ‘Address’.
• The Tabindex attribute was used, but not on all the controls.
The Tabindex order was: Name, Address, Password, Gender (male), Gender
(female), Browser (Netscape), Browser (Internet Explorer), Browser (Lynx),
Browse, Submit, Computer.

4.5.2 Results on Forms

4.5.2.1 Accessibility of Forms

Table 4.12, below, shows which examples the participants found to be the most accessible.

Table 4.12: Number of participants who found each example of forms to be the most accessible.

<table>
<thead>
<tr>
<th>Example</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>9</td>
</tr>
<tr>
<td>Example 2</td>
<td>9</td>
</tr>
<tr>
<td>All examples</td>
<td>1</td>
</tr>
<tr>
<td>equally accessible</td>
<td></td>
</tr>
<tr>
<td>No examples</td>
<td>2</td>
</tr>
<tr>
<td>accessible</td>
<td></td>
</tr>
</tbody>
</table>
4.5.2.2 Factors affecting accessibility

There was no significance between the participants' preference and whether they were using a mainstream or special browser. However in further analysis there was a significant relationship between preference for example 1 and the use of a GUI browser as opposed to a text-based or special browser (Chi square = 5.84, df = 1, p=0.015).

Nine participants found example 1 to be the most accessible because the Tabindex order in example 2 was incorrect. Many of these were using IE which supports the method of jumping from control to control, rather than reading the whole form line by line.

Those that chose example 2 used a variety of browsers and screenreaders. Two used text-based browsers, four used special browsers (one with a braille display), and three used GUI browsers with braille displays. Some of those who chose example 2 did so because they found the layout easier to navigate because the controls and their associated instructions appeared to be located on the same line.

4.5.2.3 Comments on the accessibility of form controls

Participants made comments regarding the different form controls. While reading this section it may be useful to bear in mind that visually impaired users often use the tab key to jump to the links and form controls on a page.

4.5.2.3.1 Checkboxes

Four participants found the checkboxes in example 1 were not presented in a useful way (3 x WS, 1 x IE401 + JFW). It was reported that WS announced "browser" and then "checked" or "unchecked", but did not announce the names of the browsers. This made it impossible for the participant to choose a particular item. The underlying HTML for the checkboxes in each example is different (see Appendix 10). The Label tag is used on each checkbox in example 1, but is not used in example 2. It seems that the participants who reported this problem are not able to access the content of the Label tag. It is thought this is because the page author did not implement the correct syntax for the Label tag. They implemented the 'input
type' and 'label' tags in the wrong order, and implemented the 'value' attribute rather than the 'id' attribute. The relevant code from the example and the HTML specification is compared in Appendix 10.

4.5.2.3.2 Default text

One of the participants who used WS reported that the phrase 'Please enter your address' was repeated in example 1. They did not realise that this phrase was used for both the instruction and as default text in the field. (Repetition of phrases is an important issue for visually impaired users, especially those using speech output, because it increases reading time and can be confusing.) In this case if WS had indicated to the user that the first occurrence of the phrase was an instruction, and the second default text, the participant would at least have been better informed.

4.5.2.3.3 Instructions

Three participants reported that they preferred the longer instructions in example 1. They believed that such instructions would be better for visually impaired users and inexperienced Web users.

Five participants found that the instruction for the Address control in example 1 was not automatically announced (2 x IE401 + JFW, 2 x WS, 1 x NS + JFW). One participant who used both JFW and WS was able to find the instruction using the JFW 'read line' function, but could not find it with WS. In example 2 this participant found the same problem occurred with both JFW and WS, but in this case the label could not be found with the JFW 'read line' function.

The problem with example 1 described above was probably due to a combination of factors. The participants were jumping from control to control using the tab key, but the instruction and the text field were on separate lines. As the participants jumped to each control the screenreader would not read the instruction, even with the use of a 'read line' function because the instruction was on the previous line.
4.5.2.3.4 **Browse button**

Several participants reported difficulties with using the 'Browse’ control. One said they did not understand its function and another found that it was not announced by WS. Two others found the control could not be operated (1 x IE401 + JFW, 1 x IE401 + font and contrast). It is not clear why these difficulties arose.

4.5.2.4 **Comments on accessibility features**

The two examples between them included implementations of the accessibility features Fieldset, Legend, Tabindex, Label, and Accesskey. The participants made comments regarding some of these features, particularly Tabindex. None of the participants mentioned the use of the Accesskey tag in example 2. This is not surprising because these features are not yet widely supported by mainstream or special browsers, or screenreaders, so the participants were unlikely to be aware of the features.

4.5.2.4.1 **Fieldset / Legend**

None of the participants mentioned the groupings of the controls in example 1. It is possible that the screenreaders or browsers read the Legends used in both examples, but that they did not explicitly announce the start and end of the Fieldset in example 1. This is probably because Fieldset is not yet supported by browsers or screenreaders.

4.5.2.4.2 **Tabindex**

The order in which the cursor moves to each control in example 2 was specified with the 'Tabindex’ attribute. Some participants found this order was not logical. Four participants reported that the Submit button was not the last control in example 2, but was followed by the control for 'Computer type’, which was confusing. This confirms that the Tabindex order overrides the 'visual’ presentation. Ideally, the Submit button should be the last or penultimate control on the form (perhaps followed by the Reset button) as these four participants pointed out.
This is a good example of how a page can be presented differently in different browsers. In this case, the sighted page author was presented with the form controls in a different order from that presented to those using assistive technology.

4.5.2.5 General comments

As well as making comments regarding the accessibility of the various form controls, participants also made general comments about the two examples.

4.5.2.5.1 Underscores

One participant mentioned that neither of the examples used underscores on the control labels. This is a convention used in the Windows operating system which helps users identify the keyboard shortcuts for particular commands. This participant has found underscores to be useful in the past, and thinks they would have improved the navigation of these examples. A similar comment was also made by one of the partially sighted participants who said that it is possible to present underscores on form controls, but that they don’t see it used often.

It is thought that in the future when the Accesskey tag is supported on form controls by browsers and screenreaders that the allocated letter will be underscored, following the Windows convention, and that this will be announced by screenreaders.

4.5.2.5.2 Confidence

Two participants were concerned that if they had to submit this form for real, it may not have contained the correct information (1 x LX + VE, 1 x IE3 + WE). This suggests that these participants felt they could not rely on the browser and the screenreader to present the information on the screen accurately.

4.5.2.6 Access to forms for partially sighted participants

One of the partially sighted participants reported that the size of the text in the text-entry fields was not enlarged along with the rest of the text on the page (IE5 font + contrast). This meant that the participant could not read either the default text, or the text they typed, and they would not have realised if they had made any typing errors.
The participant seemed to think that this problem was caused by the browser (IE5) but it is not clear why.

4.6 Frames

There were two examples of frames, which are presented in Appendix 8. The examples were evaluated by 18 participants.

4.6.1 Overview of differences between frames

In example 1 two frames divide the window horizontally. The top frame displays the text 'Papers on SDRU website'. The bottom frame contains three image-links to the papers with the phrases 'Haptic device', 'Talking books', and 'Auditory navigation' embedded in the images. These phrases were also used for the alternative text for the image-links. When a link is selected the relevant paper is displayed in the top frame. Both frames can be scrolled, both have 'frame names', 'top' and 'bottom', but neither frame has a 'frame title'.

In example 2 the window is divided into three frames. The top frame spans the whole window and displays the text, 'SDRU Papers available on the Web'. The other two frames divide the rest of the window vertically. The left-hand frame contains the phrase 'table of contents', followed by links whose text is the titles of papers. When a link is selected the relevant paper is displayed in the right-hand frame. The two lower frames can be scrolled, both have 'frame names', 'left' and 'main', and both have 'frame titles', 'contents' and 'papers'. The top frame can not be scrolled and does not have a 'frame name'. This example contains an 'Accesskey' on the first item in the table of contents. The font of all text is specified to be Arial, size 2.

Both examples provided 'Noframes' versions, which were available via a link in the 'Noframes' element. In example 1 the text of the link is the word 'here' ("click here to see a non-frames version") whereas the link in example 2 is 'table of contents' ("select to go to the table of contents").
4.6.2 Results on frames

4.6.2.1 Accessibility of Frames

Table 4.13, below, shows the number of participants who found each example the most accessible.

Table 4.13: Number of participants who found each example of frames to be the most accessible.

<table>
<thead>
<tr>
<th>Example</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>5</td>
</tr>
<tr>
<td>Example 2</td>
<td>12</td>
</tr>
<tr>
<td>All examples equally accessible</td>
<td>1</td>
</tr>
<tr>
<td>No examples accessible</td>
<td>0</td>
</tr>
</tbody>
</table>

4.6.2.1.1 Factors affecting accessibility

There was a trend towards significance for example 2 to be found the most accessible. Those who preferred example 2 found it easier to navigate and found the links more descriptive. Both partially sighted participants preferred example 2 because they preferred its layout and did not like the image-links in example 1. The participants who chose example 1 did so because they found it was read in a straightforward manner.

The preferences of the blind participants were not significantly related to the browser or screenreader in use: the participants in each group were using a variety of browsers and browser / screenreader combinations.

There was no significant relationship between the preference of each participant and their experience of using the Web (the average experience of those who chose example 1 was 3 years (median 3.5 years) and for example 2 the average was 2.4 years (median 2 years).
4.6.2.2 Comments on specific aspects of frames

4.6.2.2.1 'Noframes' option

Four participants mentioned the Noframes options in both examples. These participants were using browsers that do not support frames (2 x WS, 1 x LX, 1 x WWW). Participants using GUI browsers were unlikely to be aware of the Noframes options because these browsers support frames, and therefore do not present the Noframes option.

WS does not display frames but presents the user with links to each frame, and the Noframes option if one is present. One participant who used this browser said they liked having the option of either navigating the contents of the frames, or the Noframes version.

In contrast, another participant appears to have been confused by the way in which WS presented the Noframes version. From their comments it seems that they thought that example 2 did not have a Noframes version and that the link to the table of contents was not to a Noframes version. They said “I liked the explicit link in example 2 to [the] contents page, in preference to a Noframes version in example 1” (PBAa). It seems that WS does not explicitly announce which is the link to the Noframes version and which are the links to the contents of the frames.

LX and WWW are text-based browsers and were expected to present the Noframes versions to the participants. The participant who used LX found that it was easier to find the link in the Noframes version in example 2 than in example 1. This is probably because the participant was reading the page by jumping from link to link and therefore found that the link text ‘table of contents’ was more useful than the link text ‘here’.

The participant who used the WWW browser reported that this browser announced “this document is frames-enabled. You can use these links to view each frame” (KR). The participant was aware that the browser had re-formatted the frames and thought that this made the frames easier to navigate. The participant “didn’t have any problems” with either example of the frames (KR).
Neither of the partially sighted participants mentioned the Noframes versions. This is probably because they both used IE and were therefore unaware of the existence of the Noframes.

4.6.2.2 Table of Contents

Four participants made comparisons between the Tables of Contents in the two examples. Two observed that the links in example 2 contained more text than in example 1, but found the links in example 2 more descriptive. Another added that they preferred example 2 because the links in the Table of Contents were on separate lines.

4.6.2.3 General comments on frames

4.6.2.3.1 Layout of frames

Two participants commented on the difficulties that they generally have whenever they encounter frames. One said that it takes some time to build up a ‘mental picture’ of the relationship between the frames, and to understand why certain elements appear in particular frames. Another participant said “frames are always a problem because when you initially access an unfamiliar site you don’t have any idea of the layout” (CH).

4.6.2.3.2 Navigation

Two participants mentioned the using the tab key to jump between the links in the two examples. One found that in example 1 they had to tab past the links in the paper in order to get back to the Table of Contents links. However, once they were familiar with how both examples functioned, they found it easy to switch between the papers and the Table of Contents.

The other participant commented that they were able to use the tab key to jump between the links in example 2, but not in example 1. It is not clear why this problem occurred.
4.6.2.3.2.1 Frame names

One participant criticised both examples because although ‘frame names’ were provided they were very general, such as ‘top’ and ‘bottom’ (example 1) and ‘left’ and ‘main’ (example 2). The participant wanted information about the content of the frame, rather than just the physical position (IE401 and WB).

4.6.2.3.3 Use of screenreader and browser functions

The JFW screenreader has a ‘select all frames’ function which can be used to select one frame at a time and navigate each individually. One participant described how they used this function to read each frame. However, it was not necessary in example 1 because the frame containing the paper was read smoothly, followed by frame containing the table of contents.

WS re-formats pages containing frames, and provides a link to each frame with “link to a frame”. Two participants who used WS commented on the presentation of these links. One said that the links do not provide any information about the frames that they link to. Another said they would like to be informed of the number of frames before the links are read.

4.6.2.3.4 Reading the frames

Two participants described why they preferred example 1 to example 2. In example 2, two of the three frames are presented alongside each other. As describe above in the section on tables, many screenreaders tend to read the contents of the screen line by line, right across the screen, regardless of whether the window is divided into columns or frames. This means the screenreader may not have made a distinction between the two frames. Some participants may have found that example 2 was presented in this way. Two participants preferred example 1 because the top frame, containing the paper, was read first without interruption, followed by the bottom frame, containing the table of contents buttons.

Two participants commented on phrases that were repeated by the screenreader with example 2. One said that although they could smoothly navigate between the frames they found that the screenreader repeated the title of a paper three times: once when
the link was read, again when the link was selected, and finally when the paper was loaded and began to read (NS + JFW).

Another thought that the top frame of example 2 was redundant because the same text was repeated here and at the top of the Table of Contents frame. This was a decision made by the page author, who perhaps was not aware of the potential impact of such repetition.

This participant also reported that the alternative text for the buttons in example 1 was repeated, for example “button talking books, button talking books”. The word ‘button’ is not in the HTML code for the frame, but was probably inserted by JFW (or perhaps NS).

4.6.2.4 Access for partially sighted participants to frames

4.6.2.4.1 Image-links

Both partially sighted participants commented on the use of image-links in the Table of Contents in example 1. One said that the contrast between the text and the background was not high enough which made the text difficult to read. The other said that they could not read them at all. This participant said they would have found it easier to use the buttons if they had alternative text associated with them. Alternative text was provided, so it not clear why this participant could not read it. This participant said that they would normally use ‘hover text’ to read the alternative text provided by a page author, because hover text can be enlarged.

Both of these participants have their browsers set to display high contrast and large font. While the latter setting would enlarge any ASCII text on the page this would not apply to text that is embedded in a graphical image. Therefore the text embedded in the buttons in example 1 would remain the size specified by the page author.

66 Hover text, also known as ‘ToolTips’, is text that appears when the mouse pointer hovers over an item that has a label, such as an image with alternative text.
4.6.2.4.2 Layout and design

One of the partially sighted participants said they preferred example 2 because they are used to frames being laid out in this way. However, they felt that if the left-hand frame had been any narrower they might have had difficulty reading the Table of Contents because the font was enlarged.

The other partially sighted participant said that they do not usually like navigating in frames because they often do not allow for the display of larger fonts because the content becomes distorted. This in turn means that links are sometimes not available (although they were available in these examples).

4.6.2.4.3 Scroll bars and frame borders

One of the partially sighted participants had two difficulties with the top frame in example 2, although they found this to be the most accessible. They found that where the size of the font had been increased, it did not fit in the frame. This frame could not be scrolled because scroll bars were not provided.

This participant also said that, in general, they find it difficult to navigate in pages that are divided into more than three frames, particularly when the borders are invisible. The use of borders is optional but they were provided in both examples.

4.7. Imagemaps

Two examples of imagemaps were evaluated by 19 participants. Both examples contained the same image (see Appendix 8). In example 1 each section of the imagemap had alternative text: ‘blind section’, ‘deaf section’, ‘mobility section’, ‘low vision’ (without the word ‘section’), and ‘learning section’. Alternative text was also provided for the whole image, ‘Please select area of interest’. Text links were presented on one line beneath the imagemap. The text of the links was the same as that used for the alternative text for the sections of the imagemap.

In example 2 the sections of the imagemap had alternative text: ‘blindness’, ‘deafness’, ‘mobility impairments’, ‘low vision’, ‘learning impairments’. These
phrases were repeated in the text links, which formed a vertical list beneath the imagemap. Alternative text was not provided for the whole image.

4.7.1 Results on Imagemaps

4.7.1.1 Accessibility of imagemaps

The number of participants who found each example the most accessible is shown in Table 4.14, below.

Table 4.14: Number of participants who found each example of imagemaps the most accessible.

<table>
<thead>
<tr>
<th>Preferred example</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>4</td>
</tr>
<tr>
<td>Example 2</td>
<td>9</td>
</tr>
<tr>
<td>Undecided</td>
<td>1</td>
</tr>
<tr>
<td>All examples equally accessible</td>
<td>5</td>
</tr>
<tr>
<td>No examples accessible</td>
<td>0</td>
</tr>
</tbody>
</table>

In general, those that found example 1 to be the most accessible did so because it was perceived to be more simple, and because they did not like the list of links in example 2. Those that found example 2 to be the most accessible did so because the text links were in a list, and the items were on separate lines. This latter group includes the two partially sighted participants. Those that found the two examples equally accessible had no difficulties with either example.

There was a trend towards significance in the relationship between the preference of the participants for example 2 and the use of a GUI browser (Chi square = 3.25, df = 1, significance p=0.071). Of the four participants who preferred example 1, three used text-based browsers, the other used a GUI browser. Of the nine who preferred example 2, only one used a text-based browser, while the others all used GUI browsers. The two types of browser may have presented the text links differently and this may account for the differences in preference.
There was a difference in the average experience of using the Web between those who preferred each example: for example 1 it was 1.8 years; for example 2, 2.7 years. However, there was no significant relationship between preference and experience.

The blind and partially sighted participants seem to have accessed the examples in different ways. The blind participants used the keyboard to follow the text links, whereas the partially sighted participants used the mouse to click on the imagemap. Because these two groups were interacting with the imagemaps in different ways, they had different requirements for the design of imagemaps.

4.7.1.2 Comments on imagemaps

Participants made comments regarding the presentation of the imagemaps, the alternative text and the text links. None of them commented on the content of the alternative text.

One participant found they were not informed of the presence of the imagemap on the page (LX and IE401 + JFW). The participant said this was not a problem because the text links could be used. However, imagemaps are no always accompanied by text links, so in other contexts this participant may experience have difficulty accessing imagemaps.

4.7.1.2.1 Alternative text

One participant found the alternative text for the image in example 1 was incomplete (NS + JFW). They reported that JFW announced “elect area of in” and “terest” [sic]. It is not clear why the phrase was divided into two parts and the initial letter ‘S’ was missed. Perhaps the alternative text was wrapped onto two lines and the screenreader read each line separately. The letter ‘S’ may not have been announced clearly by the speech synthesiser (Keynote Multimedia).

Another participant reported that they were not informed of the alternative text for the different areas of either imagemap (IE3 + DN). Again, it is not clear why this occurred.
One of the partially sighted participants was able to access the alternative text of the different sections and found it useful in identifying the purpose of the imagemap links. In contrast, the other partially sighted participant found they could not determine what the sections of the imagemap represented and could not click on them. It is assumed that this participant could not access or use the alternative text of the map sections.

4.7.1.2.2 Text links

The participants' were split in opinion over the presentation of the text links in the two examples. Four participants liked the use of a list in example 2 because each item appeared on a separate line, which was clear. These participants were using GUI browsers (mainly IE). In contrast, two participants found this presentation was repetitive and slowed down the navigation because the phrase “list item” was announced prior to each item. These participants were both using WS, a text-based browser. The GUI browsers and screenreaders presented the list of links in different ways. It seems that WS informs the user of the underlying element, for example it announces “list item”, whereas a screenreader just reads the text without separating them.

One participant found that they could not find the text links with the ‘JFW cursor’ but could find them when they used the tab key to jump from link to link (IE401 + JFW). This means that the screenreader mode also affects the presentation of links.

4.7.1.3 Suggestion

One of the partially sighted participants suggested that it would be useful to be told that the text links below the imagemap are alternatives to those in the imagemap.

4.8 Headers

Four examples of headers were evaluated by 18 participants (see Appendix 8). The content of each example was the header of an academic paper with a title, multiple authors, and their contact details. The other examples include the same information,
but the title, the authors' names and institutions, telephone numbers and email addresses are presented differently.

4.8.1 Overview of differences between headers

Examples 1 and 3 list the names of all the authors together with superscript numbers next to each name to indicate the institutions of the authors. The other examples (2 and 4) list the authors with their institutions so superscript numbers are not required. In examples 1, 2 and 3 the email addresses of the authors are condensed. In example 4 the email addresses are given in full, and are 'mailto' links. Example 3 includes a set of links that form a table of contents, which includes all the headings in the paper. Examples 1, 2 and 3 contained a number of left and right brackets around the email addresses.

The text of all the examples is centred on the page, except example 2 which is left-justified. All the examples have a title that is marked-up as the largest HTML heading, H1, except example 2 whose title is not a heading. In example 4 the telephone numbers are preceded by the word 'telephone'.

4.8.2 Results on Headers

4.8.2.1 Accessibility of Headers

Table 4.15, below, shows the number of participants who found each example of headers the most accessible.

4.8.2.1.1 Factors affecting accessibility

In general, those who preferred example 2 were the partially sighted participants who found it easier to see, and a blind participant who found the presentation of the authors' details useful. Those who preferred example 3 liked the table of contents and its links. The participants who found example 4 the most accessible did so because it was clearer, it included the word 'telephone', and did not include superscript numbers.
Table 4.15: Number of participants who found each example of headers the most accessible.

<table>
<thead>
<tr>
<th>Example</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>0</td>
</tr>
<tr>
<td>Example 2</td>
<td>3</td>
</tr>
<tr>
<td>Example 3</td>
<td>4</td>
</tr>
<tr>
<td>Example 4</td>
<td>8</td>
</tr>
<tr>
<td>Undecided</td>
<td>1</td>
</tr>
<tr>
<td>All examples equally accessible</td>
<td>2</td>
</tr>
<tr>
<td>No examples accessible</td>
<td>0</td>
</tr>
</tbody>
</table>

The preferences of the blind participants were not significantly related to the browser or screenreader they used: the participants in each group were using a variety of browser types and browser / screenreader combinations (except two of the WS users who found them all equally accessible).

The average experience of those who chose example 3 was 1.8 years (median 1.25 years) and for example 4 the average was 2.9 years (median 2.8 years). However, there was no significant relationship between the preference of each participant and their experience of using the Web.

4.8.2.2 Comments on headers

4.8.2.2.1 Example 3: Table of Contents

Four participants commented on the table of contents in example 3. One said it was useful because it provided information about the contents of the paper. Another said that the links were useful when they used the tab key to jump to each link on a page they got an overview of the contents of the page.

4.8.2.2.2 Example 4

Two participants described why they preferred example 4 to the other examples. One said it was the most clear: the different bits of information were clearly separated; and the email addresses were ‘mailto’ links. Another said the layout made it easier to follow. This latter point is interesting because the visual layout of
examples 1, 3 and 4 is similar, with example 2 having a different layout (left-justified). Another participant found example 4 to be more verbose.

4.8.2.2.3 Reading

One participant felt that they had read a whole paper by the time they had finished reading just the header. This was because WS read it a line at a time, which took a long time and seemed like a lot of information. Another said that they found most of the information to be rather ‘mixed up’ (IE401 + JFW).

One participant commented that the authors’ names and the different parts of their addresses were not clearly separated in examples 1 and 3 when read in synthesised speech. This made the information difficult to assimilate, unless it was read word by word. The reason for this is that there are few commas or full stops to create natural pauses in the synthesised speech. Although the text is visually divided onto separate lines, screenreaders and special browsers will not necessarily pause at the end of each line.

4.8.2.2.4 Characters perceived to be superfluous

Several participants commented on the presentation of the superscript numbers and the brackets around the email addresses. Six participants found that their screenreaders read the numbers as one line, followed by the line containing the author names, for example,

1, 1, 1, 2, 2, Chetz Colwell, Helen Petrie, Diana Kornbrot, Andrew Hardwick & Stephen Furner.

In this case the purpose of the superscripts numbers was lost. Two of these six participants described the presentation of the superscript numbers as “spurious” (PBa) and “unnecessary” (JJ). Others described the repetition as “terrible” (DG, PBb). A further participant commented that the numbers increased the ‘cognitive load’ because it was necessary to navigate backwards and forwards to understand the purpose of the numbers.
One participant commented on the brackets around the email addresses. They thought that their presence was the result of bad HTML coding but, in fact the brackets were included intentionally.

4.8.2.2.5 Email addresses

One participant commented that they did not like the fact that the email addresses of the authors were not given in full in examples 1, 2 and 3. Another participant liked the way that their screenreader (WE) preceded the email addresses in example 4 with the word 'link'. This was because they were 'mailto' links.

4.8.2.2.6 Telephone numbers

Two participants commented on the presentation of the telephone numbers. One did not know whether the number given in examples 1, 2 and 3 was a telephone or fax number. Both participants liked the use of the word 'telephone' in example 4.

4.8.2.3 Comments from partially sighted participants

The two partially sighted participants found example 2 to be the most accessible. These participants raised different issues to those raised by the blind participants. One partially sighted participant preferred example 2 because the layout was more simple than in the other examples. They said that the left-justified text of this example was easier to read than the centred text of the other examples, because each line began at the same point.

The other partially sighted participant found example 2 was easier to see. This participant preferred the heading in example 2 because in the other examples the headings were too big. This is an interesting issue because the headings in examples 1, 3 and 4 were marked-up as the largest heading, H1, which is intended to be used for the first heading of a page. In example 2 the text of the heading was given no additional mark-up and so appeared as normal text. It is possible that when the text of an H1 heading was enlarged by the settings of the participant's browser it became too large that to be easily read. In this case it seems that the participant's difficulty was caused by the 'correct' implementation of HTML.
One of the partially sighted participants said that, in general, they do not find the use of font styles, such as italics, to be particularly helpful. It is assumed this is because they are less clear. Italics were used in example 3 in the table of contents. This participant said that although it is possible to set IE401 to ignore such specifications made by the page author, they had not done so in this case.

4.9 References

Three examples of references were evaluated by 18 participants (see Appendix 8).

4.9.1 Overview of differences between references

In example 1 the reference was presented as a link ‘[10]’. Example 3 also contained a link, with the phrase ‘reference 10’. Example 2 did not have a link.

4.9.2 Results on References

4.9.2.1 Accessibility of References

Table 4.16 below, shows the number of participants who found each example the most accessible.

<table>
<thead>
<tr>
<th>Number of participants</th>
<th>Example 1</th>
<th>Example 2</th>
<th>Example 3</th>
<th>All examples equally accessible</th>
<th>No examples accessible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>4</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

In general, participants found examples 1 and 3 the most accessible because they contained links. The participants who found example 3 the most accessible did so because they liked the word ‘reference’ in the link. In contrast, those that preferred example 1 did so because they found the word ‘reference’ made the link too long. Of the partially sighted participants, one preferred example 1, the other example 3.
The preferences of the blind participants were not significantly related to the browser or screenreader they used: the participants in each group were using a variety of browser types and browser/screenreader combinations. There was no significant relationship between the preference of each participant and their experience of using the Web, or the use of a braille display.

4.9.2.2 General comments on references

Several participants made comments about the presentation of the references.

4.9.2.2.1 Reading

Three participants reported that the text of the reference was not read by the screenreader after they had selected the links in examples 1 and 3. These participants were using a variety of browsers and screenreaders. One participant, who used WB, thought that this screenreader does not support links that go to another location within the same page (jumps). This participant reported that when they initially selected the link they had to switch screenreader modes and explore the screen, and only then did they discover that they were on the same page, but in a different position.

Another participant, who used JFW, found that when they selected the link, the application cursor moved to the correct position, but JFW began to read from the top of the page again.

4.9.2.2.2 Navigation

Most participants preferred the link text in example 3, ‘reference 10’ because it was easier to find amongst the surrounding text, and was more meaningful. Also, a partially sighted participant found it was easier to click with the mouse because it was a larger target. In contrast to these positive comments, four participants did not like the word ‘reference’ because it made the link too long and was ‘useless’.

Four participants commented on the fact that example 2 had no link for the reference. Three participants found this presentation could not be read efficiently, particularly in a longer paper, because the absence of links would mean the page would have to
be read sequentially. One of the partially sighted participants said “imagine having to find your way through 20 references like that” (MTa).

Another navigation issue raised by three participants was a difficulty in getting back to the link in the paragraph after selecting it and reading the reference. It is not clear whether these participants used the ‘back’ command of their browsers as this should have taken them back to the link they selected. It may be that if the participants used this command, the screenreaders did not go back to the link, but went to the top of the page.

4.10 Images

Three images were evaluated by 15 participants (see Appendix 8). These are referred to as figures 1, 2 and 3. There were several different examples of each image: six of figure 1, and three each of figures 2 and 3.

4.10.1 Overview of differences between images

All the examples of each figure had alternative text, except example 5 of figure 1. Example 6 of figure 1 included a caption and example 2 of figure 1 included a ‘D’ link.

The alternative text used for the different examples of figure 1 was as follows.

- Example 1: “Image of the Impulse Engine 3000”.
- Example 2: “Graphic of The Impulse Engine 3000”, plus D-link to the description “this is a graphical representation of the Impulse Engine 3000”.
- Example 3: “Picture depicting engine of type impulse 3000”.
- Example 4: “The Impulse Engine 3000”.
- Example 5: No alternative text provided.
- Example 6: “Diagram of an Impulse Engine”, plus the caption “Figure 1: The Impulse Engine 3000”.

The alternative text used for the different examples of figure 2 was as follows.

- Example 1: “Image illustrative of the textures’ dimensions”.

180
• Example 2: "A diagram depicting visually grooves sizes and amplitudes stated in
document".

• Example 3: "Dimensions of sinusoidal grooves used in Experiment 1".

The alternative text used for the different examples of figure 3 was as follows.
• Example 1: "Illustration of the two perceived locations of the virtual shapes".
• Example 2: "Picture of the hands".
• Example 3: "Different mental models of the location of the virtual object in
virtual space".

4.10.2 Results on Images

4.10.2.1 Accessibility of Images

The number of participants who found each example of the three figures the most
accessible is shown in Table 4.17.

Only seven participants could access all of the examples of all of the figures (2 x
WS, 2 x IE401 + JFW, 1 x IE401 + WB, 1 x NS + JFW, 1 x BS and IE3 + DN). Six
of the 15 participants could not access any of the examples of any of the figures (1 x
WS, 2 x IE3 + WE, 2 x IE401 + JFW plus braille display, 1 x IE401 font and
contrast). In addition two participants who could access figure 1, could not access
figures 2 and 3 (1 x IE5 + JFW, 1 x WWW + SR2).

Table 4.17: Number of participants who found each example of the three figures
the most accessible.

<table>
<thead>
<tr>
<th>Example</th>
<th>Figure 1</th>
<th>Figure 2</th>
<th>Figure 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Example 2</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Example 3</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Example 4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 5</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example 6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No examples accessible</td>
<td>6</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>All examples equally accessible</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Undecided</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
While there were two participants who could not access any of the images (IE401 + JFW), and one participant who could only access one of the three figures (IE5 + JFW), there were also two who used the same browser and screenreader who could access all of the figures. Only one of the two participants who used text-based browsers evaluated the images, but this participant could not access figures 2 or 3. This is surprising as it was expected that the users of text-based browsers would have been able to easily access the alternative text.

The accessibility of the images to these participants was not significantly related to the particular browser and screenreader used, or the participants’ experience of using the Web. The average amount of experience of those that either found none of the figures accessible or could only access one of the three figures (n=8) was 2.1 years (median = 1.5) whereas the average for those that could access all the images (n=7) was 4.4 years (median = 4).

Another unexpected finding was that one of the three participants who evaluated the images with WS could not access any of the examples. Again, it was expected that the alternative text would be accessible to WS users.

4.10.2.2 General comments on images

Several participants made comments on the examples of images regarding the content of the alternative text, and the D-link. There were no comments regarding the caption on example 6 of figure 1.

4.10.2.2.1 Reading the alternative text

One participant found that it was often difficult to know whether text read by the screenreader was simply part of a paragraph, or was alternative text for an image (IE401 + WB). It is possible that WB does not announce the presence of an image, but just reads any alternative text as if it was ordinary text.

Another participant reported a similar experience. They thought they could not find the images on the page because the browser was set to display only those images that had labels (IE401 + JFW). When they changed the setting to display all images, the
screenreader announced “example 1, example 2” etc. This is not how the researcher expected the pages to be read. The words ‘example 1’ etc were marked-up as headings and were expected to be read regardless of how the screenreader was set up.

4.10.2.2 Content of alternative text

One participant said that they did not think that the alternative text for any of the examples of figure 2 was useful. They found that the accompanying paragraph provided sufficient information about the image, and thought that the alternative text could have been “more creatively descriptive” (PBb).

One participant said that they liked descriptions that expanded on the text in the accompanying paragraph, rather than repeating it. They also said that if a diagram “merely illustrates what is said in [the] paragraph, then a brief description of the object is sufficient” (NO2).

Another participant liked phrases that were concise and accurate, rather than those that tried to explain the visual nature of the image. A further participant thought that the meaning of an image should have been clear from the accompanying paragraph. However, they did not say what they thought the alternative text should include in this case. Another felt that their command of the English language was not good enough to judge the importance of the differences between the descriptions.

The comments on the content of the alternative text are very general and do not indicate why the participants preferred one example to another. As the study conducted by Leonard Kasday (described in Chapter 2, Section 2.7) suggests, the issue of the content of alternative text is complex and there is little consensus.

4.10.2.2.3 D-link

The participants who chose example 2 of figure 1 commented on the D-Link. One found that this was the only example of figure 1 that they could access. Another participant seemed to think that the use of a D-link was the only possible method for accessing the alternative text of an image. They used the D-link in example 2 of
figure 1, but could not find these links in the other examples. It is assumed that this participant was using the tab key to jump between the links on the page and this is the only link they found.

Neither of these participants mentioned the content of the D-link description. The description was rather short, "this is a graphical representation of the Impulse Engine 3000". Ideally a D-link would link to a longer description of the visual content of the image, or include any information contained in the image which is vital to understanding the page.

4.10.2.3 Comments from partially sighted participants

One of the partially sighted participants made comments on the images that were different to those of the blind participants. This participant found that they could not see the images well enough to know what they represented. They said that if the contrast between the fore- and back-grounds was higher, such as a white image on a black background, they might be able to see the image more clearly. In general they find that text embedded in images is difficult to see.

4.11 Discussion on accessibility of examples

Previous studies of guidelines conducted by Tetzlaff & Schwartz (1991) and de Souza & Bevan (1990), found that although the participants experienced difficulties with using the guidelines, their designs conformed to the styleguide (Tetzlaff & Schwartz), or were judged to be 'satisfactory' in expert analysis (de Souza & Bevan). The results of this series of two studies are similar. The page authors in the first study had difficulty in navigating in the Guidelines documents, and applying the guidelines, but the participants in this second study found most of the examples accessible to some extent.

4.11.1 Discussion on Tables

Five examples of tables were evaluated. Those participants with less than two years' experience of using the Web were significantly more likely to find the examples inaccessible. The task was to interrogate each example to find the answer to a set
question. Participants reported difficulties identifying the headers of the rows and columns associated with the data cells. The results indicate a wide range of ability across the participants.

A key issue highlighted by the participants’ reports is the way in which tables are presented by special browsers, and screenreaders used with mainstream browsers (GUI or text-based). Special browsers interpret the HTML and present the elements one at a time. Screenreaders generally present the information line by line, as it appears on the screen. The participants who used screenreaders found that the headers of the tables were not presented in a useful way: often words in header cells that wrapped onto more than one line were read separately. The participants who used special browsers found that the header cells were read as one unit. Both groups had to memorise the relative positions of the headers in order to find the data cells in the associated rows or columns. These problems were somewhat overcome in example 3, the layout of which made it easier to read, particularly with a braille display. There was a trend towards significance for those using GUI browsers and screenreaders to prefer example 3. However, the participants who found this example to be the most accessible, generally had lower scores than the other participants.

This situation was expected by the researcher, but not by the WAI Page Content Guidelines Working Group. The researcher raised the issue of IDs and Headers not being supported by current browsers and questioned why the use of paragraph breaks was no longer recommended by in the Guidelines. One respondent stated that it would be surprising if people were still using browsers that read tables rows as one line. This suggests that some members of the Working Group do not know how current browsers present information.

In the first of this series of two studies nearly all the participants (page authors) used version 4.0 of either Internet Explorer or Netscape. In this second study about three quarters of the participants used Internet Explorer 4.0 or later, Netscape 4, or another mainstream browser. If authors and users are often using the same browsers then the issue of tags being supported by user’s browsers but not page author’s browsers is not so relevant.
The needs of partially sighted people regarding tables seem to be quite different from those of blind people. For the blind participants the important factors were reading the headers in a manner which made sense, and associating the data with the headers. For the partially sighted participants the important factors were consistency in the width of cells, and whether the overall width of the table required them to scroll the page.

4.11.2 Discussion on Forms

Two examples of forms were evaluated. Two participants, one of whom was partially sighted, found neither example accessible because they were not presented in a useful way. The preferences of those who could access the forms differed and seem to be related to the browser used. The different types of browser seem to present the controls of the forms in different ways, particularly when the participants were jumping from one control to another. Those using GUI browsers found example 1, which did not contain a Tabindex, more accessible because the Tabindex in the example 2 was incorrect. The incorrect Tabindex was not a problem to those who preferred example 2 because they were using text-based or special browsers and therefore were not using the tab key. This example was also preferred because the control and its instruction were presented together. Those that used a braille display also preferred example 2.

The partially sighted participant who was able to access the forms had different concerns to the blind participants. This participant found that the default text in the text fields was not enlarged along with the rest of the text. It is important that future versions of mainstream browsers allow default text to be enlarged along with the rest of the text.

Although the examples contained several accessibility features, the participants made few comments on these. This is probably because the browsers and screenreaders support these features and so the participants were not aware of their presence. The use of such features needs to be evaluated when they are supported by most browsers.
Example 2 included a heading: 'For easy navigation MS Internet Explorer is better'. The design of pages for specific browsers is a contentious issue. It was therefore expected that some participants would comment on this, but surprisingly, none did so.

In summary, the results suggest that it is important for users of text-based browsers and users of braille displays that form controls and their instructions are located on the same line. For users of GUI browsers the Tabindex order is important.

### 4.11.3 Discussion on Frames

Two examples of frames were evaluated. Both examples provided the same functionality, but used different layouts. The frames were accessible to all those who evaluated them. Some participants preferred example 1, in which the screen was divided horizontally, because their screenreaders were able to read across the screen. However, there was a trend towards significance in the participants' preferences for example 2, in which the screen was divided vertically. From participants' reports the researcher has gained insight into the methods employed and the experiences of the participants, and how these experiences differ depending on the browser or screenreader used, and whether the participants' sight status.

The blind participants who were using browsers that display frames, such as IE and NS, needed to formulate a 'mental image' of the frames and their functionality. They probably did this by using the Tab key to jump from link to link, working from frame to frame. These participants may have had difficulties with their screenreader reading across the screen, regardless of the frame boundaries in example 2.

Those participants who were using JFW with IE (the former being designed to work well with the latter) could have used the 'select all frames' function of JFW to navigate each frame. This would have assisted in the development of an understanding of the layout of the frames, particularly in example 2.
The participants who used browsers that do not display frames, such as WS, LX and WWW (all text-based), used different methods to those who used GUI browsers. The participants using text-based browsers may not have found it necessary to develop an idea of the layout of the frames in order to navigate. WS re-formats frames and gives the user the choice of navigating each frame individually, or using the Noframes version. However, users are not necessarily informed of the number of frames present.

Participants who used GUI browsers were presented with the frames and not given access to the Noframes, whereas those using WS were given the choice of navigating the frames or the Noframes versions. The participants using LX and WWW (text-based browsers) were presented with the Noframes versions, because these browsers do not display frames. The link to the Table of Contents had different link text in each example. If the participants were jumping from link to link, without necessarily reading the surrounding text, they probably found that of example 2 more useful. In example 1, when they heard the word 'here' they would have had to change screenreader modes to read the surrounding text.

All the frames, except the top frame of example 2, had ‘frame names’ which indicated the frames’ physical location. It may have been more useful if the ‘frame names’ indicated the content of the frames.

The concerns of the partially sighted participants were different to those of the blind participants regarding frames. The partially sighted participants found the image-links in example 1 difficult to use because the contrast was not high enough between the text and the background in order to read the text. Like their blind counterparts, they relied on the alternative text in order to know the purpose of the links. For these participants the presence of scroll bars and frame borders was also important. They were not able to read enlarged text within frames that did not have scroll bars because the text was too big. Also, they relied on frame borders to separate pieces of text that might otherwise run into each other.
4.11.4 Discussion on Imagemaps

Two examples of imagemaps were evaluated. All the participants could access both examples, although some had difficulty in reading the alternative text of the whole image. There was a trend towards significance for users of GUI browsers to prefer example 2. It seems that the screenreaders used with GUI browsers did not announce the presence of a list in example 2, and announced the links separately because they were on separate lines. This was preferred over the way in which the screenreaders did not separate the links that were all on one line in example 1.

The other participants who preferred example 1 in which the links were presented on one line were mostly using non-GUI browsers that presented the links separately. This was preferred over the way in which these browsers repeated the phrase 'list item' for each item of the list in example 2.

None of the participants commented on the content of the alternative text associated with the different sections of the imagemap, or on the text of the links (which contained the same text). It is therefore assumed that none of the participants noticed the different phrases 'blindness section' and 'blindness', or that this difference did not affect their comprehension of the example.

The two partially sighted participants had different experiences from each other with the imagemaps. One was able to use the alternative text to identify the map sections, whereas the other found they were not able to distinguish between the sections and therefore could not use the links.

4.11.5 Discussion on Headers

Four examples of headers were evaluated. Most participants preferred example 4 because it presented the authors' details together, it did not use superscript numbers, and it included the word 'telephone'. Other participants preferred example 3 because it provided links in a table of contents. The partially sighted participants commented on the size of the headings, the use of italicised text, and the justification of the text.
Some of the presentational aspects of the headers involved the contraction of text, and other ‘visual’ conventions used in academic papers. These have been found to cause difficulties for the blind participants. For example, the superscript numbers caused confusion. Also, the contraction of the email addresses was not useful. Participants found the use of the word ‘telephone’ useful for the identification of the string of numbers.

4.11.6 Discussion on References

The three examples of references were evaluated. All the participants who evaluated the references could access them all, but they had different preferences. Most participants preferred example 3, while others preferred example 1. These preferences are related to the presentation of the links. Most participants preferred the examples that contained a link. Some liked the word ‘reference’ in the link whereas others did not. The opinions of the two partially sighted participants were also divided: each preferred a different example, for the same reasons as the blind participants preferred each example.

The issues raised by the participants regarding the presentation of the references could be classified as usability issues rather than accessibility issues. The participants were able to access the information within the examples and commented on reading and navigating the information.

4.11.7 Discussion on Images

Several examples of three different images were evaluated. The most surprising result was that so many participants could not access the alternative text of the images. This is both surprising and worrying, because there is an apparent assumption in both the Guidelines and the discussions of the WAI that if alternative text is provided users can access it.

Of the four participants who used IE401 + JFW, two were able to access all the figures and two were not able to access any. This suggests that the results on images may be due to the abilities of the participants, rather than the capabilities of the technology, or the design of the Web page. It may be the case that the alternative
Few of the participants mentioned the use of the D-link in example 2 of figure 1. This suggests that the others may not have noticed its presence, or may not have realised its function. A major problem with the use of D-links is that users need to be informed of their purpose. It is interesting that two of the three participants who mentioned the D-link said that they found this was the only way of accessing the image. Since none of the participants commented on the content of the D-link description, this is something that should be investigated further.

Only one of the partially sighted participants was able to access the images. Their comments were different from those of the blind participants. As with the image-links in example 1 of the frames, the partially sighted participant could not access the text embedded in the figure 2 because of a lack of contrast between the image and the background.

Unfortunately, the participants’ comments on the content of the alternative text do not provide much insight into why one was preferred over another.

4.11.8 Page authors’ design decisions

As described in Chapter 3 some of the page authors in the first of these two studies were observed to make design decisions which were not specifically recommended by the Guidelines, but were based on information provided by the Guidelines. When these pages were evaluated in the current study there were three cases in which these decisions were found to increase the accessibility of these elements for participants.

For example 3 of the tables, the page author split the table because they thought it would be easier for users of screenreaders to read (and for sighted people). This decision was based on the description in the Guidelines of how screenreaders work rather than a specific recommendation. The visually impaired participants found this example easier to navigate.
For example 1 of the frames the page author created a horizontal layout because they thought it would be read more easily by a screenreader, not because this was recommended.

For example 4 of the headers the page author preceded the telephone numbers with the word 'telephone' because they thought it was important to indicate the purpose of the number, not because the Guidelines had recommended this.

4.11.9 Inaccessibility of the examples

There were participants who could not access any of the examples of three elements: the tables, images or forms. So far, the discussion has focused on the accessibility of different elements, but this can be extended to the accessibility for individual participants. There were some participants who found more than one element inaccessible. Three participants found both tables and images to be inaccessible, and one found both forms and images inaccessible. These four participants did not use the same browsers or screenreaders, but are similar on other levels. They have all been using the Web for 18 months or less, and all except one use it less than 30 times per month. However, there was no significant relationship between participants' preferences and their experience, except for the tables, where those with less experience were more likely to find none of the examples accessible.

4.11.10 Assessment of examples by Bobby

The pages evaluated by the participants have also been checked using the Bobby accessibility checker (described in Chapter 1, Section 1.3.3.7.1). This provides a different perspective on the accessibility problems identified by the participants. Each example was checked, and Bobby returned reports for the different levels of priority (1, 2 and 3). These are summarised in Table 4.18 below. In some cases Bobby recommends manual checks of potential problems that cannot be detected in an automatic process. For example, the page author is recommended to check that keyboard shortcuts are provided for 'frequently used' links, that a linear text version of tables is provided, and that there is sufficient contrast between the background, text and images.
The Bobby checker highlighted several accessibility problems with the examples. Some of these problems were the same as those highlighted by the participants, other problems found by Bobby were not mentioned by the participants.
<table>
<thead>
<tr>
<th>Priority</th>
<th>Priority 2</th>
<th>Priority 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headers not properly nested.</td>
<td>Headers not properly nested.</td>
<td>From one frame.</td>
</tr>
<tr>
<td>Frames 2</td>
<td>Frames 2</td>
<td>Frames 2</td>
</tr>
<tr>
<td>Frames 2</td>
<td>Frames 2</td>
<td>Frames 2</td>
</tr>
<tr>
<td>Adjacent links not separated.</td>
<td>Adjacent links not separated.</td>
<td>From all frames.</td>
</tr>
<tr>
<td>Frames 1</td>
<td>Frames 1</td>
<td>Frames 1</td>
</tr>
<tr>
<td>Frames 1</td>
<td>Frames 1</td>
<td>Frames 1</td>
</tr>
<tr>
<td>Frames 1</td>
<td>Frames 1</td>
<td>Frames 1</td>
</tr>
<tr>
<td>Frames 1</td>
<td>Frames 1</td>
<td>Frames 1</td>
</tr>
</tbody>
</table>

**Table 4.18: Conformance of Examples to Bobby’s Protocols.**

- **Bold text in Examples column indicates examples found to be the most accessible by a majority of participants.**

- **Example body appears at that priority level.**

- **Key:**
  - Frames 1
  - Frames 2
  - Frames 3
  - Frames 4
  - Frames 5
  - Frames 6
### Note for Table 4.18: Under priority 3 Body highlighted when the language was not specified in any of the pages.

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Figure 1</th>
<th>Figure 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Figure 1</th>
<th>Figure 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>ImageMap 2</th>
<th>ImageMap 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Form 2</th>
<th>Form 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Frames 2</th>
<th>Frames 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Frames 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Conforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Conforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Conforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortcuts for Frequently used Links?</th>
<th>Conforms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternating text</td>
<td></td>
</tr>
<tr>
<td>Reference 3</td>
<td></td>
</tr>
<tr>
<td>Reference 2</td>
<td></td>
</tr>
<tr>
<td>Reference 1</td>
<td></td>
</tr>
<tr>
<td>Reference 4</td>
<td></td>
</tr>
<tr>
<td>Reference 5</td>
<td></td>
</tr>
</tbody>
</table>
4.11.10.1 Agreement between Bobby and participants

The following problems were highlighted by the participants and Bobby.

- Frames: Missing frame titles (both examples), link text does not make sense out of context (example 1), lack of contrast on image links (example 2).
- Imagemaps: No separation of adjacent links (example 1).
- Forms: Control and instruction are not on the same line (examples 1 and 2), missing or incorrect Tabindex (both examples).

4.11.10.2 Disagreement between Bobby and participants

The researcher has identified several instances in which the examples the participants found to be the most accessible do not achieve the highest conformance to Bobby. These instances are listed below.

- Forms: Preference was equally split, although Bobby found more problems with example 2.
- Frames: Most participants found example 2 the most accessible, but Bobby found similar numbers of problems with both examples.
- Imagemaps: Most participants found example 2 the most accessible, but Bobby found more problems with this example than example 1.
- Headers: Most participants found example 4 the most accessible, but Bobby found no more problems with this example than with examples 1 and 3.

The following problems were highlighted by Bobby, but not by the participants.

- Forms: Headings not properly nested (example 2), no default text in text entry fields (both examples), shortcuts not provided for form controls (both examples), controls not grouped (example 2), labels not associated with controls (both examples).
- Frames: The provision of shortcuts for frequently used links (both examples), headings not properly nested (example 2).
- Imagemaps: shortcuts not provided for frequently used links (both examples), no alternative text on hotspot (example 2).
- Headers: shortcuts not provided for frequently used links (example 2).
• Images: shortcuts not provided for frequently used links (example 2 of figure 1 (with the D-link)), no alternative text (example 5 of figure 1).
• Tables: No linear text alternative for tables with parallel word-wrapped columns (all examples).
• All elements: Language not specified.

The following problems were reported by participants or observed by researcher, but not highlighted by Bobby.
• Headers: use of superscript numbers (examples 1 and 3).
• Tables: it is not easy to read the rows and columns (examples 1, 2, 4, 5).
• Frames: there is no access to the Noframes version for users of GUI browsers (both examples).
• Forms: default text in entry fields is not enlarged with the rest of the text with browser settings (both examples).

It is clear from the above lists that there are differences between the opinions of the participants and the Bobby testing results regarding the accessibility of the elements. Several problems were found by the participants but not reported by Bobby. There are several reasons for this. In the case of superscript numbers this was not highlighted by Bobby because it is not categorised as an accessibility issue by the WAI and developers of Bobby. If the use of superscript numbers was to be identified as an accessibility issue, it could be identified by automatic checking.

In the case of the tables, although Bobby found that all the tables conformed to priority 1, the participants did not find them entirely accessible. In the case of the frames, Bobby did not recognise that GUI browsers cannot access the Noframes versions.

There were also several problems found by Bobby that were not identified by the participants. The participants probably did not mention the lack of keyboard shortcuts (for links and form controls). This may be because they did not know this possibility existed.
The researcher is surprised that Bobby classifies the document language as an accessibility issue. Only one of the participants mentioned that their first language was not English, but they did not report difficulties with the fact that the language of the documents was not specified.

4.11.10.3 Limitations of Bobby

There are inherent problems with any automatic checker, as discussed in Chapter 1, Section 1.3.3.7.1. However, the researcher has identified other limitations with Bobby that are not related to its automation, but to a lack of definition of the terms it uses. These are listed below.

- Images: Bobby states that Longdesc or D-links should be provided for important images, but does not define 'important', just gives examples such as chart and graph.
- Links: Bobby states that shortcuts should be provided for 'frequently used' links, but does not define 'frequently'.
- Tables: Bobby states that 'lengthy' headers should be abbreviated, but does not define 'lengthy'.

4.11.11 Assessment of examples by W3C Validator

For completeness the researcher checked the examples with the W3C validator\(^{67}\). The validator identified various syntax errors in the HTML code that had not been found by the participants or Bobby. The validator identified some problems that were also found by Bobby, for example, where alternative text was missing, and where Fieldset was not used. It is not relevant to discuss all the problems found by the validator here, but there are two points that should be explored further.

The validator reported that cells in tables cannot be associated with more than one other cell. It also stated that variables in the Headers and ID tags could not start with the digit ‘1’. These two factors were implemented in examples 1 and 5 of the tables. The page authors of these examples followed the example provided in the Guidelines, but chose to substitute the variables provided in the example. There was no indication

\(^{67}\) W3C Validator available at: http://validator.w3.org/ (checked Dec 2000).
in the example that the variables needed to follow a specific syntax. Furthermore, there was no indication that the Headers and ID tags could not be used to associate cells with more than one other cell.

This finding supports the researcher’s assertion in Chapter 3 that page authors need additional support for following examples. Also the examples need to be as transparent as possible so that it is clear whether there are restrictions on the variables that affect the validity of the element.

4.11.12 Discussion of method

4.11.12.1 Difficulties with recruitment

The researcher encountered some difficulties with recruiting visually impaired participants. The original recruitment email message sent to the email lists did not provoke as many responses as had been expected. Other researchers have found that the Internet is an ideal arena for recruiting participants. Smith & Leigh found the Internet “an inexpensive and efficient method to access potentially large numbers of volunteers” (Smith & Leigh, 1997, p497). Schmidt believes “it is possible for researchers to access a large number of individual with an interest in very narrow topic domains” (Schmidt, 1997, p274). The visually impaired people required for this study could be described as having an interest in a narrow topic domain, but it was not possible to recruit large numbers of them.

It is difficult to know why more people did not respond. They may not have the time to spare, they may think ‘someone else will do it’, they may not think the study is going to improve accessibility. Those that have access to the Web from home may not be able to afford the cost of the telephone call for the time required for the evaluation or to download the zipped version. Those that only have access to the Web from work may not be allowed to spend time using the Web for personal purposes, especially during work hours (as reported in a recent radio programme (BBC, 1999)). The original message did not include a deadline, which may have inadvertently encouraged people to put the task off indefinitely. Some of these issues were addressed in a follow-up email message sent to the same email lists.
Two weeks after the original email message was sent a follow-up message was sent to the same lists as before. It was also sent to some Visually Impaired Computer Users Groups (VICUGS) in the US. The message served several purposes:

- Reminded those who had been intending to take part but had not yet done so.
- Set a deadline two weeks from the date of the email message.
- Highlighted the fact that the results of the study would be fed into the ongoing development process of the WAI Web Content Accessibility Guidelines.
- Reminded people that they were not required to complete all the tasks, but just to do as much as they wanted.
- Invited people to let the researcher know if they had been ‘put off’ for any other reason.

The benefits of sending this follow-up message were clear from the response in the following few days. One person wrote to say that they would like to take part but only had access to a friend’s computer. Several people sent answers to the questionnaire on the deadline date and apologised they had left it so late.

Ten of the 41 people who were sent personal invitations replied, saying that they were currently busy but would try to do the series of tasks in the next week or so. Out of these 41, 6 actually did take part.

The problems with recruitment mentioned above may have several implications. If left to do a task ‘when they have time’ many people may not get round to it. This would suggest that it might be preferable for people to attend the laboratory. However, this would exclude people who did not live within a few hours’ travelling distance of the University. Also, if there were so few people who were willing or able to spend a few hours performing the series of tasks in their home or office, there may be even fewer people who would be willing to travel to do the task.

There are other problems associated with having participants attend the laboratory to take part in experiments. As can be seen from the profiles of the people who took part in the test site evaluation, there is a wide range of different browsers, screenreaders, and speech synthesisers in use. Screenreaders are complex
applications that are all quite different from each other. They take many hours of training to become proficient and many users have them set up with several personal preferences. It is therefore important that people use their own screenreader for this type of evaluation. It would be difficult to set up all the different combinations of browsers and screenreaders in the laboratory, especially to customise them for each participant.

One participant contacted the researcher to ask whether he could attend the laboratory in order to take part in the experiment. He said that it would be difficult for him to take part during office hours while in his office. His employer agreed to give him time off on the basis that he would gain experience of different Web elements while taking part. This scenario suggests that if participating in an experiment involved the participants learning about the use of a technology, and if this was seen to be of benefit to their work environment, this might be a good way of recruiting participants. Any training would increase the time needed to complete any task and this would have to be considered during the planning stage of an experiment.

The one participant who did attend the laboratory afterwards informed the researcher that although they had encountered many difficulties with navigating the tables during the task, when they tried them again from their office they had much less difficulty. The researcher thought that this might have been because the correct versions of the screenreader and browser had been installed, but the participant believed it was because they were in a familiar environment and had had some practice while in the laboratory.

4.11.12.2 Limitations of method

With hindsight it is clear that there are several aspects of the method used that could be improved. Listed below are improvements that should be included in any future study.

- Include an example of an imagemap that does not include text links. This will show whether participants can access the alternative text of the imagemap sections.
• Use a real form and ask participants to complete and submit it. A summary of the submitted information could be returned to the participant. This would reveal whether they completed the form as they intended.

• In tables do not include more than one instance of any number. This will reveal where participants get any incorrect answers from.

• Include a table used for layout as well as for data.

• Measure participants' confidence in their answers to any set question.

• Ask participants to rate each example, rather than indicate which they prefer, and ask why each example is preferred.

• Present elements and examples in a different order for each participant. This will avoid a potential 'tail off' in responses.

• Present adapted examples and compare to non-adapted examples.

4.11.12.3 Decline in participation

The different elements were presented in the same order to all participants, although participants were told that they could evaluate them in any order. If the number of participants who evaluated each of the elements is considered in view of the order in which the elements were presented, the numbers decline from 20 to 15. The order in which the elements were presented is shown below, each followed by the number of participants who evaluated them: tables: 20; forms: 21; frames: 18; imagemaps: 19; headers: 18; references: 18; images: 15.

4.11.12.4 Negative choices

In many cases the participants reported that they made negative choices, especially when the choice was between two examples, such as with forms, frames and imagemaps. For example, some participants preferred example 1 of the forms because they found the Tabindex in example 2 was incorrect. The partially sighted participants preferred example 2 of the frames because they did not like the image-links in example 1. Some participants preferred example 1 of the imagemaps because they did not like the list of links in example 2.
With hindsight the researcher believes that this problem could have been avoided if participants had been asked to rate each example, and provide reasons for the rating, rather than to choose which example they found to be the most accessible.

4.12 Conclusions on the accessibility of the examples

In addition to the design of the examples, other factors that affect accessibility are the way in which the information is presented by browsers and screenreaders, and the individual differences between participants, such as their sight status and their experience of using the Web. In some cases these factors had a significant impact on the accessibility of the elements evaluated in this study. Those participants with less than 2 years' experience of using the Web were significantly more likely to find none of the table examples accessible. The use of a GUI browser was significantly related to the preference for tables (example 3) and imagemaps (example 2). Furthermore, the blind and partially sighted participants preferred different examples of the tables, the headers and the frames.

With the relatively small number of participants involved in this study there was a lower expectation of significant relationships. Were the study to be repeated with a larger number of participants more significant relationships could emerge.

The results for the tables suggest that until browsers and screenreaders support the association between cells an interim solution is required. Early versions of the Guidelines recommended the use of paragraph breaks in each cell, but this was removed when Headers and IDs were introduced in HTML Version 4.0. Without an interim solution blind users will continue to find that a table row may be read as a whole line, without any pauses between cells. Also, cells containing text wrapped onto more than one line will not be read as one unit of information. Furthermore, blind users will need to continue to visualise and memorise the layout of the columns and rows. That participants with less experience of using the Web were significantly more likely to find the tables inaccessible suggests that these participants were less experienced in memorising table layouts. It is interesting that the row-based table was perceived to be easier to navigate than the column-based tables, particularly for
those using GUI browsers, even though the number of correct answers to the set question were lower for this layout.

The results for the frames are similar to those of the tables in that a horizontal layout was found to be more accessible than a vertical layout. This is presumably because screenreaders read horizontally. Another similarity is that participants reported having to build up a mental picture of the layout in order to navigate the frames. An interesting issue is that the Noframes versions of the frames were only available to those using special or text-based browsers. Those using GUI browsers did not have access to the Noframes versions. Noframes versions are intended to enable blind users to avoid having to navigate frames, but some blind people were not able to access this. This problem has two potential solutions. The Guidelines could recommend that a link to the Noframes version be provided. Also, the User Agent Guidelines could recommend that GUI browsers provide access to the Noframes version.

The results from the forms support the recognised need for a Tabindex in forms and for the instructions associated with controls to be located on the same line.

The results for the headers highlight the difficulties that blind people experience with visual conventions, such as superscript numbers and condensed email addresses. While these conventions may be useful for visual presentation, they are not useful when presented in synthesised speech. These results also highlight the need for the context of information to be available, for example for a string of numbers to be identified as a telephone number. Furthermore, from these results it is also clear that navigational aids, such as links in a table of contents, are of benefit. Page authors need to consider these three issues. However, the issues are difficult to translate into guidelines, so it is not clear how they should be presented.

The results for the references provide a good example of how there is not necessarily a consensus on some issues, such as the length and content of link text. Some users may prefer longer link text because it is more informative, whereas others may prefer
links to be shorter because they are read more quickly, particularly if the same links are repeated several times.

The results for the images are of concern because the alternative text was not accessible to many participants, who were using different browsers, different screenreaders and had different amounts of experience of the Web. In some cases the D-link was the only source of a description of an image. The D-link approach does not involve a specific accessibility tag, but simply provides a link to a description using standard HTML. Currently the use of the D-link is only recommended for use with ‘important’ images that require detailed descriptions. (Also the Longdesc tag is intended to replace the D-link.) The researcher would not suggest that the Guidelines recommend the use of D-links for all images for two reasons. Firstly, it would be too laborious for blind users to have to use links to get to all descriptions. Secondly, it is believed that page authors may be reluctant to use this approach because it would be time-consuming to implement and to maintain. A preferable solution would be for the User Agent Guidelines to recommend that browsers and screenreaders present alternative text automatically.

The results for the images do not shed much light on the issue of the content of alternative text. A previous study found there was no consensus on the content of alternative text and this is certainly supported by the findings regarding link text from the results on references.

In contrast to the images, the imagemaps were accessible to all the participants, but this was because they were able to access the alternative text links provided for the imagemaps. With the text links the important issues were whether the links on one line were read as separate units (by GUI screenreaders), and whether the links in the list were preceded with the phrase ‘list item’ (by special browsers). These are related to the issues of table rows being read as one line, and the word ‘reference’ being repeated. The Guidelines recommend that links are separated by non-link characters and these results confirm that this is necessary. As with the headers, page authors need to be aware of the way that assistive technology presents information in order to understand the purpose of this recommendation. In addition, the User Agent
Guidelines need to emphasise the need for users to be able to easily customise their browsers and screenreaders so that they can choose how information such as list items are presented.

The results for all the elements indicate that the requirements of partially sighted users are different to those of blind users. Partially sighted users who have the text displayed in their browsers enlarged can see less of a page, or element, at one time, and therefore need to scroll to see the whole page or element. This clearly is a slower way of working and more effort is required to build up an overview of the page. Therefore design that reduces scrolling, such as narrower tables, should be recommended in the Guidelines. It is also important that the ability to scroll is provided, for example in frames, so that the user can access the entire content of a frame.

For partially sighted users the clarity of the visual appearance of a page is important, perhaps even more so than for users without a visually impairment. For example, for partially sighted users the use of borders helps to distinguish between frames. Also these users need high contrast between the foreground and background, for example in images with embedded text. Finally, for partially sighted users the font or style of text is important, for example, italicised text is more difficult to read.

Some requirements of partially sighted users may conflict with common approaches to the presentation of text, or with perceived good practice for HTML. For example, partially sighted users may prefer text to be left justified, whereas centre justified text is often used for aesthetic purposes. Also, partially sighted users may find that large headings require them to scroll too much, but the HTML specification recommends that the largest heading (H1) is used for the most important heading on a page, which is usually the first heading\(^6\). The Guidelines include recommendations for high contrast between fore- and backgrounds, but the researcher believes that more information should be provided regarding the way in which partially sighted users interact with Web pages, in particular how they have to scroll pages.

\(^6\) For details on the use of H1, see http://www.w3.org/TR/1998/REC-html40-19980424/struct/global.html#h-7.5.5 (in the HTML 4.0 Specification) (checked Nov 2000).
Three important issues emerge from the results of this study. Firstly, the alternatives provided for blind people, such as Noframes and alternative text, did not necessarily improve accessibility for some participants who were using GUI browsers. An explanation for this might be that the Guidelines are based on an outdated assumption. The first accessibility guidelines such as those from Vanderheiden, Chisholm, Ewers, & Dunphy (1997) and Wesley, Reeves, & Evenepoel (1997) were developed four years ago. At that time most blind people were using, or were thought to be using, text-based browsers. These browsers automatically displayed the alternatives because they did not support frames or images. In the last few years, the accessibility of GUI browsers has improved and, if the participants in this study form a representative sample, it seems the majority of blind people are using GUI browsers. These browsers display images and frames and therefore do not, or cannot, present the alternatives. A premise of the Guidelines is that blind people sometimes require alternatives, but this needs to be updated, because blind people now need access to the alternatives.

The second issue is that in some cases blind users require a way of accessing the visual layout of information, for example in tables and frames. Currently blind users have to read the information repeatedly in order to build up an idea of the layout. This is time-consuming, and the mental image is not always accurate. While solutions for these problems are available, they are not sufficient. For tables it is possible to create associations between cells, but this approach is not supported by any browser, and there is not guarantee that it will be supported in the future. Also there is no guarantee that this approach will be usable for blind people. For frames it is possible to provide Noframes versions, but GUI browsers, the most commonly used browsers, do not support this.

The final issue relates to the design of textual information and how this affects the way in which the information is presented in synthesised speech. It is clear that the context of specific units of information is important. Examples of situations in which context is important are links and form controls. If a user uses the tab key to get to a link, they will only hear the text of the link and not the surrounding sentence.
Similarly, when a user jumps between the controls of a form, they will not hear the whole line on which the control is located, unless they actively read it. If the instruction for the control is not on the same line it may be difficult to find it and relate it to the correct control. Examples from this study of contexts for textual information that were found to be useful involve the association of a word with specific numbers. The word ‘telephone’ immediately preceding a string of numbers and the word ‘reference’ accompanying a single number embedded in text clearly indicate the purpose of the numbers. Although this raises the related issue of whether users wish the word ‘reference’ to be repeated twenty times in a paper with this number of references.

4.12.1 Definition of ‘accessibility’

The term ‘accessibility’ needs to be clearly defined. It could refer to the ability to read and identify all the available information. An example of this type of accessibility is provided by the participants’ experiences of the images. Some participants were able to identify that images were present and to read the alternative text. The images were accessible to these participants. Other participants were not able to read the alternative text and so the images were not accessible to them.

However, the above definition does not take into account the need to be able to read information in an efficient and useful way. For example, the information in the tables was available to most of the participants in that they were able to read the text of the headers and the numbers in the cells. But being able to read the information is not sufficient, particularly in the case of tables. The reports from some of the participants state that they identified the column headers by reading right across the first few lines and using the pauses or spaces to identify the relationship between the different pieces of text. Once this relationship was recognised it could be used to identify the headers. The headers of each column had to be memorised in order to know which column header related to the data in each cell.

Although many of the participants were able to give the correct answers to the question, this is not an efficient way of working and the ability to read information cannot constitute ‘accessibility’. In order for information to be accessible users need
to be able to identify relationships between different aspects of the text, and to be able
do this efficiently, as well as being able to read it. In the design of (mainstream)
interfaces the issue of being able to work effectively and efficiently is part of the
usability of the interface.

It is clear from the Bobby reports on the examples that this checker is using different
definitions of accessibility to those used by the participants.

4.12.1.1 Usability

Some of the issues raised by the participants can be considered to be specifically
usability issues rather than accessibility issues. The inclusion of the word ‘telephone’
in example 4 of the headers increased the usability of the example, rather than the
accessibility. The issue of the phrase ‘list item’ being announced (or not) in the
examples of the imagemaps increased or decreased the usability of lists for different
participants. The inclusion (or not) of the word ‘reference’ in the links in the
examples of references also increased or decreased the usability of the references for
different participants. These issues suggest that these users may require more control
over what is read, and the way it is read. For example, screenreaders could allow
users to choose whether or not “list item” was announced for each item in a list. It is
beyond the scope of this thesis to investigate how this type of option should be
facilitated.

4.12.2 Alternative approach to improving accessibility

The issues raised in Section 4.12 above are clearly not sufficiently addressed by the
use of accessibility guidelines. This is partly because page authors have difficulty
using guidelines (as found in the study presented in Chapter 3), and partly because the
techniques suggested by the guidelines cannot adequately improve accessibility. This
in turn is because current browsers and screenreaders do not support all the
techniques, and the techniques do not cover all the aspects that visually impaired
people have difficulty with (such as preceding a string of numbers with the word
‘telephone’). Since the use of accessibility guidelines cannot ensure full accessibility
of the Web alone, alternative approaches are required to complement this approach.
For two of the three issues raised above (access to alternative content and access to
layout) an alternative approach is to use novel interfaces, such as those available via haptic devices, which can provide alternative ways of presenting content. Such devices could provide access to the original content, thus reducing the need for alternatives such as Noframes versions. Such devices are also particularly appropriate for presenting the layout of elements such as frames and tables. In order to improve accessibility through the use of haptic devices it is essential that the way in which artefacts are perceived via such devices is investigated. This is the aim of the studies presented in Chapter 6, which follows the review of the literature on touch perception of textures and objects in the next Chapter.
4.12.3 Summary of conclusions

The conclusions of this study for blind and partially sighted people are summarised in Table 4.19. Specific problems and potential solutions are identified, and suggestions for changes to the WAI Content and User Agent Guidelines are presented.

Table 4.19: Summary of conclusions from study

<table>
<thead>
<tr>
<th>Element</th>
<th>Type of problem identified (by participants or the researcher)</th>
<th>Type of user / browser / screenreader affected</th>
<th>Potential solutions</th>
<th>Changes suggested for Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tables</td>
<td>Difficult to navigate</td>
<td>Blind users (all browsers and screenreaders)</td>
<td>Support for Headers and ID tags by future browsers / screenreaders</td>
<td>WCAG: Inform authors that tags not currently supported, and provide another solution</td>
</tr>
<tr>
<td></td>
<td>Columns more difficult to read than rows</td>
<td>Blind users of GUI browsers and braille displays</td>
<td>Provide information in rows?</td>
<td>WCAG: Suggest use of rows rather than columns</td>
</tr>
<tr>
<td>Forms</td>
<td>Difficult to navigate if Tabindex incorrect</td>
<td>Blind users of GUI browsers</td>
<td>Tabindex needs to be correct</td>
<td>WCAG: emphasise importance</td>
</tr>
<tr>
<td></td>
<td>Difficult to make association between Control and Instruction</td>
<td>Blind users of GUI browsers, text-based browsers, braille displays</td>
<td>Control and Instruction need to be on same line</td>
<td>WCAG: emphasise importance</td>
</tr>
<tr>
<td></td>
<td>Default text in text-entry fields needs to be enlarged within browser settings</td>
<td>Partially sighted users using large font settings in browser</td>
<td>Support by future browsers</td>
<td>Inclusion in UA Guidelines</td>
</tr>
<tr>
<td>Frames</td>
<td>No indication of number of frames present</td>
<td>Blind users, particularly WS users</td>
<td>Support by future browsers</td>
<td>WCAG: Additional checkpoint required</td>
</tr>
</tbody>
</table>
| WCAG: Additional | Purpose of D-Link | Appearance of text, numbers, or symbols. | Users of GUI browsers, partially shielded users, and text that is difficult to read or understand.

**Inclusion in U/A Guidelines**

**Checkpoint required**

**Support by U/A ( guardian)**

**Implementation of**

**Explanation of**

**Major Point**

| WCAG: Additional | Purpose of D-Link | Appearance of text, numbers, or symbols. | Users of GUI browsers, partially shielded users, and text that is difficult to read or understand.

**Inclusion in U/A Guidelines**

**Checkpoint required**

**Support by U/A ( guardian)**

**Implementation of**

**Explanation of**

**Major Point**
<table>
<thead>
<tr>
<th>References</th>
<th>Other Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>No problems can be evidenced</td>
<td>Some information not required to be replaced</td>
</tr>
<tr>
<td>Numbers not self-evident</td>
<td>&quot;Hi, I'm...&quot; announcement or e-mail announcement of particular GUI users required</td>
</tr>
<tr>
<td>Purpose of session or activity described</td>
<td>All blind users support by June</td>
</tr>
<tr>
<td>Most blind users (not necessarily all)</td>
<td>Particularly GUI users</td>
</tr>
<tr>
<td>WCAG: additional information or reference of internal reference, if necessary</td>
<td>Number of required users</td>
</tr>
</tbody>
</table>
CHAPTER 5 – REVIEW OF THE LITERATURE ON TOUCH PERCEPTION AND THE USE OF TOUCH IN COMPUTER INTERFACES

5.1 Introduction ................................................................. 215
5.2 Touch perception ............................................................ 216
  5.2.1 Definitions ........................................................... 216
  5.2.1.1 Tactual perception ............................................. 216
  5.2.1.2 Tactile perception ............................................. 217
  5.2.1.3 Kinaesthetic perception ..................................... 217
  5.2.1.4 Haptic perception ............................................. 217
  5.2.2 Physiology ......................................................... 218
    5.2.2.1 Adaptation .................................................. 219
  5.2.3 Experience of touch ............................................. 219
  5.2.4 Active and passive touch ..................................... 220
  5.2.5 Synthetic and analytic functions ............................ 220
  5.2.6 Functional touching ............................................ 221
5.3 Perception of texture .................................................. 221
  5.3.1 Vibration ............................................................ 222
  5.3.2 Pressure ............................................................. 222
  5.3.3 Hand force and speed .......................................... 223
  5.3.4 Perception of roughness .................................... 223
    5.3.4.1 The pioneering work of Stevens & Harris ............ 223
    5.3.4.2 Investigations of roughness by Lederman .......... 225
    5.3.4.3 Selective adaptation .................................... 227
    5.3.4.4 Discrimination ........................................... 227
  5.3.5 Role of visual imagery in perception of texture .......... 227
5.4 Perception of Form .................................................... 228
  5.4.1 Perception of length / extent ................................ 229
  5.4.2 Perception of proportion ..................................... 230
  5.4.3 Perception of orientation .................................... 232
  5.4.4 Perception of curvature ...................................... 233
  5.4.5 Identification of objects by touch ......................... 233
  5.4.6 Recognition of objects ....................................... 233
  5.4.7 Illusions and after-effects in perception ................ 234
  5.4.8 Observations of participants’ behaviours ................ 234
  5.4.9 Perception of form by blind people ....................... 235
  5.4.10 Factors that appear to have impact on perception of form 235
5.5 Use of touch in computer interfaces ................................ 237
  5.5.1 Brief history of the use of touch in interfaces .......... 238
  5.5.2 Brief survey of haptic devices: their characteristics and potential applications 238
5.5.2.1 Haptic devices .........................................................239
5.5.2.2 Haptic devices with visual display .................................239
5.5.3.1 Haptic interfaces for use by visually impaired people .........240
5.5.3 Perception via haptic interfaces .......................................241
5.5.3.2 Implications of perception for interfaces .........................241
5.5.4 Modelling ..............................................................241
5.5.5 Quality of feedback .....................................................241
5.5.6 Human factors in haptic feedback devices .........................242
5.5.7 Applications of haptic feedback devices ............................243
5.5.7.1 Virtual Environments on the World Wide Web .................244

5.6 Perception of Virtual Textures ........................................244
5.6.1 Measuring confidence and roughness ...............................244
5.6.2 Comparison of real and virtual textures .........................245
5.6.3 Comparison of finger and probe exploration of textures .......246

5.7 Perception of Virtual Objects .........................................247

5.8 The implications of previous haptic research on the current studies 248

5.9 Conclusions on touch perception .....................................249

5.1 Introduction
This chapter reviews the literature on touch perception, in particular the perception of texture and form (including 3-dimensional objects). This includes both real and virtual artefacts. There is a substantial body of research on the perception of real textures and a small but growing number of researchers examining the perception of virtual textures. While there has been much research on various aspects of the perception of form, there is relatively little research that is directly comparable with the study presented in this thesis involving virtual objects. Even the literature on cross-modal matching (the nature of the task performed by the sighted participants with virtual objects in the study reported in Chapter 6) such as the work by Seizova-Cajic (1998), is too different for useful comparison.

The chapter begins with the perception of 'real' artefacts (as opposed to virtual) and then goes on to examine the use of touch in computer interfaces. Literature regarding haptic interfaces is reviewed, including a historical overview and a survey of existing devices. The literature on perception of artefacts via haptic interfaces is reviewed, including the perception of virtual objects and textures.
It will be seen that our knowledge of touch is limited although there is research on many different aspects of touch involving various techniques. The research on the perception of texture has investigated several factors that may affect perception, including vibration and pressure, as well as physical properties.

5.2 Touch perception
Our knowledge of the sense of touch is limited despite a substantial body of research, conducted over 30-40 years.

"Our knowledge of touch consists of only fragmentary concepts and findings, some dealing with basic functional properties (e.g., cutaneous sensitivity, limits of kinaesthetic space perception) and others with capabilities of the systems as a whole (e.g., the identification of three-dimensional objects)" (Loomis & Lederman, 1986, p31-36).

5.2.1 Definitions
There is an abundance of terms associated with touch, such as ‘tactile’, ‘tactual’, ‘haptic’, ‘kinaesthetic’ etc. These categories of touch have been further divided into modes of touch, such as ‘active’ and ‘passive’ (see below for further explanation). The terms have all been used in different ways in reports of different investigations.

5.2.1.1 Tactual perception
We are often aware of our sense of touch, but for many researchers this sense actually comprises two distinct senses: the cutaneous and kinaesthetic senses. The cutaneous sense "provides awareness of stimulation of the outer surface of the body by means of receptors within the skin and the associated nervous systems" (Loomis & Lederman, 1986, p31-2). The kinaesthetic sense "provides awareness of static and dynamic body posture" (op. cit, p31-2). Loomis & Lederman (1986) use the term ‘tactual perception’ to refer to all perception mediated by both the cutaneous and kinaesthetic senses.

69 The term ‘kinaesthetic’ is spelt thus in the UK, but spelt ‘kinesthetic’ in the US. The former spelling is used in this thesis, except in quotations that use the latter spelling.
5.2.1.2 Tactile perception
Tactile perception is defined by Loomis & Lederman as perception mediated solely by variations in cutaneous stimulation, that is, mediated through the skin (Loomis & Lederman, 1986). The perception of patterns drawn on the back, and the Tadoma method of speech reception (in which the hearing-impaired ‘listener’ places their hand on the jaw and lips of the speaker) are both examples of tactile perception.

5.2.1.3 Kinaesthetic perception
According to Loomis & Lederman, kinaesthetic perception “provides information about the relative positions and movements of the parts of the body (static and dynamic posture) as well as about muscular effort” (1986, p31-24). This information about static and dynamic body posture comes from muscle and skin receptors.

Kinaesthetic perception without cutaneous stimulation is rare outside of the laboratory: mostly it is contrived in experiments, for example with the use of anaesthesia or circulatory occlusion. In some experiments cutaneous stimulation only indicates contact with the stimuli whereas the relevant spatial information is conveyed by the kinaesthetic stimulation. An example is the judgement of the length of rods held between the forefinger and thumb where cutaneous stimulation indicates contact with the rod but it is kinaesthetic stimulation that is involved in the perception of the length of the rod.

5.2.1.4 Haptic perception
Schiff (1980) believes that cutaneous and kinaesthetic sensitivities should be considered together as the ‘haptic-tactual’ system because they are so closely related. To illustrate the relationship Schiff gives the example of trying to drive a car with bare feet. The difficulties experienced in this situation demonstrate that cutaneous sensitivity is involved in driving a car, as well as the joints of the toes, ankles and knees.

The term ‘haptic’ comes from Greek, meaning ‘able to lay hold of’ and was first used by Revesz (1950) (cited in Appelle, 1991). Gibson (1933) (cited in Appelle, 1991) used the term ‘haptics’ to refer to system involved in the seeking and pickup of information by hand.
For Loomis & Lederman (1986), haptic perception refers to perception in which both cutaneous and kinaesthetic senses convey information about objects and events. They do not restrict this definition to perception via the hand and believe that most of our everyday tactual perception involves haptic perception. This includes the sensing of food texture by the mouth.

Gibson (1966) (cited in Appelle, 1991) believed that the manipulation of objects involves pressure, force, stimulation of the skin, activity of muscles, joints and tendons. For Gibson the haptic perceptual system includes various subsystems:

- Cutaneous touch (stimulation of skin without movement of muscles or joints).
- Haptic touch (stimulation of skin plus movement of joints).
- Dynamic touch (stimulation of skin plus movement of the joints and muscles).
- Oriented touch (skin stimulation plus vestibular\textsuperscript{70} stimulation).

Appelle (1991) believes that the proliferation of terms relating to touch that have been presented in this section indicates that we currently have relatively little understanding of the perception of objects. Appelle suggests that this understanding can only be improved if we examine how the many components of touch work together as a system, as well as examining how each work independently.

In this and later chapters the term ‘haptic’ is used to refer to the combination of the cutaneous and kinaesthetic senses used.

5.2.2 Physiology

The hairless skin of the palms and fingers plays the most active role in tactile exploration. This skin has five major types of receptor, each of which responds to different types of stimulation. Nerve endings respond to pain. Meissner corpuscles respond to the movement of stimuli across the skin and detect velocity. Merkel’s disks respond to pressure and vibration. Pacinian corpuscles respond to light touch, vibration and detect acceleration. Ruffini corpuscles respond to shearing of the skin.

\textsuperscript{70}Vestibular stimulation is the stimulation of the channel that communicates with other channels, i.e. the nervous system.
pressure and thermal changes (Burdea, 1996). There is debate about whether these receptors work in isolation or whether, as Schiff (1980) believes, they have overlapping receptive fields.

5.2.2.1 Adaptation

When receptors respond to stimuli they produce discharges which send information to the brain. “With repeated stimulation of a receptor or set of receptors, systematic declines in sensitivities occur” (Schiff, 1980, p116). When the stimulation is continuous the discharges from the receptors drop-off, and the discharges from different receptors drop-off at different rates. The variation in the number of discharges over time is a measure of ‘sensorial adaptation’. Schiff finds that “adaptation to pressure occurs at a faster rate with more intense stimuli...we adapt quickly to large and light-weight stimuli” (Schiff, 1980, p118). An example provided by Schiff of rapid adaptation is the way in which we only become aware of our clothing when we move: the movement produces pressure changes against the skin (Schiff, 1980). This is because the receptors have ‘adapted’ to the presence of the stimuli (the clothes). These receptors are known as ‘rapidly adapting’ (RA) receptors as they have a fast drop-off rate. (‘Slowly adapting’ (SA) receptors have a slow rate of drop-off.) Adaptation occurs in cutaneous receptors and larger areas of the skin to pressure, temperature, vibration and pain (Schiff, 1980).

The spatial resolution of a receptor is defined by its receptive field, the area in which it responds to stimuli. Pacinian and Ruffini corpuscles have large receptive fields and therefore low spatial resolution. Meissner corpuscles and Merkel’s disks have smaller fields and high resolution. Different parts of the skin have different spatial resolution. The spatial resolution of different parts of the skin can be tested using the ‘two point limen’ technique. In this technique participants are touched on a part of the body with two pointers. They are then asked whether they felt one or touch points. The threshold for the perception of two points has been found to be 2.5mm on the fingertips, 11mm on the palm, and 67mm on the thigh (Burdea, 1996).

5.2.3 Experience of touch

Katz was one of the first researchers to investigate the sense of touch. Katz’ work, *The world of touch*, (1925) was translated by Krueger (1989), and commentaries
were written by Zigler (1926), Krueger (1970) (both cited in Lederman, 1982), and Krueger (1982). Katz emphasised that when we touch an object, the perceptual experience is of the object, rather than the skin being touched. For example, if a surface is explored using a stylus that is held in the hand, one is aware of the surface not of the vibrations felt by the hand. Katz developed a model in which, “The senses vary in the degree to which the resulting percepts are experienced as part of the self (are “phenomenally subjective”) or as external to the self (are “phenomenally objective”)” (Loomis & Lederman, 1986, p31-4).

According Katz’ model, the visual sense is at the phenomenally objective extreme because visual experience refers to perceptual space beyond the self. At the phenomenally subjective extreme are sensations such as hunger and pain because these sensations are experienced within the body (interoceptive senses). Loomis & Lederman find this distinction descriptive, rather than theoretical, but acknowledge its importance in the description of the experience of touch. These authors believe that touch is located between the subjective and objective extremes, and that the perceptual experience can refer to either the subjective or objective poles. If an observer explores an object actively and uses the part of the skin that is normally used for tactual exploration, then the objective pole is favoured. On the other hand, if skin that is not normally used for tactual exploration is touched passively (such as the inside of the nose or ear) the subjective pole is favoured.

5.2.4 Active and passive touch
The distinction between active and passive touch mentioned above was investigated by Gibson (1962; 1966) (both cited in Loomis & Lederman, 1986). Gibson found that when observers were exposed to passive tactile stimulation they tended to describe their experience in terms of tactile sensations. In contrast, when they engaged in active tactual exploration observers tended to describe their experience in terms of objects. Gibson believed that this difference was to do with whether or not the observer had control over the ‘pickup’ of information.

5.2.5 Synthetic and analytic functions
Heller (1983) made the distinction between synthetic functions of passive touch and analytic functions of active touch respectively. The intention of synthetic touch is to obtain an overall impression of an object. This is carried out by the ‘resting’ hand,
i.e. using passive touch. The moving hand uses analytic touch to gain an impression of the features of the object, i.e. using active touch.

It seems slightly contradictory to say that a ‘resting’ hand can obtain an ‘overall’ impression of an object. However, it is assumed that Heller means that when we hold an object in our (resting) hand we can gain information about the object, such as its weight, and hardness / softness. It is only when we actively explore the object that we can gain information about, for example, its size and the roughness / smoothness of its surface.

5.2.6 Functional touching
Schiff (1980) makes a distinction between the study of cutaneous and kinaesthetic sensitivities in classical psychophysics and the more ‘functional analyses’ of touching behaviours and the resulting percepts. Such functional approaches investigate how we gather information about the world using information-gathering systems, such as the hands. Katz (1925) and Gibson (1966) (both cited in Schiff, 1980) identified ‘sets’ of related systems, such as the shoulder, arm, wrist and hand, which all consist of skin, muscles and joints. The functional approach considers how we use these systems to extract the information we need.

5.3 Perception of texture
There have been many studies of the perception of texture to investigate the effects of factors such as vibration and pressure. Many factors have been varied such as surface structure, coatings on participants’ fingers, and the speed and force of hand movements.

According to Loomis & Lederman, the concept of texture refers to “the physical properties of objects, fluids and substances” (1986, p31-26). This includes attributes such as roughness, hardness, elasticity, and viscosity.

Katz (1925) (cited in Lederman, 1982) was one of the first to investigate the perception of texture. Katz conducted many experiments that investigated the roles

---

71 A percept is “a mental concept resulting from perceiving” (Allen, 1990, p883).
of vibration, pressure, and the force and speed of the hand in the perception of texture. These factors are discussed in turn in the following sections.

5.3.1 Vibration
Katz (1925) (cited in Lederman, 1982) believed that vibration was necessary for the perception of texture. Katz argued that texture cannot be perceived if a finger is held stationary on a surface because of the absence of vibration. Vibrations in the skin are created by relative movement between the skin and a surface.

In demonstrating the necessity of vibration Katz (1925) (cited in Schiff, 1980) applied glue to the fingers of subjects and then had them feel paper surfaces. The surfaces were reported to feel very smooth. Katz argued that they felt smooth because the glue eliminated any vibrations. Lederman (1982) presents an alternative interpretation of Katz’ results. Lederman suggests that ‘lateral shearing forces’ on the fingertips may mask the stimulation. The glue may have increased such lateral forces and therefore masked the sensation of roughness.

Schiff (1980) places a different emphasis on this study conducted by Katz. Schiff suggests that the reason that perceived roughness increased while the surfaces were reported to feel smooth is that “haptic-tactual perception of surface structures does not depend solely on coefficient or friction variables, nor on any sort of kinesthetic information” (1980, p159). Schiff supports this view with the example that relative roughness can be judged by just rolling the finger on a surface, without producing any friction. However, Schiff does not offer any empirical evidence for this assertion.

5.3.2 Pressure
In a further study Katz (1925) (cited in Lederman, 1982) investigated the roles of vibration and pressure of the perception of texture. Participants were asked to feel two different surfaces simultaneously, one with each hand. In different conditions the participants’ fingertips were coated in cellulose on one, both, or neither hand. Katz found that discrimination between the surfaces improved when either both or neither hands were coated and discrimination decreased when just one hand was coated. Katz concluded that the cellulose filled the irregularities on the surface of the
skin, which only slightly dampened the vibrations but excluded the sense of pressure. Therefore if pressure played a major role in perception of texture then discrimination should have been worse when the skin was covered. Lederman suggests, however, that the coating may have reduced the lateral force acting on the skin (Lederman, 1982).

5.3.3 Hand force and speed
Katz (1925) (cited in Lederman, 1982) also investigated the effect of hand force and speed on the perception of texture. Katz found that when surfaces were moved across stationary fingers, perceived roughness increased with increasing force. Katz also found that when surfaces were moved at 15 cms per second across stationary fingers, discrimination of roughness decreased.

In another study Katz found that judgements of roughness were the same when subjects moved their hands over ridges at speeds varying from 1 - 10 cms per second. Lederman suggests that this latter result may indicate “a form of roughness constancy in which subjects compensate for the alterations in vibratory frequency resulting from variation in hand speed” (Lederman, 1982, p135).

5.3.4 Perception of roughness
Most of the work on texture perception since the 1960s has investigated the perception of roughness and smoothness. Loomis & Lederman define it as “undulations or protrusions of a surface that are of a much smaller scale than the fingertip but large enough to permit tactual discrimination between the surface in question and one that is smooth” (Loomis & Lederman, 1986, p31-27).

5.3.4.1 The pioneering work of Stevens & Harris
The study of the psychophysics of texture began with the work of Stevens & Harris (1962) who were the first to use the magnitude estimation technique to measure the perception of roughness. Stevens used this technique to investigate the brightness of light, the magnitude of electric shocks, the length of lines and many other physical dimensions of the sensory world (see Figure 5.1 below).
In a magnitude estimation study, participants are asked to feel a standard stimulus and give it a number that represents the intensity of the stimulus for them, for example 10. This number is known as the modulus. Participants are then asked to feel a range of other stimuli and give a number that reflects the relationship of the current stimulus to the standard. This is the ‘magnitude estimation’ (ME). If the stimulus feels twice as intense as the standard they give it an ME of twice the modulus, for example 20. If it feels half as intense as the standard they give it an ME of half the modulus, for example 5. This sequence is repeated several times to build up a picture of the perception of the stimuli.

**Figure 5.1: Psychophysical functions for electric shock, apparent length, and brightness.**

(Graph from Snodgrass, Levy-Berger, & Haydon, 1985, p79.)

Stevens found that the perceived sensation of a physical characteristic is related to the actual magnitude of the physical characteristic, via a power function, which is presented thus:

\[ S = a I^b \] (Snodgrass, 1985)

Where I is the magnitude of the stimulus, S is the magnitude of the sensation, and ‘b’ is the value of the exponent. The value of 'a' reflects the particular number which the participant used for the standard stimulus, the modulus. The magnitude of the exponent (b) is important in this equation. When b is greater than one (b > 1) it means that the intensity of the sensation grows more rapidly than the intensity of the
physical stimuli (see the curve for electric shock in Figure 5.1 above). When the exponent is less than one \((b < 1)\) the reverse is true (see the curve for brightness in Figure 5.1 above) (Snodgrass et al., 1985).

Stevens & Harris (1962) investigated the perceived roughness of emery cloths that varied in grit value (particle size). Subjects made magnitude estimates of both roughness and smoothness and found that these dimensions of perception were power functions of grit value (with exponents of 1.5 and -1.5 respectively). It was concluded that this indicates that smoothness and roughness are reciprocally related.

5.3.4.2 Investigations of roughness by Lederman

Lederman and colleagues (Lederman, 1974; 1981; Lederman & Taylor, 1972) have performed a series of studies to investigate tactual perception of roughness. Aluminium plates were used as the stimuli, which had linear gratings with a rectangular cross section cut into the surface (see Figures 5.2 and 5.3 below). The depth of the grooves was sufficient to avoid the finger touching the bottom of the groove. A range of parameters was varied, including groove width, land width, contact force, and scanning velocity. An overview of the results is presented below. In all trials subjects used their fingertip(s) to scan over the surface and the magnitude estimation technique was used.

5.3.4.2.1 Surface structure

In the first experiments Lederman and colleagues varied the surface structure of the textures. The results showed a strong relationship between increased groove width and increased perceived roughness (Lederman 1974; 1981; Lederman & Taylor 1972). The research also found that perceived roughness decreased with increased land width and that the ratio of groove width to land width had a negligible effect (Lederman & Taylor, 1972).
5.3.4.2.2 Active and passive touch
Lederman (1981) also compared active and passive modes of touch. No difference in the judgements of roughness was found between the two modes. It was concluded that estimates of roughness do not require kinaesthetic perception but depend solely on the tactile sense.

5.3.4.2.3 Hand speed and force
As well as varying surface structure and the mode of touch, Lederman also investigated the role of hand speed and hand force in the perception of roughness. In one study it was found that perceived roughness decreased slightly with increasing hand speed (Lederman, 1974). In several studies perceived roughness was found to increase with finger force, although not as much as with increased groove width.
In one study the effect of finger force was found to increase with groove width (Lederman, 1981).

### 5.3.4.3 Selective adaptation

Lederman, Loomis & Williams (1982) (cited in Loomis & Lederman, 1986) investigated the adaptation effects on the perception of vibration and roughness. Subjects rested their fingertip on the contactor of a vibrator. The finger was subjected to 20Hz and 250Hz in the experimental conditions and to 0Hz in the control condition. Subjects made estimates of the magnitude of the vibration.

Large selective adaptation effects were found: the adaptation 20Hz vibration selectively attenuated the perceived magnitude of subsequent 20Hz vibration relative to that of subsequent 250Hz vibration. The converse was true for adaptation to 250Hz. The same conditions were used to study the perceived roughness of grooved plates, but there were no discernible effects of adaptation.

### 5.3.4.4 Discrimination

Lamb (1983) (cited in Loomis & Lederman, 1986) investigated roughness discrimination. Subjects were presented with rectangular arrays of raised dots, which varied in terms of the spatial period in one direction. For each trial two textured stimuli were presented, one being a standard of fixed period, the other being a comparison which was either the same as the standard or different. Subjects reported whether the comparison was the same as or different from the standard. The discrimination index increased linearly with the percentage of change in spatial period. The mode in which the surface was touched, either 'active' or 'passive' did not affect the subjects' ability to discriminate, neither did the velocity of scanning motions, or the manipulation of contact force.

### 5.3.5 Role of visual imagery in perception of texture

Heller (1989) compared the perception of abrasive papers by sighted, congenitally blind and late-blind people, using active and passive touch, in order to investigate the role of visual imagery. It was found that the accuracy of the three groups was

---

72 Adaptation is defined in Section 5.2.2.1 below.
similar, for both active and passive touch. Heller therefore concludes that visual imagery is not necessary in the perception of texture although this is in contrast to previous studies. Heller (1989) reviews several studies in which the performance of congenitally blind, late blind and sighted people are compared in a variety of different tasks. Heller found that the results were conflicting. For example, it was found that visual imagery provided benefits in a memory task for sighted people (Christian, Bickly, Tarka & Clayton, 1978) and late blind people (Warren, 1984). Late blind people have been found to perform better in the perception of form (Worchel, 1951) and other spatial tasks (Millar, 1975; 1976). In contrast, Bailes & Lambert (1986) found no difference in the performance of congenitally blind and late blind people in a form recognition and memory task. Heller (1989) also reviews studies that have compared touch and vision in the perception of texture to investigate the role of visual imagery. However, these studies involved blindfolded sighted people, rather than blind people. Similar performance was reported for the two modalities (Heller, 1982; Jones & O’Neal, 1985; Lederman & Abbott, 1981).

Heller compared the perception of abrasive papers by sighted, congenitally blind and late-blind people, using active and passive touch. It was found that the accuracy of the three groups was similar, for both active and passive touch. Heller therefore concludes that visual imagery is not necessary in the perception of texture. This is in contrast with previous studies on the perception of form.

5.4 Perception of Form
The perception of different attributes of objects has been investigated in previous studies, such as the length of objects, the proportion of one object to another, ratio of one length to another, curvature of edges, and the orientation of objects. A range of methods have been used in these investigations, such as cross-modal matching (recognition or matching of objects with one mode after perceiving them with another mode), discriminating between objects (judging whether pairs are the same or different), and matching shapes. Various apparatus and novel procedures have also been used, particularly in the perception of length, including rods with rubber stoppers for measuring length, giving length stimuli verbally, and the experimenter moving the participants’ arms.
Some researchers have compared the performance of blind and sighted people (e.g. Davidson (1972) (cited in Appelle, 1991)). Other researchers have observed the behaviours of participants, for example to identify the movements of hands (e.g. Davidson (1972) and Zinchenko & Lomov (1960)), or the use of measuring strategies (e.g Appelle, Gravetter, & Davidson (1980)).

The studies discussed here have not investigated the same aspects of the perception of objects as the study described in Chapter 6, but this section gives an overview of the types of investigations carried out, and the limits to our knowledge about touch perception of objects.

5.4.1 Perception of length / extent

In an early study of the perception of length, Jastrow (1886) (cited in Appelle, 1991) investigated judgement of distance using several methods of touch (blocks held between forefinger and thumb, moving a pencil, or participants’ arms moved by the experimenter). It was found that the touching method influenced the perceived distance. For example, a distance between a forefinger and thumb was perceived to be equal to a distance covered by a moving arm that was actually 68% shorter.

Similar results were achieved by Stanley (1966) (cited in Appelle, 1991). Stanley investigated the perception of length with different touch methods. In general, rods that were held between the forefingers (the haptic condition) were perceived to be longer than distances that were judged using outstretched arms (kinaesthetic condition).

Brown, Knauft & Rosenbaum (1948) (cited in Appelle, 1991) asked participants to reproduce distances that were presented by a visual marker by moving a pointer, in the dark. It was found that movements away from the body were more accurate than movements towards the body.

Several researchers have investigated participants’ ability to judge distances in kinaesthetic space by moving a finger or hand between the endpoints that define the distance. Loomis & Lederman (1986) cite several studies of this type in which the
‘radial-tangential effect’ has been observed (for example Davidon & Cheng (1964)). Participants are asked to judge the lengths of segments forming an L-shaped figure lying on a table. The segment lying in the median plane, or the radial length (parallel to the front of the subject’s body) is perceived to be 10% longer than the segment lying in the fronto-parallel plane, or the tangential length (pointing away from the subject’s body) (Loomis & Lederman, 1986).

Seizova-Cajic (1998) measured the ability of participants to reproduce the length of a stimulus given verbally, by either drawing a line of that length (kinaesthetic condition), or making a mark on an existing line (visual condition). In the kinaesthetic condition the reproduced length was underestimated and the underestimation increased with the actual length of the stimulus. Reproduced length in the visual condition was very close to the actual length.

Seizova-Cajic also investigated the effect of the direction of movement in kinaesthetic perception. Lengths were perceived to be longer in the medial plane (parallel to the front of the subject’s body) than in the fronto plane (directly in front of (going away from) the subject). When the standards were observed by moving the finger in the ‘away’ direction, they were perceived to be longer than the ‘toward’ direction, in both fronto and medial planes. Seizova-Cajic concluded that stimuli were perceived to be shorter via kinaesthetic perception than via visual perception and the difference between the two increased with the size of the stimuli.

Appelle, Gravetter, & Davidson (1980) had participants make same / different judgements of pairs of lengths, using vision or haptic touch. Participants in three touch conditions were either told to use measuring techniques, or told not to use measuring, or not given any restrictions. It was found that vision was constantly more accurate, but haptic accuracy increased with the length of the stimuli.

5.4.2 Perception of proportion
Appelle (1991) argues that proportion provides information of a higher order than, for example, extent, orientation or curvature, as these are properties of isolated contours. Proportion is an ‘invariant’ property because it can be transposed across objects of different absolute sizes, and therefore may be more important to
perception. This concurs with Gibson’s opinion that invariant properties are particularly important to perceptual systems (Gibson, 1962) (cited in Appelle, 1991).

In the study described above involving perceived length, Appelle et al. (1980) investigated measuring activity in the perception of proportion. Participants made ‘same’ / ‘different’ judgements on 36 pairs of 3-dimensional rectangles with a range of width-to-length proportions. Participants used either vision or touch, with the touch group in three conditions: measuring, no measuring, no restrictions. Appelle et al. found that those in the measuring condition performed significantly better than the other two haptic groups. Interestingly, the visual condition was not significantly better than the measuring haptic group. The authors hypothesized that when judging proportion subjects do not combine two separate judgements of the extent of the width and length, but use the width of the object as a unit of distance to judge the length of the object.

The above hypothesis was tested in a third experiment. The stimuli were the same as those used in the second experiment, except that the standard was presented vertically, 6mm below the horizontally oriented comparison (to be comparable with the length and width of the stimuli in the first experiment). Subjects gave the ratio of the bottom extent to the top extent.

Appelle et al. found the results from the first and third experiments to be similar except for the no-measuring condition. Appelle et al. conclude that measuring behaviour determines the ratio between the two extents of a rectangle: participants were using the width as a unit of distance to judge length of the rectangle. Appelle et al. also suggest that “proportion is neither a directly nor a spontaneously perceived attribute of form” (Appelle et al., 1980, p161).

Appelle et al. note that their conclusions contrast with those of Revesz (1950) (cited in Appelle et al., 1980). Revesz argued that it is not possible to make a successful comparison of the proportions of two different-sized rectangles by measuring, either haptically or visually. Appelle et al. argue that it is indeed possible and that measuring is the method used by subjects when judging proportion.
5.4.3 Perception of orientation
Several researchers have investigated subjects' ability to adjust the orientation of a comparison rod to match the orientation of a standard rod, for example, Lechelt, Eliuk & Tanne (1976) (cited in Appelle, 1991). In a kinaesthetic condition participants adjusted the rod using their hands. In the visual condition the rod was adjusted via remote control. It was found that stimuli that were oriented horizontally or vertically were perceived more accurately than obliquely oriented stimuli, with both visual and kinaesthetic perception. This is termed the 'oblique effect'. Lechelt, Eliuk & Tanne (1976) suggest that this arises from the proprioception of the position and orientation of the limbs.

In a follow-up study Appelle & Countryman (1986) (cited in Appelle, 1991) hypothesised that the haptic oblique effect was produced by the task of touching oblique stimuli. The movement required to touch a right-leaning stimulus with the left hand is different from that required if touching the same stimulus with the right hand: the left hand moves towards the body whereas the right moves away from the body. Appelle & Countryman eliminated the effect by changing the task performed by subjects so that they did not have to move their hands in different directions in relation to the body.

Appelle & Gravetter (1985) (cited in Appelle, 1991) suggest that the haptic oblique effect may be due to the fact that Lechelt et al.'s (1976) subjects were shown the standards prior to the experiment. Appelle & Countryman (1986) found that this effect was reduced when subjects did not have prior knowledge of the stimuli. Appelle & Countryman conclude that the haptic oblique effect is more likely to be due to the influence of haptic exploratory movements than to haptic sensitivity being different for different orientations.

It is surprising that neither Appelle & Countryman, or any previous researchers in this field, tested this hypothesis regarding visual knowledge of the stimuli with congenitally blind participants.
5.4.4 Perception of curvature

Various studies have further investigated the issue of apparent straightness of curved edges. For example, Rubin (1936) (cited in Appelle, 1991) reported that when subjects felt a concave edge by moving an outstretched arm the edge felt straight, and the effect was even stronger when the subjects moved just their forearm. Rubin concluded that this result was probably due to variations in the pressure on the fingertips.

Goodnow, Baum & Davidson (1971) (cited in Appelle, 1991) found that the centre of a curve is perceived to be displaced in the direction of the movement of the hand. Davidson (1972) (cited in Appelle, 1991) investigated this issue further by analysing video recordings of the exploratory strategies used by blind and sighted subjects. Davidson identified five different styles of exploration: gripping; pinching; fingertip sweeping; and tracing. Davidson concluded that, “exploratory activity may be confused with the property being explored” (Appelle, 1991, p181). In other words, that apparent straightness of a concave edge could be attributed to its correspondence to the natural sweep of the arm. Davidson speculated that use of the gripping style could be the best way to judge a curve because the attention is focused on a different plane to that of the sweeping arm movement.

5.4.5 Identification of objects by touch

Klatzky, Lederman and Metzger (1985) (cited in Loomis & Lederman, 1986) investigated the ability of observers to identify 100 common objects by touch. They observed an accuracy rate of 96% with 68% of the responses occurring within 2 seconds of initial contact and 95% within 5 seconds.

5.4.6 Recognition of objects

Schiff & Dytell (1972) (cited in Schiff, 1980) investigated visual recognition of objects that had been explored through haptic manipulation. The objects were recognised with less than 100% accuracy. Schiff concludes that “the capacity of a sensory system in no way guarantees that the sensitivity for discrimination is ‘useful sensitivity’ (Schiff, 1980, p155 -156).

Revesz (1950) (cited in Appelle, 1991), and more recently, Appelle, Gravetter & Davidson (1980) have suggested that haptic recognition of objects is not immediate,
as it is with vision. Instead, these authors suggest that the perception of different parts of an object is followed by a cognitive process through which the whole is constructed.

5.4.7 Illusions and after-effects in perception
A few studies have investigated the existence of illusions, similar to optical illusions, and after effects in different types of touch perception. For example, Appelle (Appelle, 1991) found that when participants are presented with long and short lengths alternately, followed by the presentation of an intermediate length, kinaesthetic after-effects occur. When a subject inspects a long standard followed by a shorter comparison, the comparison feels even shorter that it would if it was inspected prior to the standard. Appelle reports that this effect has been replicated under a range of different conditions.

Gibson (1933) (cited in Appelle, 1991) reported the existence of kinaesthetic after-effects in the perception of curvature. Gibson found that after 3 minutes’ exploration convex curves (bending away from subject) began to lose their convexity, and straight edges began to feel concave.

Fry (1975) (cited in Appelle, 1991) presented subjects with haptic versions of the Müller-Lyer73 illusion and varied the touching method between active and passive. It was found that the illusory effects were nearly twice as large for active touch than for passive touch. Revesz (1950) (cited in Appelle, 1991) and others have found that when haptic equivalents of visual illusions exist, they do not necessarily occur for the same reasons as the visual illusions, and the direction may not be the same.

5.4.8 Observations of participants’ behaviours
In addition to the observations made by Davidson (1972) of participants’ exploratory movements when judging curves, as described above, Zinchenko & Lomov (1960) (cited in Loomis & Lederman, 1986) have also observed the motions of the hand. Zinchenko & Lomov suggest that the motions of the hand while acquiring

73 The Müller-Lyer illusion is demonstrated with “a pair of arrows whose shafts are of equal length, one having outgoing and the other ingoing arrow heads at each end. The one with the outgoing heads looks considerably longer, though it is in fact the same length” (Gregory, 1990, p138).
information about an object, parallel the motions of the eye during visual perception. Tactual perception involves larger and smaller movements of the hand. The larger movements (macromotions) acquire information about the object, while the very small motions of the hand (micromotions) continue the excitation of the receptors in order to ensure the continuity of the tactile image.

5.4.9 Perception of form by blind people
Hunter (1954), Ewart & Carp (1963), and Davidson (1972) all conducted studies involving blind people. Hunter and Davidson both found that blind people were more accurate in judgements of straight, concave and convex edges. Davidson suggests this is because the blind participants mostly used a ‘gripping’ style of exploration, which was thought to be more accurate than other styles of exploration, mostly used by sighted participants. In a review of studies involving perception of form by blind and sighted people, Ewart & Carp found that some studies found blind people performed better than sighted people, others found the reverse, and still further studies found no difference between the two groups. In their own study Ewart & Carp found no difference between blind and sighted children when making same / different judgements of 2-dimensional objects.

5.4.10 Factors that appear to have impact on perception of form
Previous research has identified several factors which affect perception of length, proportion, curvature etc. These factors are summarised below. Several variables have been identified as being related to perceived extent or length. The orientation of the stimulus in relation to the body has been observed to be a key factor by several researchers. Lengths that are parallel to the body have been found to be perceived to be longer than lengths pointing away from the front of the body, for example, Davidon & Cheng (1964), Seizova-Cajic (1998). The direction of the movement of the hand or arm in relation to the body has also been observed to be an important factor in the perception of length / extent, for example Brown et al. (1948), Seizova-Cajic (1998). Stimuli that are observed with the hand moving away from the body have been found to be judged more accurately than with the hand moving toward the body. The direction of movement in relation to the stimulus has also been found to be a factor in the perception of curvature (Goodnow et al., 1971).
The part of the body used and the touching style have been observed to be important factors in the perception of length, proportion, and curvature. Jastrow (1886) and Stanley (1966) both found that lengths explored with just the hand were perceived to be longer than those explored using the whole arm. Appelle et al. (1980) found that participants judged proportion more accurately when they used a haptic measuring method. Rubin (1936) found that curved stimuli were perceived to be straight when they were felt using an outstretched arm or just the forearm. Davidson (1972) found that a 'gripping' style of touching enable blind people to perceive curved stimuli more accurately than people who used other styles of touching.

The gripping style of exploration of curved edges was thought to be more accurate than other styles because attention is focused on the fingers and hand, rather than on the sweep of the arm (1972). Ewart & Carp (1963) found a significant relationship between performance and IQ for blind people when performing same / different judgements of shapes. However, they found no difference between in performance for blind and sighted children, and concluded that 'visual imagery' is not a critical factor in this type of task. Ewart & Carp also found that some shapes were judged more accurately than others (there were more errors with the semi-circle, quarter-circle and crescent than with the square, rectangle and parallelogram). Jastrow (1886) (cited in Appelle, 1991) found that touching method influenced perceived distance.

The direction in which the exploring hand moves has an effect on the perception of extent: movements away from the body were more accurate than movements towards the body (Brown et al., 1948) (cited in Appelle, 1991). A radial-tangential effect has been identified: the segment of a L-shape lying in the median plane, or the radial length (parallel to the front of the subject’s body) is perceived to be 10% longer than the segment lying in the fronto-parallel plane, or the tangential length (pointing away from the subject’s body) (Loomis & Lederman, 1986).

Seizova-Cajic (1998) found that length was perceived to be shorter via kinaesthetic perception than via visual perception and the difference between the two increased with the size of the stimuli. Also, the direction the hand is moving (away from or towards the body) has been found to affect perceived length.
There is debate regarding whether participants can, and do, use measuring strategies
in judging proportion or ratio. Haptic oblique effects seem to exist in the perception
of obliquely oriented stimulus, although there is debate about the cause of these
effects (Lechelt et al. (1976), Appelle & Gravetter (1985), Appelle & Coutryman
(1986).

The researcher believes that some of the findings summarised above can be
explained by comparing the experimental conditions with 'real world' human
activities. For example, when we touch an object, our initial contact is probably with
the edge or end that is nearest to us, rather than the aspect that is furthest away. We
may then move our hands along the object, away from our body to feel the rest of the
object. Therefore in experimental conditions the behaviour that most closely
matches our non-experimental behaviour is likely to be more accurate.

The researcher finds that there are many factors that can compound the results in
different types of experiment. For example, as Appelle & Graveteer (1985) point
out, prior visual knowledge of the stimuli that are to be explored by touch could
affect the results.

5.5 Use of touch in computer interfaces
The sense of touch has been incorporated into computer interfaces, particularly in
virtual reality (VR) systems, to supplement or replace visual displays. Such
interfaces are often referred to as 'feedback' devices. There are three main types of
feedback given by such systems.

- Tactile feedback constitutes sensations applied directly to the skin, usually in
  response to contact with a virtual object. Tactile feedback cannot resist against
  the movement of the user’s hand.
- Force feedback constitutes sensations of weight or resistance of objects in a
  virtual world. Force feedback can resist against the movement of the user’s hand.
- Haptic feedback is used to include both tactile and force feedback. It can
  therefore resist against the movement of the user’s hand.
  (Burdea, 1996)

Haptic interfaces incorporate actuators that provide force or tactile feedback to the
user. A variety of actuators are currently in use. They range from electric motors
and hydraulic pistons to pneumatic muscles. This review will focus on force and haptic feedback in user interfaces.

5.5.1 Brief history of the use of touch in interfaces
Before the advent of VR in the 1970s there were systems that incorporated the sense of touch in 'master / slave' systems. In such systems the user operated the 'master' that was connected via mechanical linkages to the 'slave', which performed the task. The nature of the linkages meant that the slave had to be in close proximity to the master.

The first known system to provide force feedback to the user's hand from a remote slave (as opposed to a near slave) in a tele-manipulation system was developed in 1954 (Burdea, 1996). The first system to provide force feedback to the user's fingers was developed in 1966. The user wore a glove that measured the movements of the fingers. It then transmitted the information to a remote slave, and provided feedback to the user. These early systems required the master and slave to have the same configurations. This meant that masters were bulky and could only work with one slave. By the mid-1980s new control systems meant that masters became smaller and could operate a variety of slaves.

What we now call 'virtual reality' emerged in the late 1970s and at that time the first tactile and force feedback devices were developed. The first commercial force feedback devices for VR became available in the early 1990s. These complemented the visual and sound feedback already available in computer interfaces.

5.5.2 Brief survey of haptic devices: their characteristics and potential applications
This section briefly describes some of the range of existing haptic interfaces to Virtual Environments (VEs). Such devices can be distinguished in terms of the nature of the interface between the user and the device, for example whether a thimble, handle, probe, joystick, laparoscope, or a sliding foot-mounted device is used. The nature of the interface in turn defines which part of the user’s body is used to interact with the device.
The range of devices can also be distinguished by the types and nature of the items that can be displayed, for example objects, textures, fluids, surfaces and springs. A further distinction between such devices is whether or not they are used in conjunction with a visual display.

### 5.5.2.1 Haptic devices

Devices that provide haptic information include the Phantom, the Sandpaper system, the Impulse Engine 3000, and the Haptic Master. The Phantom is a ‘universal force-reflecting’ interface which displays virtual objects. It has two user interfaces: a thimble in which the user places a finger, and a stylus, which is held like a pen. The thimble and stylus are both attached to a linkage. To feel the objects the user moves the linkage around via the thimble. The thimble can be replaced with a stylus, which would allow the user to control a virtual pencil or paint brush. The Phantom therefore has several different areas of potential application: as an interface for teleoperation; as an input device for CAD; moulding virtual clay; and practising surgical procedures on virtual patients (MIT haptics group\(^{74}\)). Two studies involving the Phantom are described later in this chapter, in Sections 5.5.2.3 and 5.7.

The Sandpaper joystick system developed at the Massachusetts Institute of Technology (MIT) (Minsky, Ouh-Young, & Steele, 1990). The system simulated virtual textures that could be felt by the user (as described in Section 5.6.1).

The Haptic Master can also display virtual objects, which have either hard or elastic surfaces. It can also simulate the rigidity and weight of virtual objects, and the velocity and vorticity of fluids. Force-feedback is given via a handle that is supported by three sets of pantographs\(^{75}\).

### 5.5.2.2 Haptic devices with visual display

Two devices that provide both haptic and visual information are the Walkthrough Simulator and the WYSIWYF (What You See Is What You Feel) Display. The

---


\(^{75}\) For further details of the Haptic Master see: http://intron.kz.tsukuba.ac.jp/HM/txt.html (checked Nov 2000).
Walkthrough Simulator simulates buildings and spaces, rather than abstract objects. The ‘walker’ wears omni-directional sliding devices on their feet, a hoop around their waist, which limits their body movement, and a head-mounted display which presents the virtual scene. The Simulator can be used by designers and architects to show their designs of buildings or open spaces to clients.

The WYSIWYF Display combines a real-time video image of a real scene, including the user’s hand, with a graphical image of the virtual scene. The user views this combination of real and virtual scenes on a LCD display. An example of its potential applications is ‘virtual tennis’ in which the player uses a real tennis racket to hit a virtual tennis ball. The force-feedback is provided by a robot arm. This type of display can be used in training visual-motor skills such as those used in surgery. An advantage of the integrated display is that it overcomes the problem of the visual display being separate from the haptic display and the situation of ‘feeling here but looking there’.

5.5.2.3 Haptic interfaces for use by visually impaired people

Haptic devices have clear potential to be used by blind and visually impaired people as such devices can complement the access to visual interfaces provided by speech output. However, this potential is not reflected in the current development of haptic interfaces. There are only two known studies that involve visually impaired people: the Phantasticon (Rassmus-Gröhn & Sjöström, 1998), which combines a Phantom interface and a visual display; and the current study involving the Impulse 3000, which does not use a visual display. The authors believe that the Phantasticon can be used by visually impaired children for learning mathematics and painting colour pictures. To paint, the user can choose a colour from the screen and then paint with it by moving the Phantasticon with their finger. The harder they push, the thicker the line becomes. The user can then change the program mode and explore the structure of the picture with their finger (Rassmus-Gröhn & Sjöström, 1998).

---


77 For further details of the WYSIWYF Display see: http://www.cs.cmu.edu/afs/cs.cmu.edu/project/mst/www/virtual/virtual_desc.html (checked Nov 2000).
5.5.3 *Perception via haptic interfaces*

In the investigation of haptic perception with a view to its incorporation into VR systems there are several complicating factors. The characteristics of the receptors in the skin are not clearly understood, such as their temporal and spatial characteristics, and their adaptation to stimuli. Burdea finds that there is no consensus on the spatial resolution of receptors or the maximum forces that they can withstand.

Nevertheless, Burdea states that knowledge of the tactile and force sensing and control of humans is essential for the design of human-computer interfaces that incorporate such sensing and control.

5.5.3.1 *Implications of perception for interfaces*

Burdea concludes that the characteristics of haptic perception cannot be extrapolated into guidelines for the design of ideal haptic interfaces. This is because the range of tasks that could be performed cannot be simulated by just one interface. There may also be conflicting requirements for a system. For example, large forces are required in order to simulate hard contact with an object, but user fatigue must be kept to a minimum. Burdea finds it is also risky to base design on the few studies that have been conducted because they involve so few participants and a variety of different interfaces. There are many characteristics that must be taken into account in the design of optimal interfaces (Burdea, 1996).

5.5.4 *Modelling*

In order for a haptic interface to provide feedback to the user the system requires input for the physical modelling of objects. This includes their properties and behaviour, for example to simulate collisions, grasping, surface deformation, hard contact, gravity, and friction.

5.5.5 *Quality of feedback*

Jex (1988) (cited in Burdea, 1996) defined four tests to evaluate the quality of a haptic interface, which are described below.

- Simulate a piece of light balsa wood, with negligible inertia, friction or vibrations
- Simulate a crisp, hard stop.
• Simulate friction without sponginess or jitter.
• Simulate a mechanical catch with crisp transitions and no lag.


Jex seems to imply that 'sponginess' or 'lag' should not exist within systems. However, it may not be necessary to eliminate them altogether, just to a sufficient level that the user cannot perceive them. For example, there may be levels of time lag that exist within a system, which cannot be perceived by the user. To the researcher it seems more important that factors such as 'sponginess' or 'lag' should not interfere with the user’s task, rather than necessarily being eliminated. To ensure that the user can perform the task effectively and efficiently the perception of these factors needs to be investigated.

5.5.6 Human factors in haptic feedback devices

Burdea states that human factors studies can be used to “measure the benefits brought by haptic feedback to the simulation of realism and the user’s sense of immersion” (Burdea, 1996, p225). The researcher believes that human factors studies should not be restricted to realism and immersion. There are some applications of haptic feedback that do not necessarily have to simulate realism. For example, a tactile map of a building or campus does not have to feel ‘real’ and does not require immersion, but it does need to include representations that make sense to the user. This is where it is essential to investigate what users can perceive in a haptic virtual environment. For example, whether they can distinguish different shapes, sizes and textures, as discussed in Chapter 6 of this thesis. Also, many virtual environments do not need the user to be ‘immersed’ in the environment. There are few systems where full immersion is necessary, for example some training systems. Most systems that incorporate haptic feedback provide additional feedback to existing visual or auditory feedback in, for example, graphical user interfaces.

Burdea states that there have been few human factors studies of haptic feedback in VR. Burdea finds that such studies often involve too few participants for the results to be highly statistically significant. Also the studies are difficult to compare because they involve different systems. This reflects the constant change in the technology.
Burdea briefly reviews some of the existing human factors studies. Hannaford & Wood (1989) (cited in Burdea, 1996) found that force feedback reduced task time by 30% and error rates by 60%, compared to manual performance or visual feedback. Howe & Kontarinis (1992) (cited in Burdea, 1996) found that force feedback alone was not enough for the effective manipulation of a fragile object; that tactile feedback was also required. Akamatsu, MacKenzie, & Hasbroucq (1995) (cited in Burdea, 1996) found that tactile feedback on a mouse device decreased performance time in a task that involved positioning the cursor over a target in a graphical user interface. None of the studies above involved blind participants.

In contrast to these positive effects of haptic and tactile feedback, Burdea highlights some negative effects, such as motion sickness. This mainly occurs in immersive systems that incorporate a motion platform, for example those used for training pilots. Studies have found that motion sickness can be reduced by reducing the time lag between visual and motion cues.

5.5.7 Applications of haptic feedback devices
Burdea (1996) describes many applications in which haptic feedback can be used. Medical applications include the training of doctors in the palpation of tissue, performing epidural anaesthesia, and minimally invasive surgery. The benefits of training doctors using simulations are clear: they are able to have a realistic experience without any potential danger to patients.

Haptic feedback can be used in the rehabilitation of stroke patients with paralysed limbs. They can assist and guide the patient's limb during exercises. Haptic feedback can increase the realism of arcade games, home-based games, interactive art and virtual museums. In tele-robotics haptic feedback has been used for many years to assist the performance of tasks in adverse or remote environments. Haptic feedback has also been used in military training, for example in the training of pilots. Burdea concludes that "the spectrum of VR simulations using haptic feedback will increase as the underlying technology matures and its price becomes more affordable" (Burdea, 1996, p277).
5.5.7.1 Virtual Environments on the World Wide Web

Three-dimensional virtual environments (VEs) can be created using the Virtual Reality Markup Language (VRML) and can be accessed via the Web. The VEs that currently exist on the World Wide Web (the Web) mainly contain graphical objects and scenes, and these are not accessible to blind users. There is potential for haptic interfaces to allow blind users to interact with these VEs on a tactile basis instead of on a visual basis and thus improve the level of access to the Web.

Goralski, Poli & Vogel (1997) provide a scenario for the use of VRML. In this scenario a person who is interested in the Great Pyramids (who may be a student or an Egyptology specialist) and who has access to the Web, would be able to search the Web and quite easily find information on their chosen topic. Currently, this information is likely to be text format, accompanied by 2-dimensional graphics. However, with VRML the information could contain a simulation of the Pyramids that the user could walk around, climb and explore inside. The Web site could also incorporate text to be associated with the virtual world; more text than any book could contain. Sections of this text could be spoken and other sounds could also accompany the simulation. Goralski et al. believe that the activities that could be performed in this virtual world of Pyramids could not be performed in reality, even on a trip to Egypt. VRML therefore, according to Goralski et al. has the potential for enabling Web users to experience more of the world than is possible in the 'real' world.

5.6 Perception of Virtual Textures

As well as the research reported in this thesis, two other researchers have previously studied the psychophysics of virtual textures: Minsky (1995) and Jansson (1998).

5.6.1 Measuring confidence and roughness

Minsky (1995) used a virtual environment known as the Sandpaper system, which was operated via a joystick. The system was developed with the purpose of studying virtual textures. Minsky examined the perception of two different types of textures. The first type were 'textured surfaces', some of which were analogous to the stimuli used by Lederman (1974; 1981; Lederman & Taylor, 1972) i.e. surfaces with square-shaped grooves. The second type of texture were 'non-textures'. These were
surfaces that were either smooth or had grooved surfaces with a very large period. All participants in the study were sighted.

The first experiment measured participants' confidence that a 'textured surface' was being displayed, rather than a 'non-texture'. High levels of confidence were found for the textured surfaces and low levels of confidence for the non-textures. Force amplitude was found to contribute to the confidence ratings.

The second experiment involved the quantitative measurement of roughness. Using the magnitude estimation technique (described in Section 5.3.4.1), Minsky investigated the factors that contribute to perceived roughness of texture. A psychophysical function of force amplitude to perceived roughness was found.

The third experiment involved qualitative measurements of roughness and explored the perceptually salient features of texture. A similarity-sorting technique was used, where textures were sorted into 2, 3 then 4 groups of roughness-similarity. Force amplitude was found to be primarily salient for roughness sorting, with spatial period having secondary salience.

Minsky found that a psychophysical power law held between the perception of the virtual textures and their simulated properties although this was mediated by the force produced by the device rather than the spatial period of the simulated textures. Minsky hypothesised that this was due to the fact that the textures were felt via a joystick rather than by rubbing a finger over the surface, as participants had in previous studies of real textures.

5.6.2 Comparison of real and virtual textures
Jansson (1998) used the Phantom device (described in Section 5.5.2.1) to investigate the perception of texture, using the magnitude estimation technique. Jansson compared the perception of four real textures with four virtual textures. The real textures were sandpapers with grit sizes of 50, 80, 120 and 220. The virtual textures were not exact copies of the real textures but Jansson states that “the properties used in the simulation [were] intended to be sufficient for allowing accurate perception of roughness of the sandpapers” (Jansson, 1998, p107). The participants were
blindfolded sighted people. The Phantom device was used with a stylus attachment that was held by the participant like a pen. The real textures were also felt using a stylus so that the results could be compared.

Overall, the real textures were perceived to be less rough than the virtual textures, but this was not a significant difference. Jansson intends to repeat the experiment with blind participants, but does not expect large differences between blindfolded sighted participants and blind participants, especially not in terms of the difference between the real and virtual stimuli.

Jansson does not say why the particular grit sizes were chosen. The intention may have been to compare the results with those of Minsky (1995). Although Jansson states that the virtual textures were ‘intended to be sufficient’ to allow comparison with the real textures, it is not stated what the properties of the virtual textures were. This makes it difficult to replicate the study or compare the results with other studies. The results showed that the two groups of textures were perceived to be different (although not significantly so). It is not clear whether these differences were due to whether the textures were real or virtual, or due to the different characteristics of the textures.

5.6.3 Comparison of finger and probe exploration of textures
Early studies on the perception of texture mostly involved participants touching the stimuli with their bare finger. Overall, it was concluded that the spacing between the elements making up the surface of the texture was the main determinant of perceived roughness. For example, Stevens and Harris (1962) found that perceived roughness increased with the interelement spacing between the grit particles on sandpapers, and Lederman and colleagues found that perceived roughness increased with wider groove width on plates engraved with parallel grooves (Lederman, 1974; 1981; Lederman & Taylor, 1972). A classic psychophysical power function was obtained with a linear relationship between the log of the interelement spacing variable (e.g. grit particle distance or groove width) and the sensation of roughness as measured by magnitude estimation.
In the early 1990s an inverted U-shaped function (as opposed to a classic linear power function) was obtained for 2-D dot patterns touched with a bare finger (Connor, Hsiao, Phillips, & Johnson, 1990; Connor & Johnson, 1992) (cited in Lederman, Klatzky, Hamilton, & Ramsay, 1999). It was found that the peak of the quadratic equation was around 3.5mm. This means that for interelement spacing of less than 3.5mm a positive exponent was obtained, but for interelement spacing of more than 3.5mm a negative exponent was obtained. When a negative exponent is obtained it means that perceived roughness decreases with increased interelement spacing.

In the last decade haptic devices have been developed as user interfaces to virtual environments, and such environments can contain textured surfaces. Several researchers (including the current researcher) have recognised the need to investigate how virtual textures are perceived. This includes the investigation of the effect of using a mediating device, such as a probe or thimble, to explore the textures rather than feeling the texture directly with the finger, as is most common in real life (Jansson, 1998; Minsky, 1995; Minsky & Lederman, 1996). In order to investigate the effect of mediating devices, some researchers have used the probe of a haptic interface to explore virtual textures (Colwell et al., 1998; Jansson, 1998; Minsky, 1995; Minsky & Lederman, 1996). Other researchers have had participants explore real textures with a probe or stylus (for example, Klatzky & Lederman, 1999). It has been found that the type of touching (either finger or probe) is an important factor in the perception of textures. Klatzky & Lederman’s study was conducted since the current studies were conducted. It is therefore presented in more detail in the discussion of study of the perception of virtual textures presented in the next chapter (Section 6.6.1).

5.7 Perception of Virtual Objects

Jansson (1998) is the only other researcher known to have studied the perception of virtual objects. Jansson investigated the ability of participants to identify four types of virtual object, cubes, spheres, cylinders and cones. Each object type had three different sizes, 5, 25, 50mm (maximum height and width). The task was to identify
the form as quickly and accurately as possible. There were 3 trials in which each of the 12 objects were presented.

The accuracy rate for spheres was 100% for each size. The rate for cubes over 90% for each size. The accuracy rates for cylinders and cones were just above 70% for 5mm sizes, around 90% for 25mm, and above 90% for 50mm. Jansson reports that 52% of the mistakes were made during the first trial and that if just the second and third trial are analysed, the overall rate of accuracy would have been 95%. Jansson states that this demonstrates quite a rapid learning effect. This statement suggests that the participants were given immediate feedback on their responses, although this is not explicitly stated.

The exploration times for each object varied between 10 and 40 seconds. The spheres had the shortest exploration times, followed by the cubes, cylinders, and cones. For all object types the smallest sizes had the longest exploration times, and for all except the spheres, had the lowest accuracy rates. Jansson offers no explanation for these findings. The researcher believes that the high exploration times for the smallest objects may be because participants were searching for 5mm objects within a relatively large virtual space.

5.8 The implications of previous haptic research on the current studies

The study involving virtual textures presented in the next chapter (Chapter 6) was informed by previous studies involving both real and virtual textures. Lederman (1974; 1981; Lederman & Taylor, 1972) conducted several studies to investigate the perception of real textures using a set of 10 textures (as reviewed in Section 5.3.4.2 above). These textures had a square-shaped groove, although Lederman would have preferred to use sinusoidal-shaped grooves (however she was limited by the technology available to her at the time). Lederman varied several characteristics of the textures and found that perceived roughness was mediated by groove width. Minsky (1995) extended the work of Lederman in an investigation of the perception of virtual textures (as reviewed in Section 5.6.1 above). Minsky used square-shaped grooves which were comparable with those used by Lederman. Minsky found that
perceived roughness was mediated by the force of the device used to simulate and explore the virtual textures.

The current study extends the work of Lederman by using 10 virtual textures with varying groove widths. These textures had (or were intended to have) the same dimensions as those used by Lederman (and Minsky), but had sinusoidal-shaped grooves, which Lederman would have preferred to have used instead of square-shaped grooves. The current study also extends Minsky’s work in that virtual textures were investigated using a different haptic device to that used by Minsky.

In order to assess the perceived roughness of the virtual textures, the current study adopted the same technique used by Lederman and Minsky; the magnitude estimation technique. This technique was first used by Stevens & Harris (1962) in the investigation of perceived roughness (as reviewed in Section 5.2.4.1). Stevens & Harris found that Stevens’ Power Law that held for perceived brightness, length etc, also held for perceived roughness. This technique was therefore a valid and appropriate technique to use in the current study.

The researcher found that there were no previous studies that investigated the overall size of objects that could be used to inform the current study involving virtual objects, presented in Chapter 6. Previous studies of real objects have investigated specific characteristics of a range of objects, such as perceived length, proportion, orientation, and curvature (reviewed in Sections 5.4.1, 5.4.2, 5.4.3 and 5.4.4 respectively). The researcher believes that size is a fundamental characteristic of objects and therefore it was relevant to conduct a study involving the perception of the size of virtual objects.

5.9 Conclusions on touch perception
This chapter has identified three different types of touch (tactile, kinaesthetic, and haptic) and has focused on haptic perception of textures and objects. The literature reviewed has presented what is known regarding the perception of real and virtual textures and objects. For example, Stevens’ power law holds for texture (as well as for brightness, length, etc) (Stevens & Harris, 1962). Lederman found that groove width and hand force are the main contributing factors to the perception of roughness
Minsky found that Stevens’ power law holds for virtual textures, but is mediated by the force of the device (Minsky, 1995; Minsky & Lederman, 1996). Jansson found that real textures are perceived to be less rough than virtual textures, although not significantly so (Jansson, 1998; 1999). Jansson also found that virtual objects of different shapes can be recognised, although the accuracy rates and the recognition time depends on the different shape (Jansson, 1998; 1999).

The previous studies on the perception of texture have differed in a number of ways. The early studies involving real textures mostly used rectangular groove shape, whereas more recent studies involving virtual textures have used sinusoidal shaped grooves. The early studies also required participants to use their fingertips to feel surfaces, whereas in recent studies of real and virtual surfaces have used a probe of some kind. Most studies of real and virtual textures have involved sighted participants; only studies conducted recently have involved blind people.

Clearly, there are benefits to the use of touch in interfaces because of the potential to supplement visual output for sighted people, and speech output for blind people. Regarding access to the Web for visually impaired people, the researcher believes that the use of haptic interfaces to browse the Web as an alternative to speech output is not necessarily viable, or desirable. The use of haptic information to supplement speech output seems to be a more realistic and attainable objective.
6 Chapter 6 – Investigations of the Perception of Virtual Textures and Objects with a Haptic Interface

6.1 Introduction ........................................................................................................... 252
6.2 Method ................................................................................................................. 253
  6.2.1 Design .............................................................................................................. 253
    6.2.1.1 Virtual textures ......................................................................................... 253
    6.2.1.2 Simple objects ......................................................................................... 254
    6.2.1.3 Demonstration of complex objects ......................................................... 254
    6.2.1.4 Location of virtual space ......................................................................... 254
  6.2.2 Participants ..................................................................................................... 255
6.3 Apparatus and Materials ...................................................................................... 256
  6.3.1 The device ....................................................................................................... 256
    6.3.1.1 Different presentations of objects .......................................................... 257
  6.3.2 Stimuli ............................................................................................................. 258
6.4 Procedure ............................................................................................................. 261
  6.4.1 Textures .......................................................................................................... 261
  6.4.2 Simple objects .................................................................................................. 262
  6.4.3 Complex objects ............................................................................................. 263
  6.4.4 Location of virtual space .............................................................................. 263
6.5 Results ................................................................................................................ 264
  6.5.1 Textures .......................................................................................................... 264
    6.5.1.1 Significant relationship between stimulus and perception ....................... 265
    6.5.1.2 Difference between blind and sighted participants ................................. 265
    6.5.1.3 Positive and negative exponents ............................................................. 265
  6.5.2 Simple objects .................................................................................................. 266
    6.5.2.1 Cubes ........................................................................................................ 266
    6.5.2.2 Spheres ..................................................................................................... 266
    6.5.2.3 Rotated cubes ........................................................................................... 267
    6.5.2.4 Sheared cubes ......................................................................................... 267
  6.5.3 Complex objects ............................................................................................. 273
  6.5.4 Location of virtual space .............................................................................. 273
6.6 Discussion on the perception of virtual textures and objects .............................. 274
  6.6.1 Textures .......................................................................................................... 274
  6.6.2 Simple objects .................................................................................................. 277
    6.6.2.1 Rotated cubes ........................................................................................... 278
    6.6.2.2 Sheared cubes .......................................................................................... 278
    6.6.2.3 Contribution of experience to perception .............................................. 279
    6.6.2.4 Recent research on perception of virtual objects ................................... 279
  6.6.3 Complex objects ............................................................................................. 280
6.1 Introduction

This chapter presents two studies in which a haptic feedback device was used to present virtual textures and objects. The aim of the studies was to investigate the perception of these virtual textures and objects by blind and sighted people. The first study involves virtual textures and extends previous research on the perception of real and virtual textures. The second study involves simple virtual objects. This chapter also describes a demonstration of complex virtual objects given to the participants. The studies are the first known to examine the perception of virtual textures and objects by blind people.

It is important to examine how virtual artefacts are perceived in order to design virtual environments (VEs) with which the user can interact effectively. If the user of a VE needs to be able to accurately judge the shape and size of an object, or discriminate between different surfaces, the VE needs to be designed to allow this. In other words, the differences between objects or between surfaces need to be discriminable by the user.

For blind people the advantage of the use of touch in computer interfaces, including interfaces to VEs, is that touch can supplement the usual mode, speech output. In relation to the Web, this approach has potential for complementing the approach improving accessibility through the design of the content of Web pages. As has been

---

78 It should be noted that a calibration error was discovered in the IE3000 after the completion of the studies reported in this thesis. The impact of this error on the results reported in this Chapter is explained in detail in the Addendum (located after the Appendices).
seen in previous chapters Web page authors have difficulty implementing the
techniques suggested by accessibility guidelines (Chapter 3) and the use of such
techniques does not necessarily improve accessibility for visually impaired people
(Chapter 4). It is therefore important to examine the use of haptic devices to improve
access to the Web for visually impaired people.

It is important that the perception of virtual textures and objects by blind people is
compared to that of sighted people. This will allow interfaces and VEs to be
designed for use by both groups. There have been few human factors studies
involving haptic feedback devices, so little is known about the perception in virtual
environments (Burdea, 1996).

6.2 Method
6.2.1 Design
6.2.1.1 Virtual textures
The first of the two studies investigated the perceived roughness of ten virtual
textures. This study extends the work of Stevens & Harris (1962) and Lederman
Williams, 1982; Lederman & Taylor, 1972) who investigated the perceived
roughness of real textures. It also extends the work of Minsky (1995) and
complements the work of Jansson (1998) who have investigated the perception of
virtual textures.

The study had a mixed design with three independent variables, one between groups
and two within groups. The between groups independent variable was whether the
participants were blind or sighted. The two within groups independent variables
were a) the physical groove width of the textures felt by the participants (10 levels,
see Section 6.3.2.1.1 below) and b) the trial (each participant felt each texture 6
times). The dependent variable was the sensation of roughness of the texture as
measured by the magnitude estimation technique.

As described in the previous chapter, the magnitude estimation technique provides
data about the relationship between the physical parameters of a stimulus and
participants' perception of the stimulus. Stevens, Lederman, Minsky and Jansson have used this technique to investigate the perception of the roughness of real and virtual textures.

6.2.1.2 Simple objects

The second of the two studies investigated the perceived size or angle of simple virtual objects, such as cubes and spheres. A full factorial design was not used in this study because informal testing of the device found that the presentation of several objects was problematic, as described in Section 6.3.2.1.2 below. The researcher conducted a ‘trial run’ of the proposed procedure with three colleagues. These ‘trial participants’ found the exploration of the simple objects quite difficult and soon became fatigued and bored. The researcher therefore decided to limit the number of trials with objects to 18 as this seemed to be a reasonable number, and allowed a sufficient range of objects to be presented. (No data were collected from this trial run of the procedure.)

The between groups variable was whether participants were sighted or blind. The within groups independent variables were the size or angle (of shear or rotation) of the virtual objects and the mode of exploration (from inside or outside the object). The dependent variable was the perceived size of the virtual objects.

6.2.1.3 Demonstration of complex objects

In order to demonstrate the use of complex objects composed of simple objects participants were shown one or more such objects, such as a kitchen chair and an armchair. This demonstration explored whether such objects could be recognised, and if the composite objects could be identified. This demonstration did not have a formal set of conditions as it was intended to explore the way in which the participants perceived complex objects.

6.2.1.4 Location of virtual space

Hardwick, Furner, & Rush (1997) observed that people differed in terms of where they imagined the virtual objects to be located in relation to a haptic device. Some people had a mental image of the virtual space being outside the device, so that
virtual objects are felt to be near the hand and are touched by the end of the probe that they hold (see Figure 6.1, a, below). In contrast, others imagine the virtual space to be within the device, so that the virtual objects are touched by the other end of the probe (see Figure 6.1, b, below). Data on this concept were collected from most of the participants in these studies.

Figure 6.1: Different mental models of the location of virtual space: a) outside and b) inside.

Drawing by Andrew Hardwick, reproduced with permission.

6.2.2 Participants

The same twenty-two participants took part in both studies. Nine participants were blind and 13 were sighted. This number of participants is comparable to the numbers used in previous similar studies. Stevens & Harris (1962) used 20 participants, Lederman (Lederman, 1974; 1981; Lederman & Abbott, 1972) used a mean of 8 participants, and Heller (1989) used 10 late blind and 10 early blind participants (in a comparison of the two groups). In the current study six of the sighted participants were female and all the other participants were male. The participants ranged in ages from 18 to 65, the mean age being 32 (median = 30). The sighted participants were all university students, from a variety of disciplines. The blind participants were all employed in computer-related jobs, except two: a student on a computer science course, and a retired audio engineer.

Six of the 9 blind participants were either born without sight or lost their sight by the age of 30 months. The other 3 blind participants lost their sight between 8 and 26 years of age.
6.3 Apparatus and Materials

6.3.1 The device
The device used in the study was the Impulse Engine 3000™ (IE 3000) (shown in Figure 6.2 below). The hardware was developed by the Immersion Corporation and was on loan to the University of Hertfordshire from British Telecommunications Plc. Andrew Hardwick of British Telecommunications Plc wrote the software.

The approximate dimensions of the device were: height = 29cm; width = 29cm; depth = 25cm. As the dimensions indicate, the device can be placed on a desk next to a computer.

A user interacts with the device using a ‘probe’. In Figure 6.2 below this can be seen protruding from the front circular plate. The probe is approximately 17 cms in length and 1 cm in diameter. The probe has 3 ‘degrees of freedom’ of motion, i.e. it can be moved in 3 spatial dimensions: forwards and backwards, up and down, and left and right (and any direction between these). Any reference in this chapter to the probe moving left or right means that the end held by the user is moving left / right (while the other end is moving right / left).

The device provides feedback to the user by monitoring the position of the probe when it is moved by the user, and altering the force accordingly. The probe is connected to three motors at the centre of its length and is therefore pivoted about its centre. The force is created by the three motors, which exert resistance against the probe. This gives the user the impression that a texture or object is present and allows them to feel the properties of the texture or object.

---


80 See Hardwick et al. (1998) for a comprehensive description of the software, and Hardwick (1997) for a technical description of the device.
The motors of the IE 3000 are capable of withstanding 8 Newtons (approximately 2 lb. of force) from the user. If the user exerts more force than this against an object, they may have the sensation of pushing through the object and coming out the other side. Under normal conditions the user does not have to exert much force (far less than 8 Newtons) in order to feel objects and textures; just enough force to maintain contact with the surface.

The IE 3000 uses a combination of force and tactile feedback to display textures and objects, and their surfaces. Burdea (1996) defines this combination as ‘haptic feedback’.

6.3.1.1 Different presentations of objects
The device can present simple objects, such as cubes and spheres, from either the outside or the inside; i.e. an object can be felt from the outside or from the inside of the object. When exploring the inside of an object, it is as if the user is inside the object and they cannot feel the outside of the object. The opposite is true when a user is exploring the outside of an object in that they cannot feel the inside of the object. An analogy with the real world might be exploring the outside of a closed box but not being able to explore inside it, and then getting inside the box, closing it, and exploring the inside of it, but no longer being able to feel the outside of the box.
6.3.2 Stimuli
Two types of virtual stimuli were used in the studies: horizontal textured surfaces, and simple 3-dimensional objects, such as cubes and spheres.

6.3.2.1.1 Textures
Ten virtual textures with a range of groove widths were used. The researcher created the virtual textures by setting the values for the attributes groove width, amplitude, and groove shape. The grooves ranged in width from 0.675mm to 2.7mm (in steps of 0.225mm) (as shown in Table 6.1 below) and had a fixed amplitude of 0.1125mm\textsuperscript{81}.

<table>
<thead>
<tr>
<th>Texture number</th>
<th>Groove width (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.675</td>
</tr>
<tr>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>3</td>
<td>1.125</td>
</tr>
<tr>
<td>4</td>
<td>1.35</td>
</tr>
<tr>
<td>5</td>
<td>1.575</td>
</tr>
<tr>
<td>6</td>
<td>1.8</td>
</tr>
<tr>
<td>7</td>
<td>2.025</td>
</tr>
<tr>
<td>8</td>
<td>2.25</td>
</tr>
<tr>
<td>9</td>
<td>2.475</td>
</tr>
<tr>
<td>10</td>
<td>2.7</td>
</tr>
</tbody>
</table>

A key difference between the textures used by Lederman and colleagues (Lederman, 1974; 1981; Lederman & Taylor, 1972) and those used in this study was that those used by Lederman had a rectangular waveform whereas the textures used in this study involved sinusoidal shaped grooves, as shown in Figure 6.3 below.

\textsuperscript{81} These values are different from those reported in papers published on this work due to the calibration error referred to in footnote 78 (page 252) and explained in the Addendum, located after the Appendices.
Lederman wanted to use textures with sinusoidal grooves, but was not able to create them due to technical difficulties and so used rectangular grooves. For this study it would have been preferable to use virtual textures with rectangular grooves to match those used by Lederman, but this was restricted by the capabilities of the IE 3000. Early informal testing of the device showed that textures with rectangular grooves caused the virtual pointer to become stuck in a groove and jump about. This meant that it was impossible to move the probe across the texture with a smooth action. It was therefore decided that sinusoidal grooves should be used, particularly since Lederman would have preferred to use that shape herself.

6.3.2.1.2 Simple objects

The objects used in this study were 6 cubes (three inside presentation, three outside presentation), 6 spheres (three inside presentation, three outside presentation), 3 rotated cubes and 3 sheared cubes (a total of 18 objects). Andrew Hardwick created the objects. The cubes and spheres were presented from both the inside and the outside. The rotated cubes were only presented from the outside, and the sheared cubes were only presented from the inside. The rotated and sheared cubes are depicted in Figure 6.4 below. Table 6.2 (below) presents the objects used in the study, which are indicated by ‘*’. 
Figure 6.4: Representations of the rotated and sheared cubes.

Table 6.2: The type and size / angle of the simple objects.

<table>
<thead>
<tr>
<th>Object Type</th>
<th>Size/Angle</th>
<th>Inside presentation</th>
<th>Outside presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cube</td>
<td>1.0 cm</td>
<td>*</td>
<td>See note 1</td>
</tr>
<tr>
<td></td>
<td>1.5 cm</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2.0 cm</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2.5 cm</td>
<td>See note 2</td>
<td>*</td>
</tr>
<tr>
<td>Sphere</td>
<td>1.0 cm</td>
<td>See note 3</td>
<td>See note 4</td>
</tr>
<tr>
<td></td>
<td>1.5 cm</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2.0 cm</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2.5 cm</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Rotated cube (2cm)</td>
<td>30°</td>
<td>See note 5</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>50°</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>70°</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Sheared cube (2cm)</td>
<td>18°</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>41°</td>
<td>*</td>
<td>See note 5</td>
</tr>
<tr>
<td></td>
<td>64°</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Key: ‘*’ = object presented.

Notes for Table 6.2: A full factorial design with presentation of all virtual objects of all sizes/rotations/shears was not used for the following reasons.
1. The outside presentation of a 1.0 cm cube was found that the object was found to be difficult to locate in the virtual space. In contrast, with the inside presentation of a 1.0 cm cube the probe is restricted within the object, which means there is no difficulty in locating the object.
2. The inside presentation of a 2.5 cm cube was found to be too large for the virtual space: the probe could not touch all the sides and corners of the object.
3. The inside presentation of a 1.0 cm sphere, like the 1.0 cm cube, was found to be too small: the probe could hardly be moved at all inside the object which meant that it could not be explored.
4. The outside presentation of a 1.0 cm sphere was found to be very difficult to locate in the virtual space.
5. It was decided that 18 objects were sufficient for this study and so rotated cubes were not presented from the inside and sheared cubes were not presented from the outside.
6.3.2.1.3 Complex objects

The demonstration of complex objects was exploratory in nature. It involved a selection of complex objects created by Andrew Hardwick. These included miniature versions of a sofa, an armchair, and a kitchen chair. The objects were constructed using simple cuboid objects butted together. For example, the sofa was made up of several cuboid objects that formed the sitting area, the back rest and the arm rests. (See Hardwick et al. (1998) for details.)

6.4 Procedure

At the beginning of the study participants were given a verbal overview of the aims of the study and were told about the tasks involved in the studies. They were shown the device and told briefly how it works. The blind participants were allowed to explore the device with their hands so that they were aware of its shape, size, and the position of the motors.

The participants performed three trials with the textures and then had a break. They then performed nine trials with the simple objects. After another break they performed the final three trials with the textures, and the final nine trials with the simple objects. The trials were alternated in this way to avoid participants getting bored with the repetitive nature of the trials involving the virtual textures.

6.4.1 Textures

Prior to starting the trials with the textures, the participants were presented with three examples to give them an idea of the types of textures involved. The magnitude estimation technique was explained to the participants and they performed a practice trial from which the data was not used. This allowed them to become familiar with the technique. In all trials the participants were allowed to choose their own modulus because Stevens & Harris (1962) found that 'observers' were more accurate in their assessments of roughness when they were allowed to choose their own modulus. Participants were advised not to use a modulus of zero, or of a low number because subsequent textures may be more or less rough than the standard.
Texture 5 (groove width 1.575mm) was used as the standard in all trials because it was from the middle of the range of textures. This was the texture to which the participants assigned the modulus. The system software presented the other nine textures in a random order, and in the same random order for each trial. The participants assigned magnitude estimations (MEs) to the other nine textures.

Each participant performed six trials with the set of ten textures, giving six measurements for each texture and a total of 60 measurements.

### 6.4.2 Simple objects

Before starting the trials the participants were presented with examples of the objects. It was explained that the device could present objects from the inside and the outside, and participants were given examples of this. The way in which it is possible to push through an object was also explained. The participants were told that they would be asked to judge the size of the cubes and spheres, and to judge the angle of rotation or shear of the rotated and sheared cubes.

Prior to the study the researcher arranged the order in which the objects were to be presented to the participants. The objects could have been presented in random order, but instead they were ordered so that the different types of objects were distributed as evenly as possible. The order of the objects is provided in Appendix 11.

A multiple choice matching technique was used to investigate participants’ perception of the size or angle of the virtual objects. When the participants were presented with each object, they were asked to explore it until they felt they had a good idea of its size or angle. No time limits were imposed.

When the participant said they were ready, they were given a representation of four objects and asked to indicate which of the four matched the virtual object in terms of size or angle. Sighted participants were shown scale drawings on a card and blind...
participants were shown scale 2-dimensional representations made from fuzzy felt\textsuperscript{62}, which they could feel. There was one card or fuzzy felt board for each type of object. Each card had one extra representation whose equivalent virtual object was not presented. (See Appendix 12 for examples of the representations given to the sighted participants – the representations given to the blind participants were tactile versions of the same cards.)

The difference in the formats of the response representations shown to the blind and sighted participants means that they were performing slightly different tasks. The sighted participants were cross-modal matching, i.e. they were matching a haptic stimulus to a visual response representation. In contrast, the blind participants were using same mode, i.e. matching a haptic (or tactile) stimulus to a haptic response representation.

6.4.3 Complex objects
During the demonstration of complex objects participants were asked to feel one or two of the complex virtual objects as time permitted. Informal testing with these objects suggested that it was difficult to identify the complex objects just by feeling them. Participants were therefore told what the object was before they felt it. They were asked to describe what they could feel and report whether they could identify the different components of the object. No formal data were collected during this demonstration.

6.4.4 Location of virtual space
After all the trials were completed 20 of the 22 participants were asked where they believed the virtual textures and objects were located in relation to the device. (The other two participants did not provide data on this subject due to time constraints.)

\textsuperscript{62}Fuzzy felt is a children’s toy. Pieces of felt can be cut to any shape or size and then placed on an A4-sized ‘fuzzy’ board. The felt sticks to the board until it is peeled off. In this case the felt was glued to the board in order to avoid accidental movement or removal.
6.5 Results

6.5.1 Textures

The data from the textures were analysed using logistic regression analysis to examine the relationship between the perception of the virtual texture and the physical parameters (groove width) of the texture. Regression analyses were conducted for each participant individually and on the massed data, which allowed a comparison of the performance of blind and sighted people. The results are presented in Table 6.3, below.

Table 6.3: Results of analysis of data for blind and sighted participants.

<table>
<thead>
<tr>
<th>Participants</th>
<th>Exponent</th>
<th>t-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blind participants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>-.457</td>
<td>-5.451</td>
<td>.0006</td>
</tr>
<tr>
<td>S8</td>
<td>-.306</td>
<td>-1.561</td>
<td>n.s.</td>
</tr>
<tr>
<td>S15</td>
<td>-.760</td>
<td>-3.323</td>
<td>.01</td>
</tr>
<tr>
<td>S16</td>
<td>-.768</td>
<td>-9.417</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>S17</td>
<td>+.747</td>
<td>+7.752</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>S18</td>
<td>-1.701</td>
<td>-17.094</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>S19</td>
<td>-.449</td>
<td>-2.845</td>
<td>.0217</td>
</tr>
<tr>
<td>S20</td>
<td>+.897</td>
<td>+27.120</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>S21</td>
<td>+.841</td>
<td>+8.349</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>S22</td>
<td>+1.027</td>
<td>+12.685</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>S32</td>
<td>+1.896</td>
<td>+27.120</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>S3</td>
<td>-.577</td>
<td>-1.626</td>
<td>n.s.</td>
</tr>
<tr>
<td>S4</td>
<td>-.766</td>
<td>-3.550</td>
<td>.0075</td>
</tr>
<tr>
<td>S5</td>
<td>+.394</td>
<td>+2.380</td>
<td>.0445</td>
</tr>
<tr>
<td>S6</td>
<td>-1.601</td>
<td>-5.858</td>
<td>.0004</td>
</tr>
<tr>
<td>S7</td>
<td>+.331</td>
<td>+1.332</td>
<td>n.s.</td>
</tr>
<tr>
<td>S8</td>
<td>-1.173</td>
<td>-5.967</td>
<td>.0003</td>
</tr>
<tr>
<td>S9</td>
<td>+.616</td>
<td>+5.577</td>
<td>.0005</td>
</tr>
<tr>
<td>S10</td>
<td>-.750</td>
<td>-3.456</td>
<td>.0086</td>
</tr>
<tr>
<td>S11</td>
<td>-.157</td>
<td>-1.554</td>
<td>n.s.</td>
</tr>
<tr>
<td>S12</td>
<td>-1.511</td>
<td>-16.789</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>S13</td>
<td>+.205</td>
<td>+6.10</td>
<td>n.s.</td>
</tr>
<tr>
<td>S14</td>
<td>+.079</td>
<td>+.878</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Key: n.s. = non-significant result.

264
6.5.1.1 Significant relationship between stimulus and perception

Overall, with the data from the blind and sighted participants combined, there was a highly significant relationship between the participants' perception of virtual texture and its simulated physical characteristics ($F_{(1,216)} = 12.09$, $p < 0.001$). Seven of the thirteen sighted participants and eight of the nine blind participants individually showed a significant relationship between perception of virtual texture and its simulated physical characteristics.

6.5.1.2 Difference between blind and sighted participants

The data were further analysed to compare the results from the blind and sighted participants. The actual means for the groups of blind and sighted participants were -.093 and -.362 respectively. A 2-tailed $t$-test found $t = .78$, df = 21, $p = .4469$. Therefore there was no significant difference between the two groups of participants. This lack of significant difference between the blind and sighted participants becomes less meaningful in the light of the fact the participants had negative and positive exponents as these different ways of perceiving textures need to be separated out.

6.5.1.3 Positive and negative exponents

As can be seen in Table 6.3 above, participants had both negative and positive exponents. A negative exponent means that the narrower grooves were perceived to be rougher than the wider grooves whereas a positive exponent means the wider grooves were perceived to be rougher. Of the eight blind participants who showed a significant relationship between the stimulus and their perception, half had a negative exponent and the other four had a positive exponent. For all the sighted participants who showed a significant relationship between the stimulus and their perception, the exponent was negative. Overall the majority of participants who showed a significant relationship between the physical textures and their perception also showed a negative exponent.
6.5.2 Simple objects

6.5.2.1 Cubes

For the virtual cubes, the sets of sizes of the cubes explored from the inside and the outside varied (see Table 6.2, above), so this factor could not be included in an analysis of variance. Instead, two separate analyses of variance were conducted, one for inside exploration and one for outside exploration, using visual status as a between groups variable and object size as a within groups variable.

The analysis of variance for exploration of the cubes from the inside found that there was no significant difference in the accuracy of the perception of the size of the virtual objects between blind and sighted participants ($F_{(1,20)} = 1.25, \text{n.s.}$) and no significant interactions involving visual status. Size did have a significant effect on the accuracy of perception ($F_{(2,40)} = 20.79, p < 0.001$). Figure 6.5 (below) shows that the effect of size on the accuracy of perception did show a particular trend. Cubes of all sizes were overestimated in size when explored from the inside. This can be thought of as a "Tardis" effect (see Colwell, Petrie, Kombrot, Hardwick, & Furner, 1998) as objects appear bigger from the inside than the outside. The effect was greatest for the smallest size object, but the effect was smallest for the medium size object (the overestimation for the largest object was of an intermediate value).

The analysis of variance for exploration of the cubes from the outside found no significant differences in the accuracy of the perception of the size of the objects due to either visual status or object size. As can be seen from Figure 6.5 below, this is because the accuracy of perception of the virtual cubes from outside was quite accurate, with only very small deviations from the calculated values.

6.5.2.2 Spheres

For the spheres, a full factorial analysis was possible using visual status as a between groups variable and exploration mode (from the inside versus from the outside) and virtual object size (1.5cm, 2.0cm and 2.5cm) as the within groups variables. The

---

83 The Tardis is a time travel machine in the popular British television series, Dr. Who. From the outside it is only the size of a telephone booth, but inside it is multi-roomed.
dependent variable used was the absolute deviation of the estimate of the size of the virtual object from its calculated size.

The analysis of variance found that there was no significant difference in the accuracy of the perception of the size of the virtual objects between blind and sighted participants ($F_{(1,20)} = 0.05, \text{n.s.}$) and no significant interactions involving visual status. However, both exploration mode ($F_{(1,20)} = 53.89, p < 0.000$) and virtual object size ($F_{(2,40)} = 6.72, p < 0.005$) had highly significant effects on accuracy of perception.

There was also a significant interaction between exploration mode and size ($F_{(2,40)} = 14.81, p < 0.000$). These effects are illustrated in Figure 6.6, below. The main effect for exploration mode is that exploration from inside the object leads to an overestimation of size, whereas exploration from outside leads to an underestimation of size (a further demonstration of the Tardis effect). However, this effect is moderated by size, being strongest for the smaller spheres. From the point of view of size, accuracy of perception is greatest for the largest sphere and decreases with size.

6.5.2.3 Rotated cubes

The analysis of variance for the estimation of the angle of rotation of a 2cm cube (explored from the outside only) found that there was no significant difference between sighted and blind participants ($F_{(1,20)} = 3.39, \text{n.s.}$). There was a significant effect for angle ($F_{(2,40)} = 34.22, p < 0.001$), which is illustrated in Figure 6.7 below. This shows that the smallest rotation (30 degrees) was somewhat overestimated, the middle angle (50 degrees) was quite accurately estimated and the largest angle (70 degrees) was very much underestimated.

6.5.2.4 Sheared cubes

The analysis of variance for the estimation of the angle of shear of a 2cm cube (explored from the inside only) found no significant difference between sighted and blind participants ($F_{(1,20)} = 0.42, \text{n.s.}$). There was a significant effect for angle ($F_{(2,40)} = 5.80, p < 0.01$), with the smallest angle (18 degrees) being slightly overestimated.
and the two larger angles (41 and 64 degrees) underestimated, as illustrated in Figure 6.8 below. There was also a significant interaction between angle and visual status ($F_{(2,40)} = 4.83, p < 0.01$). As can be seen in Figure 6.8 below, sighted participants were very accurate in their estimations of shear, whereas it was blind participants who overestimated the smallest angle and underestimated the two larger angles. However, this difference was not significant.

In terms of the overall accuracy of individual participants, one of the blind participants had an exceptionally high number of correct estimations of size across all the objects (15 out of 18 compared to a mean of 9 across both groups (median = 9)). A possible explanation of this high score is discussed in the discussion in Section 6.6.2.3 below.
Figure 6.5: Graph showing the estimation of the size of virtual cubes: inside and outside presentation.
Figure 6.6: Graph showing the estimation of size of virtual spheres: inside and outside presentation.
Figure 6.7: Graph showing estimation of the rotation of the virtual cubes.

Actual Rotation (degrees)

Estimated Rotation of 2cm Virtual Rotated Cubes

Estimated Rotation
Actual Rotation

0  10  20  30
0  10  20  30

0  10  20  30  40  50  60  70

0  10  20  30  40  50  60  70
Figure 6.8: Graph showing estimation of shear of the sheared cubes by blind and shielded participants.
6.5.3 Complex objects
During the demonstration of the complex objects the participants reported that they could make sense of most of the complex objects, and did so quite quickly. However, this was not true for all the complex objects. For example, the participants found that the kitchen chair was extremely difficult to make sense of.

Most participants said that although they could feel the shape of the components of the objects, such as the arm-rests and seats, they would not necessarily have been able to work out what they represented without being told.

6.5.4 Location of virtual space
Of the 20 participants who provided data on the location of the virtual space, 4 said that they imagined the objects to be located outside of the device, i.e. that the objects were touched by the end of the probe that they were holding, rather than the other end. Three of these 4 participants were blind. Fifteen participants thought that the objects were located inside the device and that the objects were touched by the other end of the probe. Of these 15, 6 were blind. One of the 20 participants imagined the objects to be located half way along the length of the probe, i.e. that the objects were not touched by either end of the probe, but by the middle of it. Analysis of this data using a Chi Square found that these differences were not significant (Chi square = 1.9, df = 1, n.s).

The participants differed in terms of which part of the probe they believed was touching the objects. Some believed that the objects were touched by one end of the probe (either the end they held or the other end) whereas others believed that as they moved along an edge of an object, the length of the probe was touching the object.

Participants also differed in the way they imagined touching the back of an object (outside presentation). Some imagined that they were feeling it with the end of the probe that they were holding. Others imagined that the probe was going through the object and
that there was a knob at the other end of the probe that prevented the probe from being
drawn through the object.

6.6 Discussion on the perception of virtual textures and objects
In this section an overview of the results of these studies will be presented. The findings
on the textures and objects will then be considered in turn, particularly in the light of
studies conducted by other researchers in this field. This is followed by some interesting
observations on the use of the IE3000, and on other aspects of the study.

The results presented above show there were no significant differences in the perception
of virtual textures and objects between blind and sighted participants. For the virtual
textures the majority of participants showed a significant relationship between the
simulated physical characteristics and their perception. Most of those that showed this
significant relationship also showed a negative exponent (meaning that the narrower
groove widths felt rougher than the wider grooves). For the simple objects an interesting
effect, known as the Tardis effect, was obtained for both cubes and spheres: objects
generally felt larger when they were explored from the inside than when they were
explored from the outside. For the rotated and sheared cubes there were varying levels of
under- and overestimation of angles.

6.6.1 Textures
Since the current study was conducted, other researchers have investigated the perception
of virtual textures, and real textures via a mediating device, such as a probe (see below).
These studies replicate the current results, and suggest possible explanations of the
differences between the current results and those for real textures in terms of the negative
exponents obtained.

Following on from Lederman’s previous research on the perception of real textures with
a bare finger, Klatzky & Lederman (1999) have recently investigated the perception of
real textures via 2mm and 4mm probes. Klatzky & Lederman found that with the 2mm
probe, the (log-log) data fitted a quadratic equation and the inverted U-shaped function
peaked at a certain interelement spacing and then declined. Therefore with a 2mm probe
the function switches from positive to negative.

Klatzky & Lederman conclude that when the interelement spacing is narrower than the
probe (i.e. less than 2mm) the texture feels more rough (a positive exponent). This is
because the probe cannot get into the space and it is only the ridges that are touched. In
contrast, when the interelement spacing is wider than the probe the texture feels less
rough (negative exponent). Klatzky & Lederman suggest this is because the probe can go
into the interelement space and therefore the underlying surface is touched as well as the
ridges. Since the underlying surface is smooth, the percept of the texture includes a
combination of both the underlying surface and the ridges. The texture is therefore
perceived to be less rough than when the percept does not include the underlying surface.

Wall & Harwin (2000) conducted a study to investigate the perceived roughness of
virtual textures via haptic perception. The textures were explored using the Phantom
device. The textures were of similar dimensions to those used in the current study. As in
the current study, and that conducted by Klatzky & Lederman, the majority of
participants showed a negative exponent (roughness increased with decreased groove
width).

The findings of Klatzky & Lederman and Wall & Harwin are in contrast with those of
earlier studies in which textures were explored with the finger. In these earlier studies a
positive exponent was obtained (wider groove widths felt more rough). Klatzky &
Lederman suggest that this difference between the exploration with a probe and a finger
is due to the fact that with small groove widths the fingertip cannot touch the bottom of
the valley of the groove, whereas a probe tip can touch the bottom.

The exact size of the IE3000’s probe point (used in the current study) is not known, but it
is thought to be ‘infinitely’ small, and therefore smaller than the width of the grooves of
the textures. It would therefore be expected that a negative exponent would be obtained
for the perceived roughness of virtual textures explored via a probe as it has been in the studies described above by Klatzky & Lederman and Wall & Harwin.

Penn et al. (2000; in press) have compared the current results with those of a study using the Phantom device. The virtual textures used by Penn et al. had the same dimensions as those in the current study. The experiments were repeated using two different endpoints fitted to the Phantom: a thimble and a stylus. The roughness of the textures was judged by blind and sighted people. As in the current study, and those conducted by Klatzky & Lederman, and Wall & Harwin, the majority of participants were found to have a negative exponent (i.e. the narrower groove widths were perceived to be more rough than the wider groove widths). Penn et al. also found that the magnitude of the negative exponent was greater with the thimble endpoint than with the stylus.

As in the current study, Penn et al. found no significant difference between the performances of the blind and sighted participants. However, a higher proportion of sighted participants showed a significant relationship between the physical properties and perceived roughness than in the current study. It is thought that a possible explanation for this is that the Phantom device is of superior quality to the Impulse Engine 3000, and does not produce as much vibration and noise, which may interfere with participants’ perceptions.

In the current study the majority of participants showed a negative exponent, although a minority showed a positive exponent. One possible explanation for this latter finding could be that the participants who had a positive exponent were moving the probe more quickly than the others were. However, Lederman has investigated the effect of finger speed (Lederman, Loomis, & Williams, 1982) and probe speed (Lederman, Klatzky, Hamilton, & Ramsay, 1999) and found that this has little or no effect on perceived roughness.

Another possible explanation is the participants thought the tip of the probe was the same size as the end they were holding, i.e. approximately 1cm in diameter. If these
participants thought the other end of the probe was this size this may have had an effect on their perception of the textures. This issue requires further investigation in order for it to be more fully understood.

The lack of difference found in this study between the perception of virtual texture by blind and sighted participants is in line with the expectations of Jansson (1998). In Jansson’s study, the magnitude estimation technique was used to assess the perception of the roughness of real and virtual textures by blind-folded sighted people. Jansson stated an intention to repeat the study with blind people, and expected the results to be the same as for the sighted people.

Jansson found that overall, virtual textures were perceived to be rougher than real textures (although not significantly so). It is possible that this difference between the two sets of results is due to the sighted participants in the current study not being blindfolded, as Jansson’s participants were. The researcher believes that if the sighted participants in the current study were blind-folded they may have been even less able to distinguish between the textures. The sighted participants were observed to watch the probe as they moved it over the textures. They may have been using this visual information with the haptic information to judge the roughness of the textures. It will be interesting to compare the results of the current study with those of Jansson’s intended future study.

6.6.2 Simple objects

This section presents observations on the results for the simple objects, and compares the results to past and recent research on the perception of virtual objects.

The results of the current study that larger objects were judged more accurately than smaller object are similar to those of Jansson (1998) (reviewed in Chapter 5, Section 5.7).

\[\text{Paul Penn of the University of Hertfordshire, is currently investigating the role of the perceived size of the probe tip in roughness perception of virtual textures, but no data is available yet.}\]
Jansson also found that larger objects were mostly identified more accurately than smaller objects.

6.6.2.1 Rotated cubes

A possible explanation for the fact that the 30° and 70° cubes were perceived to be within 10° of each other is that these two objects were confused with each other. As can be seen in Figure 6.9 below (and Appendix 12) these objects are almost mirror images of each other. The researcher believes that the participants may not have been able to discriminate between the angles of these two objects.

![Figure 6.9: Squares rotated anti-clockwise by a) 30° and b) 70°.](image)

6.6.2.2 Sheared cubes

Participants were observed to have some difficulties in exploring these objects. The cube with the largest degree of shear (64°) had two associated difficulties (see Appendix 12). Firstly, the angled sides did not feel smooth, but felt jagged. As the probe moved along the sides it made a different noise from that made when moving along smooth surfaces. Secondly, it was often difficult to perceive the corners of the object because when the probe moved from one side of the sheared cube to an adjacent side, it did not go right into the corner, but cut across it. However, these difficulties did not seem to prevent the participants from accurately judging these objects. A possible explanation for the overall
accuracy in the estimations of the sheared cubes is that the different cubes were so
different from each other they were unlikely to be confused with each other.

6.6.2.3 Contribution of experience to perception

One participant had an exceptionally high number of correct matches. In discussion with
this participant he stated that he had worked as a piano tuner for approximately 10 years
(about 7 years previously). This work involved using a screwdriver to tighten screws
within a cavity in the piano where it would not have been visible to a sighted person. The
screwdriver had to be used to guide the screw into position, then to keep it in position,
and feel when it was home. This experience may have been a factor of his high score.
Another possible factor is that this participant lost his sight aged 8 years, which means he
has some experience and memory of sight, for example he said he knows what a tree
looks like. The combination of this experience of sight and of piano-tuning may have
contributed to this participant’s high number of correct matches.

6.6.2.4 Recent research on perception of virtual objects

Since the current study investigating the perceived size of virtual objects was conducted,
another researcher, Bruns (1998) has investigated the perception of the size of virtual
objects and the effect of elasticity (hardness / softness) on the perception of size. Bruns’
study was conducted to replicate aspects of the current study, using the same device, the
IE 3000.

The objects used were cubes and spheres. The sizes of the cubes and spheres were 2, 3, 4
and 5 cms (slightly larger than those used in the current study). The objects were
presented from the inside and from the outside of the object. The elasticity of the objects
was measured in Newtons, a measure of surface resistance. Each size of object was
presented with four different levels of elasticity, 5, 6, 7 and 8 Newtons. Participants used
the magnitude estimation technique to assess the elasticity of each object. In order to
estimate the size of an object participants used a ruler with an occluding sleeve. The
sleeve was moved along the ruler to indicate the perceived size of each object.
The results of Bruns' study are similar to those of this study in a number of ways. Bruns also found no difference in the perception of objects between blind and sighted participants. Bruns also found that all objects were underestimated in size, and that larger objects were more underestimated than smaller objects. Furthermore, Bruns observed the Tardis effect: objects in the outside presentation were perceived to be smaller than objects of the same size in the inside presentation.

In the perception of elasticity the magnitude estimations did reflect the differences in elasticity. There was no difference between the blind and sighted participants. Also, there was no effect of elasticity on the perceived size of the objects.

As described in above in Section 6.2.1.4, Hardwick, Furner, & Rush (1997) observed that people using the IE 3000 differed in terms of where they imagined the virtual objects to be located in relation to the device. Bruns did not ask the participants where they perceived the objects to be located in relation to the device, either inside the device or outside. It would have been useful to compare the responses with those in the current study.

### 6.6.3 Complex objects

The complex objects demonstrated to the participants were made of simple component objects butted together with a very small space between the components. A problem associated with this method of creating objects is that the probe pointer can slip into the space and become temporarily stuck. This meant that the participants had to extract the probe from the space before continuing to explore the object. This may have made it more difficult to explore the objects effectively.

### 6.6.4 Observations on the use of the device

The way in which a user of the IE 3000 can explore virtual objects differs in several ways from the way in which real objects are felt. The device requires the user to feel textures
and objects with the probe. This is not necessarily an intuitive way of interacting with objects. It is possible to push the probe through a virtual object, but this is not usually possible with real objects. In order to feel the inside of a real object, an 'entrance' is needed, for example a door into a room. However, to feel inside a virtual object such as a cube, no entrance is required.

Despite these differences between interacting with real and virtual textures and objects, none of the participants reported specific difficulties with using the probe, pushing through objects, or feeling the insides of objects.

6.6.4.1 Control of the probe

When feeling virtual objects, many participants were observed to get temporarily 'lost' in the virtual space. This seemed to occur in three different circumstances: when feeling a new type of object; when searching for an object; and when searching for and feeling a very small object. When feeling an object for the first time, participants were observed to have difficulty keeping the probe in contact with the object. As the participants moved the probe along one side of an object, and then to the end, the probe tended to slip off into the 'virtual space'. The participants then had to explore this space with the probe in order to re-gain contact with the object and feel the next side. These difficulties seemed to result in the participants losing track of where they were in relation to the object, i.e. they had no reference points to use as navigational aids. This made the initial recognition of objects quite difficult, but most participants found that after a few minutes' practice they could trace around an object quite easily, maintaining contact with the object. Figure 6.10 (below) illustrates how participants developed control over the probe. In a) the probe slips off the virtual object at the end of each side. In b) the probe remains in contact with the object while moving around it. These initial difficulties were also observed when participants were exploring a type of object for the first time. They had difficulty maintaining contact with an object before they had built up a 'mental image' of it.
6.6.5 Improvements to the method

Since conducting these studies the researcher has identified several ways the method used could be improved in any future studies. It is acknowledged that the use of fuzzy felt did not allow the presentation of precisely accurate comparisons. Also, it would have been preferable to enable participants to specify the size or angle they perceive an object to be, instead of forcing a choice between several comparisons. This would have produced a more accurate picture of the perceived size or angle of objects. The occluding sleeve used by Bruns (1998) to enable participants to indicate the perceived size of virtual objects appears to have been a better technique than that used by the researcher. It allowed participants to customise their responses, rather than having to choose from a selection.

One solution to the problem of locating very small objects in the virtual space is to use the technique used by Jansson (1998) (reviewed in Chapter 5, Section 5.7) who placed objects within cubes with dimensions twice those of the stimuli in order to aid the participant in locating the stimuli. This technique could be used in future studies.

6.7 Conclusions on perception of virtual textures and objects

The two studies presented in this chapter investigated the perception of virtual textures and objects by blind and sighted participants. Participants were asked to judge the roughness of virtual textures and the size or angle of simple virtual objects. Some objects were presented in both outside and inside presentations, other objects were presented in just one presentation. The study of virtual textures extended previous research on the
perception of real and virtual textures. It was the first known study of virtual textures involving blind people. The study of the perceived size and angle of simple virtual objects was the first known to involve blind or sighted people.

In general, participants were able to discriminate between the textures, and there was no difference in perception between the two groups. An interesting finding was that the majority of participants who showed a significant relationship between the textures and their perception of them, also showed a negative exponent. This result has also been obtained in more recent studies involving textures (real and virtual) that were explored by a mediating device, such as a probe. These results are in contrast to the large number of studies conducted by Stevens, Lederman and their colleagues who obtained a positive exponent when real textures were explored with the finger.

It has been suggested that the reason for this difference in exponents for textures with small groove widths is due to the fact that while a fingertip cannot reach the bottom of the valley of the groove, the point of a probe which is smaller than the groove width can reach the bottom of the valley. While the main determinant of the perceived roughness of real and virtual textures is groove width, another factor is the relationship between the size of the groove width and the finger or probe that is used to feel the texture. If the groove width is larger than the finger or the probe, the texture will feel less rough than if the groove width is smaller than the finger or the probe. With the type of textures used in the literature the probe is generally smaller than the groove width and the finger larger than the groove width.

With the simple virtual objects there was, in general, no difference between the perception of the two groups. The 'Tardis' effect was observed in relation to the cubes and spheres. Previous testing with the 1E3000 had found that people differed in terms of where they perceived the virtual space to be located in relation to the device. This was also found to be the case in the current studies. Some participants thought the virtual textures and objects were located inside the device, while others thought these artefacts were outside the device.
The nature of the device has been found to introduce certain limitations in its use. In particular the display of smaller or larger objects than currently used, the display of objects with angled sides, and butting objects together. Objects that were smaller than 1cm were very difficult to find in the virtual space. The restrictions on the movement of the probe meant that objects larger than 4.5 cm could not be displayed. Objects with angled sides, such as the sheared cubes used in this study, caused difficulties because these sides did not feel as smooth as vertical or horizontal sides. As the probe was moved along these sides the device made more noise than when on smooth sides. The small gap between the components of complex objects trapped the virtual pointer, which may have interfered with building up a mental image of the object.

6.7.1 Design issues

These studies have illustrated the potential of using haptic technology to simulate real world objects or to create totally new virtual objects. They have also identified some of the problems with the current haptic technology. In designing haptic interfaces, designers need to exercise care and not assume that the virtual world will be perceived in exactly the same ways as the real world, particularly given the current limitations of haptic devices which use probes and joysticks. However, the current devices do provide realistic feeling textures and objects, which replicate the psychophysical properties of real textures and can be judged like real objects. Such devices have enormous potential for enhancing VEs for both blind and sighted people.

The researcher and colleagues published a paper on the current studies (Colwell et al., 1998). The paper concluded with a list of issues regarding the design of haptic virtual environments. These issues are summarised below.

6.7.1.1 Virtual Textures

- It cannot be assumed that physical variations in roughness of virtual textures can be easily detected or discriminated from one another, i.e. virtual textures may not be perceived in the same way as their real counterparts.
Users may vary in their perception of virtual texture in terms of the size of the differences, which they can detect. Users may also vary in their perception of virtual textures in terms of what feels rougher and what feels smoother.

6.7.1.2 Virtual Objects

- Users may perceive the sizes of larger virtual objects more accurately than the sizes of smaller virtual objects.
- Users may feel virtual objects to be bigger from the inside and smaller from the outside (the 'Tardis' effect).
- If it is important for users to perceive size accurately, virtual objects may need to deviate from their real world dimensions in the virtual world.

6.7.1.3 Interaction with the interface

- Users may not understand complex objects from purely haptic information; multimedia information may be required to give a sense of complex objects and what they mean.
- Users may have different mental models of where the virtual space is located.
- Users' mental models may vary in relation to what part of the device is "touching" a virtual object.
- Users do not appear to be disturbed greatly by being able to push through the surfaces of objects, but care should be taken to ensure that the users understand what is happening.
- Most users quickly learn strategies on how to explore virtual objects with a particular device, although some may find it useful if these are provided.

6.7.2 Virtual Environments on the World Wide Web

A logical extension to the scenario for virtual environments from Goralski, Poli & Vogel (1997) (described in Chapter 5) would be that haptic devices like the IE 3000 could extend the graphical and textual information available and allow users to touch and feel the Pyramids. Although the findings of the study reported here suggest that, in principle, it would be possible to create virtual environments that would make sense through touch, the researcher believes the hardware and the software need much improvement before such a scenario could be implemented.
6.7.3 Haptic access to the Web for visually impaired people

In the conclusion to Chapter 4 the researcher suggested that the adaptation of the content of Web pages is not sufficient for improving Web accessibility for visually impaired people, and that haptic devices, such as the IE 3000, could provide a complementary approach to improving access. In particular the potential for providing access to the visual layout of pages, and perhaps removing the need for alternative versions of pages was highlighted. In the current studies and those conducted by others it has been found that haptic devices can display virtual artefacts, such as textures and objects, and that people can recognise these artefacts and perceive them at their approximate size. The next stage of the development of such devices has already begun with the development of force-feedback mice (as described in Chapter 1, Section 1.3.4.3) which is likely to continue in the future.

6.7.4 Possible future work

The next logical step in the development of haptic devices to improve access for visually impaired people to the Web is to investigate the presentation of Web pages haptically, although this is perhaps a rather large step. In order for a visually impaired person to navigate a frameset or a table using a haptic device, other types of feedback would also be required, such as speech output, to provide background and orientation information. Further investigation of how these different sources of information would be presented by the system and managed by the user are required and should involve evaluations with potential users to ensure their requirements are considered.

Research is also continuing with the fundamental levels of perception via haptic devices. Since the current studies were conducted other researchers have conducted studies using both the IE3000 and the Phantom. The findings have generally been similar to those of the current studies. The continuation of this work indicates that a systematic investigation, like that involving real textures and objects, is underway for virtual textures and objects.
In future studies it would be interesting to observe the movements made with the probe while objects are being explored, and investigate whether they are similar to those found by Zinchenko & Lomov (1960). (Presented in Chapter 5, Section 5.4.8, Zinchenko & Lomov observed the micro- and macro-motions of the hand during the acquisition of information about an object.) This would provide further insight into how people explore virtual objects.
CHAPTER 7 – CONCLUSIONS TO THE THESIS

7.1 Introduction

This thesis has investigated the complex issue of Web accessibility for visually impaired people. This chapter provides an overview of the studies reported in the thesis and the literature reviewed. The chapter also presents the researcher’s conclusions on the definition of accessibility and how access to the Web may be further improved for visually impaired people.

Web access for visually impaired people has improved dramatically in the last few years. This improvement has come from the incorporation of two different approaches: the design of content and the presentation of content. The approach to the design of content involves the development of HTML tags that provide better access to the content, and the inclusion of such tags in Web pages. The approach to the presentation of content involves the further development of access technology and novel interfaces such as those provided by haptic devices.

With regard to the design of content, Web accessibility has a higher profile than ever, with large organisations such as the World Wide Web Consortium\(^5\) and the HTML Writers Guild\(^6\) actively promoting accessibility. Also organisations such as the

---

Royal National Institute for the Blind in the UK, and similar organisations in other countries, are campaigning for Web accessibility in several ways (for example the RNIB’s Campaign for good web design87).

With regard to the presentation of content of Web pages, a wide range of Web browsers and screenreaders is now available, which can present content in a more useful way. Some devices are very sophisticated, although perhaps difficult to learn and use. Such access technology and other devices are constantly developing and will provide improved access to both computers and the Web.

The researcher’s conclusions on the approaches to both the design of content and the presentation of content are presented below.

7.2 Design of Content

Although access technology is rapidly developing, and Web accessibility has a high profile, Web technology is also developing. This means that the ‘access lobby’ always has to ‘catch up’ with Web technology. The increased use of scripts that present animated graphics means that access technology has another hurdle to jump, and more retro-engineering needs to be done to make inaccessible sites accessible.

As suggested in the introduction to the thesis, there are several factors that provide motivation for organisations and companies to improve the accessibility of their Web sites, including legislation. Legislation has a key role to play in promoting accessibility: governments in the US, the UK and Australia are committed to making all government sites accessible. A recent case in Australia regarding the official Olympics Web site is an example of these commitments being tested in court88. A visually impaired person complained that they could not access much of the information on the site and took the Australian Olympics committee to court. The court found that the site was not accessible and ordered the committee to remedy the situation. However, the company who set up the site said they could not make the

necessary changes in time for the beginning of the Games. The complainant is now expected to seek damages for not being able to access the site during the Games.

Interestingly, the experts chosen to evaluate the accessibility of the site used the WCA Guidelines\(^8\) as a basis for their opinion. They also tested the site with the Bobby accessibility checker\(^9\) and made recommendations based on the results as well as on the Guidelines. The researcher believes that even if some of the recommendations are implemented, this will not guarantee that the site will truly be accessible to visually impaired people. As the study of the Guidelines reported in this thesis found, the implementation of some solutions recommended in the Guidelines, such as alternative text for images and Headers and IDs for tables, do not necessarily make these elements accessible.

The above is a good example of the difficulties in defining what is meant by ‘accessibility’. The Olympic Games Web site was found to be inaccessible by a visually impaired person and experts on Web technology and accessibility, and the Bobby checker. If the recommended changes are made, the site may be judged to be accessible in terms of the WCA Guidelines and Bobby, but will the site be accessible to the visually impaired person who complained?

In the researcher’s opinion, when defining accessibility we must include the experience of visually impaired people, and those with other disabilities: it is not sufficient to only consider compliance with a set of guidelines, or the output of an automatic checker. That is not to say that guidelines or checkers do not have an essential role to play in informing Web authors (and others) about accessibility issues – they do, but the most important issue is whether people with disabilities can access the information they wish to access.

7.2.1 Accessibility for visually impaired people

However, a limitation to this argument, as has been illustrated in this thesis, is that accessibility is not the same for all users. It has been found that users of different

---


browsers and screenreaders have different experiences of accessibility with the various HTML elements. For example, some participants were able to access the examples of the frames, whereas others could not. Also, participants with more experience of using the Web were able to access some elements, whereas those with less experience could not, for example those with less experience were less likely to be able to access the tables. So, regarding the Web site for the Olympic games, even if the particular person who complained finds the improved site to be accessible, users of other browsers or screenreaders may not be able to access it to the same degree.

In addition to the differences in experience of accessibility, the study found that the participants had different preferences for the way in which information was presented. For example, some participants liked text links to be presented in a list, whereas others did not like this because their screenreaders announced 'list item' for each item, which was felt to be too repetitive. Another example is that some participants liked the word 'reference' before the number that was a link to a reference, whereas others found this would also be too repetitive in a paper with many references. Therefore any definition of accessibility needs to acknowledge that individuals may have different preferences.

7.2.2 Use of guidelines
In addition to the issue of whether people with disabilities can access Web pages designed based on the recommendations in the WCA Guidelines, is the issue of whether people using the Guidelines can understand and implement the recommendations. The study reported in this thesis which examined the use of the WCA Guidelines found that participants had difficulties with the structure, and with the content of the documents. For example, they had difficulty finding the guidelines that were relevant to their task. This was mainly due to the organisation of the documents containing the guidelines and techniques. The participants also had difficulty in interpreting the guidelines and applying them in their task. This was mainly because the documents referred to technology, such as screenreaders, that the participants were not familiar with, and because the documents did not sufficiently describe the methods that people with disabilities use to access the Web.
Such difficulties with the use of guidelines are not specific to accessibility guidelines: similar difficulties have been identified by other researchers who have investigated the use of guidelines for the design of user interfaces. In this literature there are several common findings. For example, it was found that guidelines were not necessarily read, or used, in the way the authors intended, and different types of users of guidelines had different requirements of the document. In particular, the importance of the role of examples has been emphasised: users of guidelines may follow examples to the exclusion of the text (which the examples are intended to illustrate). There were also some common conclusions in the literature on how the use of guidelines can be improved. A common suggestion was that tools should be developed to overcome some of the difficulties that have been identified. However, only one such tool has so far been developed, and this has not been evaluated with potential users.

Another suggestion from the literature was that guidelines could be presented in a hypertext system in order to overcome difficulties with identifying guidelines. However, the current study involving the WCA Guidelines indicates that the use of hypertext does not necessarily overcome such difficulties, and such a complex hypertext system may even introduce its own inherent difficulties.

A further issue identified by the study of the use of the WCA Guidelines is the fact that the guidelines recommend the use of tags that are not yet supported by current browsers or by assistive technology. This issue was explicitly raised by some of the page author participants: they were frustrated that they had spent time implementing tags which they later found could not be tested in their browsers. They were also frustrated because the implementation of the tags had not solved the accessibility problems they were aiming to solve because the technology used by visually impaired people did not support these tags. This issue was implicitly raised when the pages were tested by visually impaired people: none of these participants mentioned these tags, let alone finding that they improved accessibility, and the researcher assumes this is because none of the participants were aware of the presence of the new tags.
Studies found in the literature conclude with several recommendations that are relevant to the future development of the WCA Guidelines. For example, the consideration that users of guidelines are unlikely to read such a document in full, but will dip into it to find the information they require for the immediate task (Tetzlaff & Schwartz, 1991). Also, the consideration that designers may find it difficult to understand and implement concepts that are new to them without sufficient explanation (Tetzlaff & Schwartz, 1991).

7.2.3 Conclusions on the design of content
The difficulties identified above with the use of accessibility guidelines and access for visually impaired people to Web pages lead the researcher to conclude that it is essential that assistive technology, browsers and new interfaces, such as haptic and force-feedback devices, are further developed and evaluated in order to improve Web accessibility through the presentation of content.

7.3 Presentation of Content
Studies involving haptic interfaces, such as the studies reported in this thesis, and work reported in the literature, represent advances toward improved presentation of content. This section summarises the literature on this topic and the studies presented in this thesis.

7.3.1 Touch perception
The literature on touch perception and the use of touch in computer interfaces was reviewed in Chapter 5, and more recent research was reviewed in the Discussion of Chapter 6. Three different types of touch were identified: tactile (stimulation of the skin), kinaesthetic (perception of the movement and positions of the muscles and joints), and haptic perception (perception of the combination of the tactile and kinaesthetic senses). Tactile and kinaesthetic perception are rarely used in isolation, for example, tactile perception may be used to feel the surface of an object, and kinaesthetic perception may be used to measure the length of a rod held between finger and thumb. However, most of our touch perception is haptic perception. For example, when we hold a cup of tea, we can feel that our fingers are in contact with the cup, and the temperature of the cup, and we are aware of the how far the cup is
from our mouth, and whether we are holding the cup upright. When we consider common tasks, they mostly involve some kind of touch perception, and this perception is mostly haptic perception.

In the investigation of touch perception, researchers have made several distinctions while trying to understand and explain the nature of this perception. For example, Gibson (1962) made the early distinction between active and passive touch, where active touch involves movement of the 'observer', and passive touch involves the 'observer' being touched. Loomis & Lederman (1986) extended Gibson's ideas by differentiating between whether or not the 'observer' has control over the touching. Also the distinction between individual touch sensitivities (tactile, kinaesthetic, haptic) (Loomis & Lederman, 1986) or separate sensory modes (Millar, 1981), and a whole perceptual system (Loomis & Lederman, 1986), or a unity of modes (Millar, 1981). Schiff (1980) makes a similar distinction that is related to the task involved: between the study of touch sensitivities in the study of psychophysics, and the functional analyses of touching behaviour.

7.3.2 Texture perception

With regard to the perception of textures there is a substantial body of literature. For real textures, Stevens (1962), and later Lederman (1974; 1981; Lederman & Taylor, 1972) found that a Power Law holds for roughness perception. In brief, the power law states that the perceived roughness of texture is directly related to the actual roughness of the texture. The main determinant of roughness of real textures explored with the finger has been found to be groove width: perceived roughness increases with groove width, and therefore has a positive exponent (Lederman 1974; 1981; Lederman & Taylor, 1972).

With the advent of haptic computer interfaces, the study of texture perception has been extended to include virtual textures. Minsky (1995) and Jansson (1998) have found that the power also holds for virtual textures, although there is greater variation in the exponents than with real textures. The effect of the use of a mediating device, such as a probe, on the perception of texture has been investigated, for both real (Klatzky & Lederman, 1999) and virtual textures (Wall & Harwin,
2000). It has been found that textures explored with a probe have a negative exponent, i.e. perceived roughness increases with decreased groove width. This is in contrast with textures felt with the finger. It has been concluded that the determining factor is the relationship between the size of the finger or probe in relation to the groove width (Klatzky & Lederman, 1999).

For the (real and virtual) textures explored with a probe in the literature, the probe was always narrower than the groove width. Therefore the probe touched the bottom of the groove valleys, which was the underlying surface. This meant that the texture felt smoother overall because the percept included the underlying surface as well as the ridges. In contrast, for the textures explored with the finger in the literature, the groove widths were always narrower than the finger. Therefore the finger did not touch the bottom of the groove valleys and only touched the grooves. This meant that the texture felt smoother overall because the percept only included the ridges, and not the underlying surface.

The study involving virtual textures found there was no difference in the perception by blind and sighted people. Perceived roughness was found to be related to the physical properties of the textures for the majority of participants. For most of these participants the exponent was negative. This was in contrast to the results obtained by Lederman with real textures that indicated a positive exponent (1974; 1981; Lederman & Taylor, 1972). However, the results were in common with those obtained by other studies of real textures explored with a probe (Klatzky & Lederman, 1999) and virtual textures explored with a probe (Wall & Harwin, 2000).

7.3.3 Object perception

With regard to the perception of objects, the literature is less conclusive than that on the perception of texture because there have been fewer replications of experiments by researchers. This means that there are many papers reporting examples of different tasks, which involve various objects, but few definite conclusions. This situation is also reflected in the small amount of literature on the perception of virtual objects. The existing literature on virtual object has found that sighted people have
accuracy rates of 70-100% in identifying virtual objects (depending on the particular object) (Jansson, 1998).

The study involving virtual objects observed an interesting effect, known as the Tardis effect. The cubes and spheres generally felt larger when they were explored from the inside than when they were explored from the outside. For the rotated and sheared cubes there were varying levels of under- and overestimation of angles. Again there were no differences in perception between the blind and sighted people.

More recent studies on the perception of virtual objects have had similar results to those of the current study. For example, as in this study, Bruns (1998) found no difference in the perception of objects between blind and sighted participants, and that the size of objects was underestimated by all participants. Bruns also replicated the 'Tardis effect' in that objects in the outside presentation were perceived to be smaller than objects of the same size in the inside presentation.

7.3.4 Conclusions on the presentation of content
The main conclusion the researcher draws from the studies reported in Chapter 6, and those in the literature, is that haptic computer interfaces, such as the Impulse Engine (and the Phantom – used in other studies) can simulate textures and objects. This simulation is sufficient for blind and sighted people to accurately perceive the size (Bruns, 1998; Colwell, Petrie, Kornbrot, Hardwick, & Furner, 1998), angle (Colwell et al., 1998), and elasticity (Bruns, 1998) of objects, and for sighted people to recognise objects (Jansson, 1998). Furthermore for the roughness of textures to be judged accurately by blind people (Colwell et al., 1998) and sighted people (Colwell et al., 1998) (Jansson, 1998). Future work will need to investigate the differences in perception of real and virtual artefacts further, as being currently done by Penn et al. (2000; in press).

The study involving the haptic device was the first formal study involving the IE3000, extending the exploratory work conducted by Hardwick et al. using the same device (Hardwick, Furner, & Rush, 1998; Hardwick, Rush, Furner, & Seton, 1996). It was also the first to investigate the perception of virtual objects, and to investigate perception by blind people of virtual textures and objects. This study identified for
the first time the Tardis effect, which has been replicated in another studies since. This study also confirmed the negative exponent in the perception of virtual textures.

7.4 Researcher’s final conclusions
It has been seen that accessibility guidelines for the design of content are not sufficient for improving Web accessibility for visually impaired people because there are other factors involved, including the way that browsers and screenreaders present the content. In order to improve this presentation of content new interfaces, such as haptic devices, are being developed. The studies involving the IE3000 lead the researcher to believe that haptic computer interfaces do have potential for improving access to the Web for visually impaired people. In particular, such interfaces could improve access to information that is visual in nature, such as text presented in columns or frames, and data presented in tables. In order to achieve this the work on the perception of virtual objects and textures needs to be repeated using other haptic interfaces and a wider range of artefacts. This is already underway in several universities, including the University of Hertfordshire and will pave the way for the development of devices that allow the navigation of Web pages containing elements such as frames and tables.

Regarding the design of content there are also several areas that require further investigation. For example, the HTML elements recommended in the Guidelines that could not be evaluated in the current study, because they are not yet supported by browsers, need to be evaluated with visually impaired people to ensure that these elements do improve accessibility. Also, all the elements recommended by the Guidelines need to be evaluated with people with other disabilities.

7.5 Contribution to knowledge
This thesis makes important contributions to the bodies of knowledge regarding accessibility of the Web and the haptic perception of virtual textures and objects. Regarding the accessibility of the Web, this was the first study to involve the WCA Guidelines. This study confirmed the difficulties reported in the literature regarding the use of guidelines in user interface design. It found that the hypertext format recommended as a solution to some of the recorded difficulties with guidelines has
its own inherent problems. This study was the first to examine the experience of visually impaired people in using the Web. It identified the need to improve access to alternatives provided by Web page authors, for example, alternative text for images, and NoFrames versions of frame sets. The study also identified the need to improve access to the layout of information, for example in tables, frames. It has been concluded that the haptic interface examined in this thesis represents a type of future interface that could provide the improved access to the Web that is required by visually impaired people.
8 REFERENCES


9 WEB ADDRESSES
Below are the URLs cited in this thesis. All were checked Dec 2000.

9.1 Web Accessibility Initiative

Web Accessibility Initiative home page, available at: http://www.w3.org/WAI/

Web Content Accessibility Guidelines, available at: http://www.w3.org/TR/WAI-WEBCONTENT/

WAI Authoring Tool Working Group home page at: http://www.w3.org/WAI/AT


User Agent Working Group home page, available at: http://www.w3.org/WAI/UA/


Education and Outreach Working Group home page, available at: http://www.w3.org/WAI/OE/

WAI Quicktips card, available at: http://www.w3.org/WAI/References/QuickTips/

Techniques for Web Content Accessibility Guidelines 1.0, available at: http://www.w3.org/TR/WAI-WEBCONTENT-TECHS/

9.2 Accessibility guidelines

RNIB’s Hints for designing accessible Websites, available at:
http://www.rnib.org.uk/digital/hints.htm

AUS Standards, available at:
http://www.lawlink.nsw.gov.au/4a2565da0002e87d/1fb2c8a6db60f3d14a2565f5000554ad/2fcd6b983169f487ca2568640007a952?OpenDocument

HARMONY Guidelines for the Creation of Accessible WWW Documents, available at:

Stella O’Brien’s ‘Ten hazards in accessible Web design’, available at:
http://www.dcaip.org/dcweb/pwd/art-sobl.html

9.3 Accessibility checking

Webpage Accessibility Self-Evaluation Test (version 2), available at:
http://www.psc-cfp.gc.ca/cepmppmpee/access/testver1.htm

9.4 Screenreaders

ASA: http://www.microtalk.com/
Hal: http://www.dolphinuk.co.uk/
Jaws: http://www.hj.com/
Outspoken: http://www.alva-bv.nl/alvacorp/alva_corp_home.html
Simply Talker: http://www.econointl.com/index.htm
Slimware Window Bridge: http://www.synthavoice.on.ca/
Window-Eyes: http://www.gwmicro.com/
WinVision: http://www.artictech.com/

9.5 Browsers

Lynx: http://lynx.browser.org
Internet Explorer 4.01 accessibility features:
http://www.microsoft.com/enable/products/IE.htm
Net-Tamer: http://www.nettamer.net/
Netscape Navigator: http://home.netscape.com/
9.6 Authoring tools

Details of DreamWeaver, available at:
http://www.macromedia.com/software/dreamweaver/


Details of HotMetalPro, available at: http://www.hotmetalpro.com/

9.7 Other Web accessibility resources

AWARE centre’s home page, available at: http://aware.hwg.org/

Americans with Disabilities Act home page:
http://www.usdoj.gov/crt/ada/adahom1.htm

Web Accessibility Symbol, available at:

Bobby accessibility checker, available at: http://www.cast.org/bobby

Webwatch discussion list archive:
http://lyris1.telelists.com/htbin/lyris.pl?enter=webwatch&text_mode=0

Betsie home page, available at: http://www.bbc.co.uk/education/betsie/about.html

RNIB Campaign for good web design home page, available at:
http://www.rnib.org.uk/digital/welcome.htm

9.8 Other Web resources


HTML 4.0 specification, available at: http://www.w3.org/TR/html4/

W3C Validator, available at: http://validator.w3.org/

9.9 Miscellaneous

iFeel mouse: product overview, available at:
   http://www.logitech.com/cf/products/productoverview.cfm/79
Details of Impulse Engine, available at:
http://www.immersion.com/impulseengine.html
10 APPENDICES

10.1 Appendix 1: Purpose of new HTML tags (Chapter 3)

HTML 4.0 enables page authors to associate data cells with the appropriate header cells. This is intended to improve navigation of tables by users of assistive technology, particularly screenreaders. Header cells can be allocated IDs and data cells, or other headers, can then be associated with the headers using the 'headers' tag.

HTML 4.0 also enables page authors to provide a Summary for a table. This is intended to provide users of assistive technology with additional information about the table, in case they cannot read the actual table. Graphical browsers do not display summaries.

The other element that can be used with tables is Caption. This was available in previous versions of HTML and is not directly related to accessibility. Caption is supported by most browsers (mainstream and special) and is 'visible' in graphical browsers.

The 'Noframes' tag can be used to present information to browsers which do not support frames. A page author can use this to provide the same information and functionality that is available in the frames. However, often the only information authors put in the Noframes version is about how to obtain a browser that does support frames.

Mailto links are email addresses which also form the anchors of links. Selecting the link can prompt an email package to open, with the address automatically inserted into a new message, which can be sent immediately.

'D' links are intended to provide access to a longer description that is not available in the alternative text. The text of the link is just a capital letter D.

D-links were first proposed by WGBH91. In HTML 4.0, D-links have been succeeded by the Longdesc attribute, which is defined in the HTML 4.0 specification92.

Tabindex is a tag suggested by the WAI WCA Guidelines for improving the accessibility of forms. It does not affect the visual appearance of the form.

91 See http://main.wgbh.org/wgbh/pages/ncam/accessncam.html
92 See http://www.w3.org/TR/Rec-html40/
10.2 Appendix 2: Questions sent to participants in the evaluation of the Guidelines (Chapter 3).

Please answer the following questions:

How many Web pages have you created?
How many of these pages are you currently maintaining?
Which of the following elements have you used in your Web pages?
Images, Imagemaps, Tables, Frames, Forms
Would you prefer to use an authoring tool during the experiment?
If so, which one?
Which browser(s) do you prefer to view your pages with?
Do you prefer to use a Mac or a PC?

Please provide a telephone number where you can be contacted.
Thank you.
10.3 Appendix 3: File given to participants working with tables and images (Chapter 3).

<!-- HTML code goes here -->

## Use of a haptic device by blind and sighted people

Chetz Colwell \sup{1}, Helen Petrie \sup{1}, Diana Kornbrot \sup{1}, Andrew Hardwick \sup{2} & Stephen Furner \sup{2}

\sup{1} Department of Psychology, University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK
\+44 1707 284629
{c.g.colwell | h.l.petrie | d.e.kornbrot}@herts.ac.uk

\sup{2} British Telecommunications plc, Research Laboratories, Martlesham Heath, Ipswich, 
\+44 1473 646846
{andrew.hardwick | stephen.furner}@bt-sys.bt.co.uk

## 1. Summary

## 2. Introduction

### 2.1 The Impulse Engine 3000

The device used in the current studies was the Impulse Engine 3000 (shown in [Figure 1](#Figure 1)). This force-feedback device was developed by the Immersion Corporation ([6](#[6])) and was used with software written by Andrew Hardwick (see [3](#[3])) for a comprehensive description. The system can display virtual textures and objects which the user can feel using a probe. The probe is the length and diameter of a thick pen and has 3 degrees of freedom of motion, i.e. it can move in 3 spatial dimensions: forwards and backwards, up and down, and left and right. The system provides force-feedback to the user by monitoring the position of their hand and altering the force accordingly ([5](#[5])). The force is created by three motors which exert resistance against the probe. This gives the user the impression that a texture or object is present.

![Figure 1](figure1.gif)
3. Study 1: Perception of Virtual Textures

The first study involved virtual textures with varying groove widths. The dimensions of the virtual textures were as close as possible to those used by Lederman &amp;#91;7&amp;#93;, the difference being that those textures involved grooves with a rectangular waveform whereas the textures used in the current study involved sinusoidal shaped grooves. The widths of the grooves varied from 0.375mm to 1.5mm in steps of 0.125mm and had a fixed amplitude of 0.0625mm (shown in Figure 2). There were no visual representations of the virtual textures. A magnitude estimation technique &amp;#91;11&amp;#93; was used to assess the roughness of ten textures with six trials per participant.

4. Study 2: Recognition of Virtual Objects

Table 1: Mean perceived size/angle of virtual objects with percent over- and under-estimation (data from sighted and blind participants combined)

<table>
<thead>
<tr>
<th>Object Type</th>
<th>Actual Size/Angle (cm/degrees)</th>
<th>Perceived Size/Angle Inside Presentation</th>
<th>Perceived Size/Angle Outside Presentation</th>
<th>Mean and standard deviation (cm/degrees)</th>
<th>Over/under estimation (Percent of actual)</th>
</tr>
</thead>
</table>

312
Mean and standard deviation (cm/degrees)

<table>
<thead>
<tr>
<th>Over/under estimation (Percent of actual)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cube</td>
</tr>
<tr>
<td>1.8 (0.40)</td>
</tr>
<tr>
<td>+80%</td>
</tr>
<tr>
<td>See [Note 1 below](#note 1 below)</td>
</tr>
<tr>
<td>-</td>
</tr>
<tr>
<td>1.5</td>
</tr>
<tr>
<td>1.7 (0.30)</td>
</tr>
<tr>
<td>+13%</td>
</tr>
<tr>
<td>1.6 (0.50)</td>
</tr>
<tr>
<td>+7%</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2.4 (0.20)</td>
</tr>
<tr>
<td>+20%</td>
</tr>
<tr>
<td>2.0 (0.50)</td>
</tr>
<tr>
<td>0%</td>
</tr>
<tr>
<td>2.5</td>
</tr>
<tr>
<td>See [Note 2 below](#note 2 below)</td>
</tr>
<tr>
<td>-7%</td>
</tr>
</tbody>
</table>

Notes: 1. Preliminary investigations showed that a 1 cm edge cube was too difficult for participants to find in the outside presentation, so it was omitted.
2. Preliminary investigations showed that a 2.5 cm edge cube was too big for the virtual space available.

5. Discussion
Figure 3: Different mental models of the location of the virtual object in virtual space: (a) outside the device; (b) inside the device.

6. Conclusions

7. References


Go to top of page
10.4 Appendix 4: Questions asked of page author participants (Chapter 3).

Date: Initials: GLs version used:
Text / online

You and your experience:

Course? Age?

Do you ever view your pages with an alternative browser to test how your pages are displayed?
   If so, which one(s)?

Have you ever asked other people to look at your pages to see if there are any problems with them?
   If so, was it useful? Why was it useful?

Do you ever use an authoring tool?
   Always / Sometimes / Never
   If so, which tool?

Using the Guidelines:

Which sections of the guidelines did you consult during this experiment?
<section and sub-section numbers>

Which sections of the Techniques did you consult during this experiment?
<section and sub-section numbers>

Did you find the distinction between the Priorities 1, 2, and 3 useful?
   For the guidelines?
   For the techniques?

Did you find the information you needed in the guidelines?
   Always / Sometimes / Never
   If not always, what couldn't you find?

Did you have any problems locating the information you required to implement the advice given in the guidelines?
   If not, what further information did you require?

Did you find that in following the advice contained in the guidelines you had to compromise the visual aspects of your pages?

Did you have any other problems using the guidelines?

If you used an authoring tool to implement the advice in the guidelines did you have any problems with this?
Overall, how useful did you find the guidelines?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not useful</td>
<td></td>
<td>Fairly useful</td>
<td></td>
<td>Very useful</td>
</tr>
</tbody>
</table>

Overall, how useful did you find the Techniques?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not useful</td>
<td></td>
<td>Fairly useful</td>
<td></td>
<td>Very useful</td>
</tr>
</tbody>
</table>

Can you suggest any improvements that could be made to the guidelines?

**Accessibility**

Do you think that the guidelines are a useful way of learning about accessibility?

Are there other ways in which you would prefer to learn about accessibility? For example,
- Book
- Website
- Course / lecture
- Workshop
- Other

Were you already aware of accessibility issues before participating in this study? If so, what were you aware of, and how did you find out about it?

Having participated in this study, how important do you think it is that your Web pages are accessible to people with disabilities?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not important</td>
<td></td>
<td>Fairly important</td>
<td></td>
<td>Very important</td>
</tr>
</tbody>
</table>

**Future:**

Do you think you would refer to the guidelines when designing pages in the future? If not, can you say why not?
10.5 Appendix 5: Discussion with WAI on results of study (Chapter 3)

Overview of the issues raised with the WCA Guidelines Working Group on the basis of the results of the study. (The reports on ‘invisible’ and unsupported tags were submitted in November 1998. The other reports were submitted in March 1999.)

10.5.1 Report regarding ‘invisible’ tags

The researcher reported that some participants were confused when they implemented tags but were not able to see or test the results. This occurred with tags that are new to HTML 4.0 that were regarded as ‘special’ such as Longdesc, Table summary, and Headers and IDs.

The researcher suggested that the Guidelines or Techniques documents could indicate that some tags might not currently be ‘visible’ in graphical user interface (GUI) browsers, but that they will be ‘visible’ to other browsers and assistive technology in the future.

10.5.1.1 Discussion on ‘invisible’ tags

None of the respondents acknowledged the difficulties experienced by the participants. One commented that some screenreaders currently use ‘Dynamic HTML’ and Microsoft Active Accessibility to present ‘non-visible’ tags to the user. It is not clear whether this respondent is referring to non-visible tags from previous versions of HTML, or tags from HTML 4.0. They did not indicate which screenreaders they were referring to.

Another respondent made the following comments.
• The Guidelines should indicate that many of the new tags, such as Table Summary and Longdesc, are better supported by ‘aural’ browsers than by graphical browsers.
• The Guidelines over-emphasise accessibility: there are no ‘accessibility features’ in HTML. For example, Longdesc is used by search engines to index ‘important’ images on a page.
• The Guidelines cannot say that a tag will not be visible to a page author because it cannot be assumed that they will be using a graphical browser. Also, some browsers already support Longdesc so to say ‘will be available’ would not be correct.

This respondent did not say which browsers already support Longdesc, nor did they make a suggestion as to how the issues concerning other tags should be addressed.

The researcher replied to the above comments by re-stating the difficulties that the participants had experienced concerning tags such as Longdesc and Accesskey. The researcher acknowledged that the term ‘visible’ was perhaps misleading and so the term ‘available’ was used to avoid the assumption of the use of a graphical browser. The researcher pointed out that although the participants could have used a validator to check their page, they did not.
10.5.2 Report regarding unsupported tags

The researcher reported that participants were confused when tags were recommended by the Guidelines but not yet supported by the browsers they were using. The researcher suggested that the Guidelines or Techniques documents could, at least, indicate the tags that are not yet supported by all browsers, or at best, list the browsers that do currently support the relevant tags.

10.5.2.1 Discussion on unsupported tags

One respondent commented that the issue of how ‘User Agents’ should support the new tags had been raised with the WAI User Agent working group, and that in particular, the tags that are beneficial to accessibility should be prioritised. They noted that IE 4.0 does support Accesskey, but only on form controls, not on anchors.

Another respondent made the following comments regarding lists of browsers.

- It would not be appropriate to list all the browsers that support each kind of feature. Partly because almost every feature is not supported by some browser, but also because it would not be possible to list all the existing/ emerging browsers.
- If only the major browsers were listed then page authors might not consider the other existing browsers.
- The Object tag is supported by browsers for TV devices and Accesskey is supported on some mobile devices.

Another respondent also commented on the idea of lists of browsers, and on new tags, and other aspects of the Guidelines.

- Browsers should not be listed, but it is important to provide warnings about ‘buggy’ features, such as Object, that can cause browsers to crash if they are not supported.
- HTML 4.0 should be the default version for the Guidelines. If the Guidelines state that some tags are new to HTML 4.0 there is a danger that page authors would not use the tags because only a few browsers support HTML 4.0.
- The Guidelines should emphasise that browsers should not be used as validators. Authors should be advised that before they check whether a browser renders a tag they should validate their pages, read the guidelines, and read the HTML specification.
- If the Guidelines are clear in their advice and provide ‘good’ examples, it will be clear from the Guidelines if they were implemented correctly.

The following reports were submitted in March 1999. The researcher submitted a brief description of the study, the participants involved and the procedure used.

10.5.3 Report regarding priorities

The researcher reported that one participant did not like the fact that some tags are priority 1, even though they are not yet supported by any browser. This means that a page author ‘must’ implement these tags if they are to achieve the highest level of conformance to the Guidelines. The researcher acknowledged that there was a sentence in the section ‘How the Guidelines are organised’, which notes that not all the tags recommended in the Guidelines are supported by all browsers.
"The current document and the Techniques document refer to some technologies (e.g. features of HTML, CSS, or SMIL) that may not be supported yet in some browsers, multimedia tools, etc." (WAI93).

The researcher suggested that such a statement could be used for all the relevant guidelines, not just in one section. Alternatively, the note mentioned above could be moved to a more prominent location, such as the introduction.

10.5.3.1 Discussion on priorities
The discussion on the priorities of unsupported tags covered much of the same ground as the general discussion on unsupported tags described above. However, one person responded with the following comments.

- If the priority 1 rating of this Checkpoint is lowered because browsers do not support the tag, the accessibility gains are lost. Although on the other hand authors do not like to use tags that do not work for them.
- It is a challenge to get authors to understand that the purpose of following these Guidelines is that features that may not be supported by their browser may be crucial to people who are using a different browser.

The researcher pointed out that she was not suggesting that the priority of the Checkpoint be lowered but that authors should be allowed to make informed choices regarding the implementation of tags. The researcher suggested that a paragraph similar to the following could be inserted, perhaps in the introduction to the Guidelines.

Page authors should note that some of these technologies may not be supported by all browsers. However, browsers that do not support such technologies will not be adversely affected by their use.

The following responses to the above suggestion were made by two people.

- The proposed wording is acceptable, but the utility of the Guidelines would be harmed by an explicit statement about what is supported because the situation changes too quickly.
- It is not yet known which browsers will support which tags, and it cannot be predicted whether such tags will be supported by main browsers or just special devices.

The researcher responded that this point could be made more explicit so that authors may be less surprised when recommended tags are not supported, and would perhaps understand that this issue is not predictable. The following response was made.

- Where possible the Guidelines say "until user agents..." to indicate that a tag is not yet supported but is expected to be so soon. (WC, 15 Mar 1999)

The researcher pointed out that the Checkpoint regarding Headers and IDs for tables does not use this phrase. The researcher asked what page authors should do until browsers support Headers and IDs, and that perhaps the previous technique of the use of paragraph breaks should be included here.
Two comments were made regarding the way that tables are presented.

- The issue of the use of paragraph breaks will be raised again with the working group.
- The use of paragraph breaks was intended to solve a 'legacy' problem when some browsers ignored all table mark-up and ran all the contents into one line. It would be surprising if people were still using such browsers.

The researcher did not respond to this latter comment at the time because it was hoped that the second of the series of studies, involving visually impaired users, would shed further light on this topic.

10.5.4 Report regarding new tags

The researcher reported that participants appeared to be less confident in implementing new tags such as Accesskey, Longdesc, and Headers and IDs. Also that participants seemed to require additional support in implementing these new tags. Some looked up details of the syntax of such tags in HTML reference books, but did not find any mention of these tags. The researcher suggested that further support be provided for implementing the tags. Participants said they would like examples of how the new tags might be rendered visually and in synthesised speech.

10.5.4.1 Discussion on new tags

The following responses were made.

- What kind of additional support should be provided for using new tags?
- The WAI User Agent group is also considering how such new tags will be used and the Guidelines group does not want to duplicate their efforts.

The researcher responded that it was assumed that it was known how the new tags would be used by browsers / user agents. The researcher further suggested that the Techniques document should indicate that some tags are new and so it is not known how they will be used.

The following comments were made regarding the rendering of the new tags.

- It is already known how some new tags, for example Longdesc, are to be used because they are already supported by the latest browsers, such as IE 5, or by special browsers.
- Blynx and Tiny are browsers that support Accesskey, but it is not author-specified; it is automatically generated.
- The HTML 4.0 specification purposefully does not discuss how new tags should be presented. The details of the semantics of Longdesc were postponed until they could be addressed by the WAI. It seems that this time has now come, but this working group is not the appropriate forum for these discussions. It should be discussed by a combination of working groups.
- It is the responsibility of the User Agent working group to decide how tags should be rendered. It is the responsibility of the WCA Guidelines working group to ensure that the necessary information is included in Web pages.
10.5.5 Report regarding testing / validation
The researcher reported that page authors had difficulties with using validators because they require certain information. The researcher suggested that the Techniques document could indicate which kind of information would be required. The researcher also suggested that further information be provided on the technique of viewing a page without images loaded when the image has already been viewed.

10.5.5.1 Discussion on testing / validation
The following comments were made by respondents.
- The Guidelines will soon require authors to comply with a formal grammar, such as DTD, and so further information on this topic will be provided in the Techniques document.
- The technique for viewing pages without images will be more specific.

10.5.6 Report regarding examples
The researcher reported that participants had difficulty in allocating names to variables when following examples. The researcher suggested that the variables in the examples be reviewed and changed where necessary. The problem of participants implementing the first example, rather than finding the most appropriate example was also reported. The researcher also suggested that the examples could be listed in a way that made it clear which example was to be used in which circumstances.

10.5.6.1 Discussion on examples
One of the editors responded by saying that the variables concerned would be re-named as 'header1', 'header2' etc. They also said that the idea of listing the examples was a good one and would be born in mind while the Techniques document was being edited. This idea had not been taken up at the time of writing.

10.5.7 Report regarding navigation
The problems that participants experienced with finding information initially and finding information again were described, with reference to the table of contents, the location and wording of links, and the complexity of the URLs. Several possible solutions to these problems were suggested. The table of contents could list HTML topics, rather than listing the individual guidelines; the URL could be made less complex; link text could be re-phrased and links located at the end of each guideline; and links with the same text should go to the same destination. Also the section numbering could be standardised across the Guidelines and Techniques documents and a site map could be provided.

10.5.7.1 Discussion on navigation
The response from member of the groups included several comments and questions. Each is addressed separately below.

One question was, did the participants who had difficulty using the table of contents to find the topic they were looking for, use the Checklist of Checkpoints which is organised by topic and links back to the Guidelines?
The researcher was surprised by the suggestion that the Checklist should be used as a table of contents. This seems to be a ‘workaround’ rather than a solution to the problem. The researcher replied to the respondent that only one of the participants was observed to methodically work through the Checklist as a table of contents. The others did not use the Checklist. Also the Checklist is presented on the Web site as an evaluation tool rather than a navigation aid.

The respondent also stated that others had also made the suggestion that links that go between the documents should indicate which document they point to and the working group is addressing this problem.

In more recent versions of the Guidelines the links that point to a different document indicate this fact with the terms ‘Guidelines’ and ‘Techniques’ in the link text.

The respondent also made the following responses.

- The URLs between the documents are relative and so should be no longer than a few words.
- The fact that the Guidelines and the Techniques documents are organised differently (the Guidelines by issue, the Techniques by issue and HTML topic) would make it difficult to standardise the numbering systems.
- The Guideline, ‘links with the same text should point to the same location’ has been violated. This will be remedied.
- The suggestions of changing the link text to include information about its destination (such as “definition of important” rather than “important” and “techniques for providing text alternatives for images” rather than “text alternatives”) are interesting proposals. They will be tried out.
- Other WAI Guidelines have located links at the ends of paragraphs rather than embedding them. This is a good solution.
- If other solutions do not make navigation less confusing a method of creating an accessible site map will have to be found.

10.5.8 Report regarding readability

The problems the participants experienced with reading the Guidelines were reported. This included a variety of suggestions from the participants and from the researcher. The provision of printable versions of the documents; the more restricted use of bold text; increased use of bullet points or shorter phrases; and providing a list of user groups affected by each guideline at the end of its description rather than listing the groups within the paragraph in bold text.

10.5.8.1 Discussion on readability

Several responses were made to the suggestions on phrasing.

- The phrases are already being made shorter. Editing continues in each version of the Guidelines.
- It is not necessary to make the phrases shorter and use more bullet points. The Checklist is provided for people who do not read the entire document. However, page authors must read the entire documents in order to understand the Guidelines completely and use them effectively. Those who do not read the entire document should not be catered for.
• It would not be possible to provide separate versions of the documents that contained all the information in bullet points because there can only be one 'official' version of the Guidelines.

Comments were also made on the printing of documents.
• The documents are available in zipped, postscript and PDF versions, which can be printed.
• Most people can print from their browsers. Again there can only be one 'official' version which has to be specified by a URI (universal resource identifier) and is therefore on-line. People can and do print the documents for their own convenience.

The point being made by the researcher was that a printed version of the Guidelines, whether from a zipped, text or PDF version, does not contain the links available in the on-line version. This point was not picked up by any of the respondents, except to point out that there can only be one 'official' version of the Guidelines.

One respondent commented on the listing of user groups, and the use of bold text.
• The idea of providing a list of user groups affected by each Guideline at the end of a paragraph is a good idea. This will be considered.
• Others have also commented that the bold text makes the documents difficult to read. Much of the bold text is currently being removed.

In more recent versions of the Guidelines the groups of users affected by each guideline are mentioned within the paragraph but not listed at the end of the paragraph. It is not known why the working group did not take up this latter idea.

10.5.9 Report regarding audience

The researcher reported that some participants thought they did not have enough knowledge to follow the Guidelines and Techniques. Others thought they were too basic and repetitive for more experienced page authors. A related issue is that the Guidelines do not cater for page authors who use authoring tools. The researcher suggested that the Guidelines should state their intended audience, and indicate that some tools do not yet support implementation of the Guidelines.

10.5.9.1 Discussion on audience

There was no discussion of these points by the working group. However, more recent versions of the Guidelines contain a brief statement regarding the intended audience, "The guidelines are intended for all Web content developers (page authors and site designers) and for developers of authoring tools." (WAI, op. cit.). This statement is expanded in a W3C fact sheet for the WCA Guidelines, which states, "The guidelines are written for a variety of audiences – people who are designing Web sites; people who are checking existing Web sites for accessibility; organizations that require a given level of accessibility for their Web sites; and others who are interested in ensuring that people with disabilities can access information on the Web."94

The researcher is not sure that the Guidelines can be used effectively by all these different types of people.

10.5.10 **Report regarding disability and assistive technology**

The researcher reported that participants had said they wanted (or were observed to need) more information about disability and assistive technology. In particular participants wanted statistics regarding the proportion of the population who is disabled and how users might interact with some of the new tags. The researcher suggested that this information could be provided, either in the main body of the Guidelines or in the existing appendix that contains definitions.

10.5.10.1 **Discussion on disability and assistive technology**

The following comments were made on this issue.

- The question of how many people are disabled is irrelevant, the important points are that some are, and that there are ways of designing Web pages so that they can be used by these groups.
- The primary purpose of the Guidelines is to explain ‘good design’ rather than why pages should be designed for disabled users.
- Explanations of the problems users experience and how the solutions work should perhaps be provided in the Techniques document, rather than the Guidelines.
- The issue is not how many people with disabilities are using the Web, but how many would use it if it was more accessible.
- The important issue is to raise awareness of the key problems and achieve consensus on how to solve, reduce or avoid these problems.

The researcher responded that it is natural that page authors may want to know how many people are likely to be affected if they do, or do not, implement the advice of the Guidelines. However, the fact that many of the Guidelines are aimed at improving access for those using the Web ‘eyes free’ or ‘hands free’ means the consideration of the numbers of people with disabilities becomes less relevant. The researcher expressed concern that no descriptions are provided of how assistive devices might render certain tags and how the user might interact with such tags.
10.6 Appendix 6: The software and hardware used by the visually impaired participants (Chapter 4).

Note: The last 4 rows represent the two participants who performed the task twice using two different browsers and screenreaders.

<table>
<thead>
<tr>
<th>Browser and version</th>
<th>Screen reader and version</th>
<th>Speech synth.</th>
<th>Braille display</th>
<th>Operating system</th>
</tr>
</thead>
<tbody>
<tr>
<td>IE 4.01 plus Lynx 2.8</td>
<td>Jaws 3.20.31</td>
<td>Eloquence</td>
<td>Alva ABT340</td>
<td>Win NT 4.0</td>
</tr>
<tr>
<td>IE 4.01 Window Bridge 2.56</td>
<td>Apollo 1</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>IE 4.01 Jaws 3.2</td>
<td>Dectalk Express</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>VIP 2.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>IE 4.01 Large font / high contrast</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Win 98</td>
</tr>
<tr>
<td>IE 5.0 Jaws 3.2</td>
<td>Dectalk Express</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>WWW SR/2 V2.01 Windows Bridge 2.56</td>
<td>Prose 2020</td>
<td>-</td>
<td>-</td>
<td>OS/2 V3</td>
</tr>
<tr>
<td>IE 3.0 WindowEyes 2.1</td>
<td>Sounding Board</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>IE 4.01 Jaws 3.2.3.1</td>
<td>Doubletalk</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>Lynx 2.8.2-pre VocalEyes</td>
<td>Dectalk Board</td>
<td>-</td>
<td>-</td>
<td>DOS</td>
</tr>
<tr>
<td>IE 4.01 Jaws 3.2.3.1</td>
<td>Eloquence</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>IE 5 Large font / high contrast</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Win 98</td>
</tr>
<tr>
<td>WebSpeak 2.50</td>
<td>-</td>
<td>Softvoice</td>
<td>-</td>
<td>Win 98</td>
</tr>
<tr>
<td>Netscape 4.05 Jaws 3.2.31</td>
<td>Keynote Multimedia</td>
<td>Braille Window</td>
<td>-</td>
<td>Win 98</td>
</tr>
<tr>
<td>IE 4.01 Jaws 3.2</td>
<td>Apollo 2</td>
<td>Alva ABT380</td>
<td>Win 95</td>
<td></td>
</tr>
<tr>
<td>WebSpeak 2.5</td>
<td>-</td>
<td>Softvoice</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>IE 3 Window Eyes 2.1</td>
<td>Multivoice</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>Braillesurf plus IE 3 Draculanet</td>
<td>Proverbe (French)</td>
<td>Clionotebraille 20</td>
<td>Win 95</td>
<td></td>
</tr>
<tr>
<td>IE 4.01 Jaws 3.2</td>
<td>Apollo 2</td>
<td>Alva 280 (80 cell)</td>
<td>Win 95</td>
<td></td>
</tr>
<tr>
<td>WebSpeak 2.25</td>
<td>-</td>
<td>Softvoice</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td>IE 4.01 Jaws 3.2 plus</td>
<td>Apollo</td>
<td>-</td>
<td>-</td>
<td>Win 95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>---</td>
</tr>
<tr>
<td>Hal95 2.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WebSpeak</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>&quot;</td>
</tr>
</tbody>
</table>
10.7 Appendix 7: Messages and questionnaire sent to visually impaired people (Chapter 4).

First message:
As part of a study which is evaluating the Web Content Accessibility Guidelines of the Web Accessibility Initiative, people who are blind or visually impaired are invited to test the accessibility of some web pages.

This is the second stage of the study. In the first stage web page authors were asked to create an accessible web page using the Guidelines. This has provided some insight into the ease-of-use of the Guidelines, and I now want to evaluate the accessibility for visually impaired people of the pages the authors created. The pages have therefore been put on a web site. The evaluation mainly involves viewing different examples and reporting which were found to be more accessible. The address of the site is: http://phoenix.herts.ac.uk/sdrultest/. On the welcome page there are links to a questionnaire and to a zipped version of the site that can be downloaded.

Instructions for testing the pages:
The web site contains examples of various web elements such as: images, imagemaps, forms, frames, tables, and different approaches to the presentation of article headers and references. The task involves viewing all the examples of a particular element, such as images, then answering a few questions about which example was the most accessible for you, and why.
You can evaluate as many elements as you like, in any order. If you do not have time to do all of them it would be useful if you could start with the tables, forms and frames.

The questions are listed at the end of this message, and are also available on the web site. For the tables there are additional questions which require you to interrogate the tables in order to find an answer.

You can use any one of a variety of strategies for entering the answers and sending them in, for example: create a reply to this message, or save this message as a text file; or save the questions from the web page as a text file. You may wish to switch between your browser and email package to answer the questions, or make notes and enter the answers after viewing all the elements. All answers will be kept confidential.

Please send the answers to Chetz Colwell <c.g.colwell@herts.ac.uk. If you receive this message via a mailing list, please do not send answers to the list (otherwise I cannot be responsible for their confidentiality).

I am a research student at the Sensory Disabilities Research Unit of the University of Hertfordshire, UK. I am sponsored by the Economic and Social Research Council of the UK. Please contact me if you require any further information, or if you need assistance with testing the pages, or if you are interested in the results of the study. The results of the study will be available in August 1999, both on the web site, and from Chetz on request. Contact details are provided at the end of this message.
Thank you very much for your time and effort.

Questions:
Thank you for taking the time to evaluate these pages. The 14 questions are listed below and are also available on the web site.

The answers you provide will remain confidential: your name or email address will not be associated with your answers. The results from this study may be published, but all participants will remain anonymous.

Please note that the questions for tables are slightly different to those for other elements. You will need to interrogate each table in order to find an answer which will be different for each table.

About you and your computer:
1) How long have you been using the Web?
2) How many times do you use the Web per month on average?
3) Which browser did you use to view the pages for this evaluation? Which version?
4) Which screenreader did you use? Which version?
5) Did you use a speech synthesiser? If so which one?
6) Did you use a Braille display? If so which one?
7) Which operating system did you use? DOS, Windows, or Other. Which version?

Questions for elements:
8) Tables.
Which example did you find more accessible? Example 1, Example 2, Example 3, Example 4, Example 5, or None were accessible. If you were able to access the tables, please answer the following question for each example: What size was the 2 cm sphere perceived to be with the outside presentation? Example 1: Example 2: Example 3: Example 4: Example 5: Please describe any particular problems you had with the tables.
Please describe any features you particularly liked in the tables.

Forms.
9) Which example did you find more accessible?
   Example 1, or Example 2.
   Please describe any particular problems you had with the forms.
   Please describe any features you particularly liked in the forms.

10) Frames.
    Which example did you find more accessible?
    Example 1, or Example 2.
    Please describe any particular problems you had with the frames.
    Please describe any features you particularly liked in the frames.

11) Imagemaps.
    Which example did you find more accessible?
    Example 1, or Example 2.
    Please describe any particular problems you had with the imagemaps.
    Please describe any features you particularly liked in the imagemaps.

12) Headers.
    Which example did you find more accessible?
    Example 1, Example 2, Example 3, or Example 4.
    Please describe any particular problems you had with the headers.
    Please describe any features you particularly liked in the headers.

13) References.
    Which example did you find more accessible?
    Example 1, Example 2, or Example 3.
    Please describe any particular problems you had with the references.
    Please describe any features you particularly liked in the references.

14) Images.
    Figure 1: Which example did you find more accessible?
    Example 1, Example 2, Example 3, Example 4, Example 5, or Example 6.
    Figure 2: Which example did you find more accessible?
    Example 1, Example 2, or Example 3.
    Figure 3: Which example did you find more accessible?
    Example 1, Example 2, or Example 3.
    Please describe any particular problems you had with the images.
    Please describe any features you particularly liked in the images.

Please send your answers to Chetz Colwell <c.g.colwell@herts.ac.uk.
Thank you for participating in this study.
Regards,
Chetz

---
Follow-up message:

Hi,
I apologise if you receive this message more than once, and / or if you have already taken part in my study.
You may recall that I recently sent a message to this list inviting blind people to evaluate the accessibility of a test site. This is a reminder for anyone who intends to take part, but hasn’t yet done so yet. I need the questionnaires back by Friday 14th May.
I have had fewer replies than I had hoped for, so I would be grateful if you could find the time to evaluate the site, or at least some of it – you do not have to evaluate all the different elements.
I have now included a zipped version of the site so that it can be viewed off-line. If you have been put off taking part in this study for any reason, please let me know. This will help to design future studies in this area to be as ‘user-friendly’ as possible.
The address of the test site is: http://phoenix.herts.ac.uk/sdru/test/ On the welcome page you will find links to the questionnaire and the zipped version of the site. The results of the study will be fed into the development of the Web Content Accessibility Guidelines of the Web Accessibility Initiative. They will also be available on the test site, hopefully by September. I will let all participants know when the results are available.
I look forward to hearing from you soon, preferably by the 14th May.

Thank you.

Regards,
Chetz
PS: I am supported by the Economic and Social Research Council of the UK.
10.8 Appendix 8: Screenshots of examples evaluated by visually impaired people (Chapter 4).

The arrows in figures 1, 3, 4, and 5 represent the association of data cells with headers (and headers with other headers). ‘Tn’ indicates cells to which IDs were allocated. Arrows indicate an association made between data and header cells, and other header cells. Links are represented with underlined text.

Table example 1.

<table>
<thead>
<tr>
<th>Object Type</th>
<th>Actual Size (cm)</th>
<th>Perceived Size: Inside Presentation</th>
<th>Perceived Size: Outside Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cube</td>
<td></td>
<td>Mean (cm)</td>
<td>Over/under estimation</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>18</td>
<td>+13%</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>19</td>
<td>+15%</td>
</tr>
<tr>
<td>2.5</td>
<td>See Note 1</td>
<td>2.6</td>
<td>+10%</td>
</tr>
<tr>
<td>Sphere</td>
<td></td>
<td>Mean (cm)</td>
<td>Over/under estimation</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>24</td>
<td>+15%</td>
</tr>
<tr>
<td>2.5</td>
<td>25</td>
<td>25</td>
<td>+5%</td>
</tr>
</tbody>
</table>

Notes
1. Preliminary investigations showed that a 1 cm edge cube was too difficult for participants to find in the outside presentation, so it was omitted.
2. Preliminary investigations showed that a 2.5 cm edge cube was too big for the virtual space available.

Figure 10.8.1: Four rows from example 1.
Table example 2.

Example 2

Please answer the following question for this table: What size was the 2 cm sphere perceived to be with the outside presentation?

Preceding caption:

Table 1. Mean perceived size of virtual objects with percent over- and under-estimation data from sighted and blind participants combined.

<table>
<thead>
<tr>
<th>Object Type</th>
<th>Actual Size (cm)</th>
<th>Perceived Size: Inside Presentation</th>
<th>Perceived Size: Outside Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (cm)</td>
<td>Over/under estimation</td>
<td>Mean (cm)</td>
</tr>
<tr>
<td>Cube</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>+60%</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>+50%</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>See Note 2</td>
<td></td>
</tr>
<tr>
<td>Sphere</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>12</td>
<td>+27%</td>
<td>15</td>
</tr>
<tr>
<td>23</td>
<td>21</td>
<td>+5%</td>
<td>23</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>See Note 2</td>
<td></td>
</tr>
</tbody>
</table>

Notes

1. Preliminary investigations showed that a 1 cm edge cube was too difficult for participants to find in the outside presentation, so it was omitted.
2. Preliminary investigations showed that a 2.5 cm edge cube was too big for the virtual space available.

Figure 10.8.2: First three rows of example 2 with no association.
Table example 3.

Example 3

This example is in two tables: Table 1 contains data for cubes; Table 2 contains data for spheres.

Please answer the following question for these tables: What size was the 5 cm sphere perceived to be with the outside presentation?

Preceding caption:

Table 1 and 2: Mean perceived size of virtual objects with percent over- and underestimation (data from sighted and blind participants combined).

Table of experiment results for object CUBE

<table>
<thead>
<tr>
<th>ACTUAL SIZE IN CM</th>
<th>1</th>
<th>15</th>
<th>2</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>INSIDE PRESENTATION - PERCEIVED SIZE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean in cm</td>
<td>1.8</td>
<td>1.7</td>
<td>2.4</td>
<td>See Note 2</td>
</tr>
<tr>
<td>Over/Under estimation</td>
<td>+20%</td>
<td>+15%</td>
<td>+20%</td>
<td></td>
</tr>
<tr>
<td>OUTSIDE PRESENTATION - PERCEIVED SIZE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean in cm</td>
<td>See Note 1</td>
<td>1.6</td>
<td>2.0</td>
<td>2.4</td>
</tr>
<tr>
<td>Over/Under estimation</td>
<td>+7%</td>
<td>0%</td>
<td>-7%</td>
<td></td>
</tr>
</tbody>
</table>

Notes

1. Preliminary investigations showed that a 1 cm edge cube was too difficult for participants to find in the outside presentation, so it was omitted.
2. Preliminary investigations showed that a 25 cm edge cube was too big for the virtual space available.

Table of experiment results for object SPHERE

<table>
<thead>
<tr>
<th>ACTUAL SIZE IN CM</th>
<th>1</th>
<th>15</th>
<th>2</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>INSIDE PRESENTATION - PERCEIVED SIZE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean in cm</td>
<td>2.1</td>
<td>2.3</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>Over/Under estimation</td>
<td>+20%</td>
<td>+15%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>OUTSIDE PRESENTATION - PERCEIVED SIZE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean in cm</td>
<td>2.2</td>
<td>1.6</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>Over/Under estimation</td>
<td>-20%</td>
<td>-20%</td>
<td>0%</td>
<td></td>
</tr>
</tbody>
</table>

Notes

1. Example 3 | Example 4 | Example 5 | Back to Example 3 | Back to Example 3

Figure 10.8.3: Row headers used in both tables in example 3, showing association between header cells.
Table example 4.

Please answer the following question for this table: What size was the 2 cm sphere perceived to be with the outside presentation?

**Preceding caption:**
Table 1. Mean perceived size of virtual objects with percent over- and under-estimation data from sighted and blind participants combined.

<table>
<thead>
<tr>
<th>Object Type</th>
<th>Actual Size (cm)</th>
<th>Inside presentation: Mean perceived Size</th>
<th>Inside presentation: Over/under estimation</th>
<th>Outside presentation: Mean perceived Size</th>
<th>Outside presentation: Over/under estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cube</td>
<td></td>
<td>Inside presentation: Mean perceived Size</td>
<td>Inside presentation: Over/under estimation</td>
<td>Outside presentation: Mean perceived Size</td>
<td>Outside presentation: Over/under estimation</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>18</td>
<td>+20%</td>
<td>See Note 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>17</td>
<td>+10%</td>
<td>16</td>
<td>-7%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>14</td>
<td>+10%</td>
<td>20</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>See Note 2</td>
<td>k</td>
<td>2.4</td>
<td>7%</td>
</tr>
<tr>
<td>Sphere</td>
<td>1.5</td>
<td>21</td>
<td>+27%</td>
<td>1.2</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>23</td>
<td>+15%</td>
<td>1.9</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>15</td>
<td>0%</td>
<td>2.3</td>
<td>5%</td>
</tr>
</tbody>
</table>

**Notes**
1. Preliminary investigations showed that a 1 cm edge cube was too difficult for participants to find in the outside presentation, so it was omitted.
2. Preliminary investigations showed that a 2.5 cm edge cube was too big for the virtual space available.

---

**Example 4**

Figure 10.8.4: First two rows of example 4, showing association between data cells and headers.
Table example 5.

Please answer the following question for the table: What size was the 2 cm sphere perceived to be with the outside presentation?

Preceding caption:

Table 1: Mean perceived size of virtual objects with percent over- and under-estimation (data from sighted and blind participants combined).

<table>
<thead>
<tr>
<th>Object Type</th>
<th>Actual Size (cm)</th>
<th>Perceived Size: Inside Presentation</th>
<th>Perceived Size: Outside Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (cm)</td>
<td>Over/under estimation</td>
<td>Mean (cm)</td>
</tr>
<tr>
<td>Cube</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>+80%</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>+30%</td>
<td>29</td>
</tr>
<tr>
<td>Sphere</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>+27%</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>25</td>
<td>0%</td>
<td>23</td>
</tr>
</tbody>
</table>

Notes:
1. Preliminary investigations showed that a 1 cm edge cube was too difficult for participants to find in the outside presentation, so it was omitted.
2. Preliminary investigations showed that a 25 cm edge cube was too big for the virtual space available.

Figure 10.8.5: Four rows from example 5, showing association between header and data cells.
Digital Talking Books on a PC: A Usability Evaluation of the Prototype DAISY Playback Software

by Sarah Morley
Sensory Disabilities Research Unit, Department of Psychology,
University of Hertfordshire,
Hatfield, Herts. AL10 9AB, UK
+44 (0)1707 284 639
s.morley@hesanet.uk

1. ABSTRACT

The paper describes the design and evaluation of the first system to play digital talking books on a PC: the DAISY Playback Software. The features of the software for navigating through structured digital audio are described. A detailed usability evaluation of the prototype software was designed and conducted to assess its current usability, in which 13 independently aged participants completed a series of realistic tasks and answered detailed usability questions on the system. Recommendations for improvements are presented which might inform designers of similar systems, such as other digital talking book systems or WWW browsers.

1.1 Keywords
digital talking books, blind and visually impaired readers, auditory navigation, structured information access, evaluation methodology
Digital Talking Books on a PC: A Usability Evaluation of the Prototype DAISY Playback Software

by Sarah Morley
Sensory Disabilities Research Unit, Department of Psychology,
University of Hertfordshire,
Hatfield, Herts, AL10 9AB, UK
+44 (0)1707 284 629
smorley@herts.ac.uk

1. ABSTRACT

This paper describes the design and evaluation of the first system to play digital talking books on a PC: the DAISY Playback Software. The features of the software for navigating through structured digital audio are described. A detailed usability evaluation of this prototype software was designed and conducted to assess its current usability, in which 13 blind or partially sighted participants completed a series of realistic tasks and answered detailed usability questions on the system. Recommendations for improvements are presented which might inform designers of similar systems, such as other digital talking book systems or WWW browsers.

1.1 Keywords

digital talking books, blind and visually impaired readers, auditory navigation, structured information access, evaluation methodology
Form example 1.

Example 1

A form

Personal details:
Please enter your name:

Please enter your address:

Please select your gender:
- Male
- Female

General options:
Please select the browser(s) you generally use:
- Netscape
- Internet Explorer
- Lynx

Please select the type of computer you are using:

Next time you log in you will need a password, please enter the password you wish to use:

If you would like to submit your CV, please use the 'Browse' button to select the file:

Thank you for completing the form.
Form example 2.

Example 2

For easy keyboard navigation MS Internet Explorer is better

A form

Use this form to provide your personal details

Name

Address

Password (to be used next visit)

Gender:
- [ ] Male
- [ ] Female

Select the browser(s) you generally use
- [ ] Netscape
- [ ] Internet Explorer
- [ ] Lynx

Select the type of computer you are using

If you would like to submit your CV, use the 'browse' button to select the file:

File

Thank you for completing this form.

[Submit] [Reset]
Use of a haptic device by blind and sighted people

Chetz Colwell 1, Helen Petrie 1, Diana Kombrot 1, Andrew Hardwick 2 & Stephen Furner 2

1 Department of Psychology, University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK
+44 1707 284639 (c g colwil l hi petre I d e loombrot}Wherts ac uk

2 British Telecommunications plc, Research Laboratories, Martlesham Heath, Ipswich, Suffolk, IP5 9RZ, UK +44 1473 646346
{andrewhardw I stephenfurner}Whtsyst.co.uk
Example 2

Haptic Virtual Reality for Blind Computer Users

Chen Coid, Helen Petrie & Dana Kornhtaet
Department of Psychology
University of Newbraldle, Hefield
Hertfordshire AL10 9AB UK
+44 7937 384259
(e.g. e-mail: h.petrie@de.ac.uk)

Andrew Haworth & Stephen Farren
British Telecommunications plc
Research Laboratories
Martlesham Heath, Ipswich
Suffolk IP5 9EZ UK
+44 1473 646346
(alrew.haworth@tele.farren@bts.yilo.co.uk)
Please note that the links in this example do not go anywhere.

Example 3

Use of a Haptic Device by Blind and Sighted People

Chetz Colwell 1, Helen Petz 2, Dana Kombut 1, Andrew Hardwick 2 & Stephen Furey 2

1 Department of Psychology, University of Herford, Hatfield, Herford, AL10 9AB, UK
+44 1707 354629 (e.g. colwell I h.petz @ the.herts.ac.uk)

2 British Telecommunications plc, Research Laboratories, Marlow, Bucks, UK
+44 1494 466686
(acrew.hardwick @ stephen.fury @ @b-te-htco.co.uk)

1 Example 3 | Example 4 | Back to Example 2 | Back to Example 2 |
Example 4
Use of a haptic device by blind and sighted people

Chris Colwell, Helen Price, Diana Kornbrot
Department of Psychology, University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK
Telephone: +44 1707 234529 Fax: +44 1707 235059
cw.colwell@herts.ac.uk / h.pere@herts.ac.uk / d.kornbrot@herts.ac.uk

Andrew Hardwick & Stephen Furner
British Telecommunications plc, Research Laboratories, Martlesham Heath, Ipswich, Suffolk, IP5 3RE, UK Telephone: +44 1473 664564
andrew.harward@bt.com / stephen.furner@bt.com

For assistance or further information please contact Chris Colwell c.w.colwell@herts.ac.uk.

Disclaimer
Imagemap example 1.

Example 1
Example 2

- Blindness
- Deafness
- Mobility Impairments
- Low Vision
- Learning Impairments

For assistance or further information please contact Chet Colwell: c@rehab@heres.ac.uk

Disclaimer
Haptic perception incorporates both kinaesthetic sensing (i.e. of the position and movement of joints and limbs) and tactile sensing (i.e. through the skin). At present, most VIs use visual displays, with some use of auditory and very little haptic information.

References

Example 2

Haptic perception incorporates both kinesthetic sensing (i.e., the position and movement of joints and limbs) and tactile sensing (i.e., through the skin). At present, most VR systems use visual displays, with some use of auditory and very little haptic information.

References

Reference example 3.

Example 3

Haptic perception incorporates both kinesthesia sensing, i.e. the position and movement of joints and limbs, and tactile sensing, i.e. through the skin [reference 10]. At present, most VRs use visual displays, with some use of auditory and very little haptic information.

References


For assistance or further information please contact Chris Coyle <c.coyle@bath.ac.uk>

Diplomata

348
Figure 1.

Images: Figure 1

The page contains six versions of the same image. The image is from a real article and so the paragraph which accompanied the image in the article is also provided, which puts it into context. In the article the image was also accompanied by the following note: "Figure 1. The Impulse Engine 3000."

Please read the paragraph then through the rest of the page. Then decide which example you prefer.

Accompanying paragraphs:

The device used in the current studies was the Impulse Engine 3000 (shown in Figure 1). The system can display virtual textures and objects which the user can feel using a probe. The probe is the length and diameter of a human finger and has 3 degrees of freedom of motion, i.e. it can move in 3 spatial dimensions forwards backwards, up and down, and left and right. The system provides feedback to the user by monitoring the position of their hand and altering the force accordingly (7). The force is created by three motors which exert resistance against the probe. This gives the user the impression that a texture or object is present.

Example 1

Example 2

Example 3

Example 4

Example 5

Example 6
The first study involved virtual textures with varying groove widths. The dimensions of the virtual textures were as close as possible to those used by Lederman [7], [8], [9]. The difference being that those textures involved grooves with a rectangular waveform whereas the textures used in the current study involved sinusoidal shaped grooves. The widths of the grooves varied from 0.375mm to 1.5mm in steps of 0.125mm and had a fixed amplitude of 0.0625mm (shown in Figure 2).

Example 1

![Example 1](image1)

Example 2

![Example 2](image2)

Example 3

![Example 3](image3)
Figure 3

This page contains three versions of the same image, each illustrating a different approach used by a page author. The image is from a real article and so the paragraph which accompanies the image in the article is also provided, which puts it into context. In the article the image was also accompanied by the following label, "Figure 3. Different mental models of the location of the virtual object in virtual space: (a) outside the device; (b) inside the device."

Please read the paragraph then through the rest of the page. Then decide which example you prefer.

Preceeding paragraph:

Handwerk, Rush, Purton & Sessions (4) observed an interesting phenomenon associated with the Impulse Engine 3030, whereby people differ in terms of where they think the virtual space is located in real space. Some people have a mental image of the virtual space being outside the device, so that virtual objects are felt to be near the hand and are touched by the end of the probe that held figure 3a. In contrast, others imagine the virtual space to be within the device, so that virtual objects are touched by the other end of the probe (figure 3b).
### Tables example 1

```
<TABLE BORDER=2>
<TR>
<TH id="t1">Object Type</TH>
<TH id="t2"> Actual Size (cm) </TH>
<TH id="t3" COLSPAN="2"> Perceived Size: Inside Presentation </TH>
<TH id="t4" COLSPAN="2"> Perceived Size: Outside Presentation </TH>
</TR>
<TR><TD></TD><TD></TD>
<TH headers="t3" id="t5"> Mean (cm) </TH>
<TH headers="t3" id=' t6"> Over/under estimation</TH>
<TH headers="t4" id="t7"> Mean (cm) </TH>
<TH headers="t4" id="t8"> Over/under estimation</TH>
</TR>
<TR><TH id=" t9" rowspan="4">Cube</TH>
<TD headers="tr"> <CENTER>1</CENTER> </TD>
<TD headers="t5"> <CENTER>1.8</CENTER> </TD>
<TD headers="t6"> <CENTER>+80%</CENTER> </TD>
<TD headers="t7"> <CENTER>See &lt;A HREF="#note 1 below">Note 1</A&gt;</CENTER> </TD>
<TD headers="t8"> <CENTER>-</CENTER> </TD>
</TR>
<TR><TD headers="t2"> <CENTER>1.5</CENTER> </TD>
<TD headers="t5"> <CENTER>1.7</CENTER> </TD>
<TD headers="t6"> <CENTER>+13%</CENTER> </TD>
<TD headers="t7"> <CENTER>1.6</CENTER> </TD>
<TD headers="t8"> <CENTER>+7%</CENTER> </TD>
</TR>
<TR><TD headers="t2"> <CENTER>2</CENTER> </TD>
<TD headers="t5"> <CENTER>2.4</CENTER> </TD>
<TD headers="t6"> <CENTER>+20%</CENTER> </TD>
<TD headers="t7"> <CENTER>2.0</CENTER> </TD>
<TD headers="t8"> <CENTER>0%</CENTER> </TD>
</TR>
<TR><TD headers="t2"> <CENTER>2.5</CENTER> </TD>
<TD headers="t5"> <CENTER>See &lt;A HREF="#note 2 below">Note 2</A&gt;</CENTER> </TD>
<TD headers="t6"> <CENTER>-</CENTER> </TD>
<TD headers="t7"> <CENTER>2.4</CENTER> </TD>
<TD headers="t8"> <CENTER>-7%</CENTER> </TD>
</TR>
```

352
### Tables example 2

```html
<table BORDER="2">
  <tr>
    <th>Object Type</th>
    <th>Actual Size (cm)</th>
    <th COLSPAN="2">Perceived Size: Inside Presentation</th>
    <th COLSPAN="2">Perceived Size: Outside Presentation</th>
  </tr>
  <tr>
    <td>Cube</td>
    <td>1.5</td>
    <td>M1.5</td>
    <td>+27% M1.8</td>
    <td>-20% M1.2</td>
  </tr>
  <tr>
    <td>2</td>
    <td>2.3</td>
    <td>M2.3</td>
    <td>+15% M1.8</td>
    <td>-10% M1.2</td>
  </tr>
  <tr>
    <td>2.5</td>
    <td>2.5</td>
    <td>M2.5</td>
    <td>0% M2.3</td>
    <td>-8% M2.1</td>
  </tr>
</table>
```
<table>
<thead>
<tr>
<th></th>
<th>1.8</th>
<th>+80%</th>
<th>See <a href="#note1">Note 1</a></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.5</td>
<td>1.7</td>
<td>+13%</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.4</td>
<td>+20%</td>
<td>2.0</td>
</tr>
<tr>
<td>Sphere</td>
<td>1.5</td>
<td>2.1</td>
<td>+27%</td>
<td>1.2</td>
</tr>
<tr>
<td>Sphere</td>
<td>2</td>
<td>2.3</td>
<td>+15%</td>
<td>1.5</td>
</tr>
<tr>
<td>Sphere</td>
<td>2.5</td>
<td>2.5</td>
<td>0%</td>
<td>2.3</td>
</tr>
</tbody>
</table>

1. See [Note 1](#note1) below.
2. See [Note 2](#note2) below.
Tables example 3 (two tables)

```
<TABLE WIDTH="100%" BORDER=1>
<CAPTION> Table of experiment results for object CUBE</CAPTION>
<TR>
<TH WIDTH="45">ACTUAL SIZE IN CM</TH>
<TD WIDTH="10%" ALIGN=CENTER>1</TD>
<TD WIDTH="10%" ALIGN=CENTER>1.5</TD>
<TD WIDTH="10%" ALIGN=CENTER>2</TD>
<TD WIDTH="10%" ALIGN=CENTER>2.5</TD>
</TR>
<TR>
<TH WIDTH="30%" id="1">INSIDE PRESENTATION - PERCEIVED SIZE</TH>
<TD WIDTH="20%"></TD>
<TD WIDTH="20%"></TD>
<TD WIDTH="20%"></TD>
<TD WIDTH="20%"></TD>
</TR>
<TR>
<TD WIDTH="30%" headers="1"><STRONG>Mean in cm</TD>
<TD WIDTH="20%">1.8</TD>
<TD WIDTH="20%">1.7</TD>
<TD WIDTH="20%">2.4</TD>
<TD WIDTH="20%">See Note 2</TD>
</TR>
<TR>
<TR>
<TD WIDTH="30%" headers="1">Over/Under estimation</TD>
<TD WIDTH="20%">+80%</TD>
<TD WIDTH="20%">+13%</TD>
<TD WIDTH="20%">+20%</TD>
<TD WIDTH="20%">-10%</TD>
</TR>
</TR>
<TR>
<TH WIDTH="30%" id="2">OUTSIDE PRESENTATION - PERCEIVED SIZE</TH>
<TD WIDTH="20%"></TD>
<TD WIDTH="20%"></TD>
<TD WIDTH="20%"></TD>
<TD WIDTH="20%"></TD>
</TR>
<TR>
<TD WIDTH="30%" ALIGN=CENTER headers="2"><STRONG>Mean in cm</TD>
<TD WIDTH="20%">See Note 1</TD>
<TD WIDTH="20%">1.6</TD>
</TR>
```

355
<table>
<thead>
<tr>
<th>AC'TUAL SIZE IN CM</th>
<th>INSIDE PRESENTATION - PERCEIVED SIZE</th>
<th>OUTSIDE PRESENTATION - PERCEIVED SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>2.5</td>
<td></td>
</tr>
</tbody>
</table>

**Over/Under estimation**

<table>
<thead>
<tr>
<th>2.0</th>
<th>2.4</th>
</tr>
</thead>
</table>

**Table of experiment results for object SPHERE**
### Tables example 4

This table shows the relationship between the actual sizes of objects and their perceived sizes.

<table>
<thead>
<tr>
<th>Object Type</th>
<th>Actual Size (cm)</th>
<th>Inside presentation: Mean perceived Size</th>
<th>Inside presentation: Over/under estimation</th>
<th>Outside presentation: Mean perceived Size</th>
<th>Outside presentation: Over/under estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cube</td>
<td>1</td>
<td>1.8</td>
<td>+80%</td>
<td>2.0</td>
<td>0%</td>
</tr>
<tr>
<td>Cube</td>
<td>1.5</td>
<td>1.7</td>
<td>+13%</td>
<td>1.6</td>
<td>+7%</td>
</tr>
<tr>
<td>Cube</td>
<td>2</td>
<td>2.4</td>
<td>+20%</td>
<td>2.0</td>
<td>0%</td>
</tr>
<tr>
<td>Cube</td>
<td>2.5</td>
<td>2.8</td>
<td>+10%</td>
<td>2.5</td>
<td>+5%</td>
</tr>
<tr>
<td>Cube</td>
<td>3</td>
<td>3.2</td>
<td>+15%</td>
<td>3.0</td>
<td>+10%</td>
</tr>
</tbody>
</table>

Note: See [Note 1 below](#note1).
<table BORDER="2" summary="This table charts the errors in perception of several different objects">
   <tr>
      <th id="t1">Object Type</th>
      <th id="t2"><p align="center">Actual Size (cm)</th>
      <th id="t3" COLSPAN="2"><p align="center">Perceived Size: Inside Presentation</th>
      <th id="t4" COLSPAN="2"><p align="center">Perceived Size: Outside Presentation</th>
   </tr>
   <tr>
      <td>&nbsp;</td>
      <td>358</td>
      <td>&nbsp;</td>
      <td>&nbsp;</td>
   </tr>
</table>

Tables example 5
<table>
<thead>
<tr>
<th></th>
<th>Mean (cm)</th>
<th>Over/under estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cube</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.8</td>
<td>+80%</td>
</tr>
<tr>
<td>1.5</td>
<td>1.7</td>
<td>+13%</td>
</tr>
<tr>
<td>2</td>
<td>2.4</td>
<td>+20%</td>
</tr>
<tr>
<td>2.5</td>
<td>See Note 2</td>
<td></td>
</tr>
<tr>
<td><strong>Sphere</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>2.1</td>
<td>+27%</td>
</tr>
<tr>
<td>2</td>
<td>2.3</td>
<td>+15%</td>
</tr>
<tr>
<td>2</td>
<td>2.4</td>
<td>-7%</td>
</tr>
</tbody>
</table>

Note 1: See HREF="#note 1 below"

Note 2: See HREF="#note 2 below"
### Frames example 1

Frames.html

```html
<html>
<head>
<title>Frames: Example 1</title>
</head>

<frameset rows="85%, 15%">
  <frame src="papers.html" name="top">
  <frame src="contents.html" name="bottom">
</frameset>

<noframes>
This site uses frames. Click <a href="contents.html">here</a> to see a non-frames version of this site.
</noframes>

</html>
```

Contents.html

```html
<html>
<head>
<title>Table of Contents</title>
</head>

<BODY BGCOLOR="#FFFFFF">

<center><a href="papers/colwell.html" target="top"><IMG SRC="HD.gif" ALT="Haptic device"></a>

</center>
```

360
<a href="papers/morley.htm" target="top"><img src="TB.gif" alt="Talking books"></a>

<a href="papers/morlpetr.htm" target="top"><img src="AN.gif" alt="Auditory Navigation"></a>

<P>
| <A HREF="../frames.html" target="_top">Back to Main Frames page</A> |

</center>
</body>
</html>

Papers.html
<html>
<head>
<title>Created with DiDaPro</title>
</head>
<body>

<h1>Papers on SDRU website</h1>

</body>
</html>

Frames example 2
Frames.html
<html>
<head>
<title>Frames: Example 2</title>
<frameset rows="13%, 83%">

<frame src="title.html" noresize scrolling=no >

<frameset cols="22%, *">
<frame name="left" src="contents.html" title=contents noresize>
<frame name ="main" src="papers.html" title=papers noresize>
</frameset>

</frameset>

<NOFRAMES> Select to go to the <A href="contents.html" title="table_of_contents">table of contents</A>
</NOFRAMES>

</body>
</html>

Contents.html
Use of a haptic device by blind and sighted people: perception of virtual textures and objects

Digital Talking Books on a PC: A Usability Evaluation of the Prototype DAISY Playback Software

Auditory Navigation in Hyperspace: Design and Evaluation of a Non-Visual Hypermedia System for Blind Users

Back to Main Frames page
<html>

Form example 1

<HTML>
<HEAD>
<TITLE>Forms: Example 1</TITLE>
</HEAD>

<BODY BGCOLOR="#FFFFFF">

<CENTER><H1>Example 1</H1></CENTER>

<CENTER><H1>A form</H1></CENTER>

<FORM action="" method="POST">

<FIELDSET>

<LEGEND align="top">Personal details:<BR></LEGEND>
Please enter your name:<BR>
<INPUT TYPE="text" NAME="name">

<P>Please enter your address:<BR>
<TEXTAREA NAME="address" ROWS=5 COLS=27>
Please enter your address:<TEXTAREA>

<P>Please select your gender:<BR>
<INPUT TYPE="radio" NAME="gender" VALUE="male" CHECKED>
<LABEL for="male">Male</LABEL>
<INPUT TYPE="radio" NAME="gender" VALUE="female">
<LABEL for="female">Female</LABEL>
</FIELDSET>

<FIELDSET>

<LEGEND align="top">General options:<BR></LEGEND>

<P>Please select the browser(s) you generally use:<BR>
<LABEL for="Netscape">Netscape</LABEL>
<INPUT TYPE="checkbox" NAME="browser" VALUE="NS">

<LABEL for="Internet Explorer">Internet Explorer</LABEL>
<INPUT TYPE="checkbox" NAME="browser" VALUE="IE">

<LABEL for="Lynx">Lynx</LABEL>
<INPUT TYPE="checkbox" NAME="browser" VALUE="Lynx" CHECKED>
</FIELDSET>

<P>Please select the type of computer you are using:<BR>
<SELECT NAME="comp">

<OPTION SELECTED>PC
<OPTION>Mac
<OPTION>Other

</SELECT>

</FORM>

</BODY>
</html>
Next time you log in you will need a password, please enter the password you wish to use:

If you would like to submit your CV, please use the 'browse' button to select the file:

Thank you for completing this form.

Form example 2

For easy keyboard navigation MS Internet Explorer is better.

A form

Use this form to provide your personal details.

Address:

Password (to be used next visit):

Gender:
Male <BR><INPUT TYPE="radio" tabindex="4" NAME="gender" VALUE="male" CHECKED>Male <BR><INPUT TYPE="radio" tabindex="5" NAME="gender" VALUE="female">Female

Select the browser(s) you generally use:
<BR><INPUT TYPE="checkbox" tabindex="6" NAME="browser" VALUE="NS">Netscape <BR><INPUT TYPE="checkbox" tabindex="7" NAME="browser" VALUE="IE">Internet Explorer <BR><INPUT TYPE="checkbox" tabindex="8" NAME="browser" VALUE="Lynx" checked>Lynx

Select the type of computer you are using:
<SELECT NAME="comp">
<OPTION SELECTED>PC
<OPTION>Mac
<OPTION>Other
</SELECT>

If you would like to submit your CV, use the 'browse' button to select the file:
<BR> <LABEL for="file">File:</LABEL><INPUT accesskey="B" tabindex="9" TYPE="file" id="file" NAME="cv">

Thank you for completing this form.

<INPUT TYPE="submit" tabindex="10" VALUE="Submit">
<INPUT TYPE="reset" VALUE="Reset">

</FORM>

</BODY>
</HTML>

Header example 1

<HTML>
<HEAD>
<TITLE>Headers: Example 1</TITLE>
</HEAD>

<BODY BGCOLOR="#FFFFFF">
<CENTER><H1>Example 1</H1></CENTER>

<H1>Use of a haptic device by blind and sighted people</H1>

<CENTER>Chetz Colwell <SUP>1</SUP>, Helen Petrie <SUP>1</SUP>, Diana Kornbrot <SUP>1</SUP>, Andrew Hardwick <SUP>2</SUP> &amp; Stephen Furner <SUP>2</SUP> <br>&lt;br&gt; <SUP>1</SUP> Department of Psychology, University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK &lt;br&gt;+44 1707 284629 {c.g.colwell | h.l.petrie | d.e.kornbrot} @herts.ac.uk
Header example 2

<HTML>
<TITLE>Headers: Example 2</TITLE>
<HEAD>
<H1>Example 2</H1>
Haptic Virtual Reality for Blind Computer Users
<br>
Chetz Colwell, Helen Petrie &lt;BR&gt;
Diana Kornbrot&lt;BR&gt;
Department of Psychology&lt;BR&gt;
University of Hertfordshire, Hatfield&lt;BR&gt;
Hertfordshire AL10 9AB UK&lt;BR&gt;
+44 1707 284629&lt;BR&gt;
{c.g.colwell | h.l.petrie | d.e.kornbrot}@herts.ac.uk&lt;BR&gt;
<br>
Andrew Hardwick &lt;BR&gt;
Stephen Fumer&lt;BR&gt;
British Telecommunications plc &lt;BR&gt;
Research Laboratories&lt;BR&gt;
Martlesham Heath, Ipswich&lt;BR&gt;
Suffolk IP5 7RE UK&lt;BR&gt;
+44 1473 646846&lt;BR&gt;
{andrew.hardwick | stephen.fumer}@bt-sys.bt.co.uk&lt;BR&gt;
</HEAD>
<BODY BGCOLOR="#FFFFFF">
</BODY>
</HTML>

Header example 3

<HTML>
<TITLE>Headers: Example 3</TITLE>
<HEAD>

</HEAD>

</BODY>
</HTML>
Please note that the links in this example do not go anywhere.

Chetz Colwell, Helen Petrie, Diana Kornbrot, Andrew Hardwick & Stephen Furner. Department of Psychology, University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK +44 1707 284629 {c.g.colwell | h.l.petrie | d.e.kornbrot} @herts.ac.uk

British Telecommunications plc, Research Laboratories, Martlesham Heath, Ipswich, Suffolk, IP5 7RE, UK +44 1473 646846 [andrew.hardwick | stephen.furner] @bt-sys.bt.co.uk

<table>
<thead>
<tr>
<th>TABLE BORDER=0 WIDTH=&quot;90%&quot; ALIGN=LEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD WIDTH=&quot;20%&quot;</td>
</tr>
<tr>
<td>TD WIDTH=&quot;30%&quot;</td>
</tr>
<tr>
<td>TD WIDTH=&quot;50%&quot;</td>
</tr>
<tr>
<td>TD WIDTH=&quot;45%&quot;</td>
</tr>
<tr>
<td>TD WIDTH=&quot;15%&quot;</td>
</tr>
<tr>
<td>TD WIDTH=&quot;15%&quot;</td>
</tr>
<tr>
<td>TD WIDTH=&quot;15%&quot;</td>
</tr>
<tr>
<td>TD WIDTH=&quot;15%&quot;</td>
</tr>
</tbody>
</table>

Headers: Example 4

<HTML>
<HEAD>
<TITLE>Headers: Example 4</TITLE>
</HEAD>

<BODY BGCOLOR="#FFFFFF">
<CENTER><H1>Example 4</H1></CENTER>
Use of a haptic device by blind and sighted people

Chetz Colwell, Helen Petrie, Diana Kornbrot.
Department of Psychology, University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK
Telephone: +44 1707 284629 Fax: +44 1707 285059

Andrew Hardwick & Stephen Furner.
British Telecommunications plc, Research Laboratories, Martlesham Heath, Ipswich, Suffolk, IP5 7RE, UK
Telephone: +44 1473 646846

Images Figure 1

Example 1

Example 2

Example 3

Example 4

Example 5
Example 6

Figure 1: The Impulse Engine 3000

Images figure 2

Example 1

Example 2

Example 3

Images figure 3

Example 1

Example 2
Example 3

Imagemap example 1

Imagemaps: Example 2

370
Imagemap example 2

<HTML>
<HEAD>
<TITLE>Imagemaps: Example 2</TITLE>
</HEAD>

<BODY BGCOLOR="#FFFFFF">
<CENTER><H1>Example 2</H1></CENTER>

<A NAME="imap2"></A>
<IMG SRC="images/imgmap1.gif" WIDTH=394 HEIGHT=78 ALIGN=bottom usemap="#imgmap1" border="0">
<map name="imgmap1">
<area shape="rect" alt="Blindness" coords="0,0,71,77" href="dis2/blind.html#blind">
<area shape="rect" alt="Deafness" coords="72,0,162,77" href="dis2/deaf.html#deaf">
<area shape="rect" alt="Mobility Impairments" coords="163,0,236,77" href="dis2/mobil.html#mobil">
<area shape="rect" alt="Low Vision" coords="237,0,314,77" href="dis2/lv.html#lovis">
<area shape="rect" alt="Learning Impairments" coords="315,0,393,77" href="dis2/learn.html#learn">
<area shape="default" nohref>
</map>

<UL>
<L1><A HREF="dis2/blind.html#blind">Blindness</A></L1>
<L1><A HREF="dis2/deaf.html#deaf">Deafness</A></L1>
<L1><A HREF="dis2/mob.html#mobil">Mobility Impairments</A></L1>
<L1><A HREF="dis2/lv.html#lovis">Low Vision</A></L1>
<L1><A HREF="dis2/learn.html#learn">Learning Impairments</A></L1>
</UL>

</BODY>
</HTML>
Haptic perception incorporates both kinaesthetic sensing, (i.e. of the position and movement of joints and limbs), and tactile sensing, (i.e. through the skin) <A HREF="#10">[10]</A>. At present, most VEs use visual displays, with some use of auditory and very little haptic information.

Haptic perception incorporates both kinaesthetic sensing, (i.e. of the position and movement of joints and limbs), and tactile sensing, (i.e. through the skin) [reference 10]. At present, most VEs use visual displays, with some use of auditory and very little haptic information.

References

10.10 Appendix 10: Comparison of syntax used for the checkbox (Chapter 4)

a) Used with the Label tag by the page author of example 1, and b) used in the HTML 4.0 specification.

a) The page author implemented the following code:

```html
<INPUT TYPE="checkbox" NAME="browser" VALUE="NS">
<LABEL for="Netscape">Netscape</LABEL>
```

b) In the HTML 4.0 specification⁹⁵ the ‘input type’ and ‘label’ tags are used in the reverse order:

```html
<LABEL for="fname">First Name</LABEL>
<INPUT type="text" name="firstname" id="fname">
```

The page author for the example 1 of forms implemented the ‘input type’ and ‘label’ tags in the wrong order. Also, they implemented the ‘value’ attribute rather than the ‘id’ attribute.

⁹⁵ See http://www.w3.org/TR/Rec-html40/
10.11 Appendix 11: Order in which virtual objects were presented (Chapter 6).

1.5 cm³ Cube (outside presentation)
1.0 cm³ Cube (inside presentation)
2.0 cm Sphere (outside presentation)
1.5 cm Sphere (inside presentation)
41° Sheared cube
70° Rotated cube
2.5 cm Sphere (outside presentation)
2.0 cm³ Cube (outside presentation)
2.0 cm Sphere (inside presentation)
30° Rotated cube
18° Sheared cube
2.0 cm³ Cube (inside presentation)
1.5 cm Sphere (outside presentation)
1.5 cm³ Cube (inside presentation)
2.5 cm Sphere (inside presentation)
64° Sheared cube
50° Rotated cube
2.5 cm³ Cube (outside presentation)
Figure 10.12: Representation of the cubes in the 'outside' presentation.

Blind participants were shown tactile representations of virtual objects shown to shielded participants (Chapter 6).
Figure 10.12: Representation of spheres in the outside presentation.
Figure 10.12.3: Representation of cubes in the inside presentation.
Figure 10.12.4: Representation of the spheres in the inside presentation.
Figure 10.12.5: representation of the sheared cubes (inside presentation).
Figure 10.12.6: representation of the rotated cubes (outside presentation).
Addendum: Explanation of the calibration error with the IE3000 (Chapter 6)

Some months after the completion of the studies reported in Chapter 6, and after the results had been published, a calibration error was discovered in the IE3000 by the researcher's colleagues at BT. Initially it was thought that it was just the virtual textures that were affected by this error. It was established that the dimensions of the virtual textures used in the study were in fact 1.8 times larger than was originally thought. The dimensions of the textures were intended to be as close as possible to those used by Lederman, (Lederman, 1974; 1981; Lederman & Taylor, 1972) but this was not the case. The data from the virtual textures study were re-analysed using the correct dimensions and the correct results are presented in Chapter 6.

Initially, the effect of the calibration error was less clear in relation to the virtual objects. When the error was discovered it was thought that it had no impact on the size of the virtual objects. However, in more recent discussions with BT it has emerged that the size of the cubes and spheres was indeed affected by the error. (The angle of the rotated and sheared cubes was not affected because the angle would remain the same whatever the size of the object.)

The main problem caused by this error is that the choices presented to the participants did not accurately represent the virtual objects they were touching. For the smaller objects the participants did have one or two potential choices, but this is not the case for the larger objects, as can be seen in Table1.

Table 1: the intended and actual sizes of the virtual cubes and spheres, and the potential choices available to the participants.

<table>
<thead>
<tr>
<th>Object</th>
<th>Intended size</th>
<th>Actual size</th>
<th>Potential choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubes (inside presentation)</td>
<td>1.5, 2.0</td>
<td>1.8, 2.7</td>
<td>1.5, 2.0</td>
</tr>
<tr>
<td>Spheres (inside presentation)</td>
<td>1.5, 2.0</td>
<td>2.7, 3.6</td>
<td>2.5, None</td>
</tr>
<tr>
<td>Cubes (outside presentation)</td>
<td>1.5, 2.0</td>
<td>2.7, 3.6</td>
<td>2.5, None</td>
</tr>
<tr>
<td>Spheres (outside presentation)</td>
<td>1.5, 2.0</td>
<td>2.7, 3.6</td>
<td>2.5, None</td>
</tr>
</tbody>
</table>

The data for the cubes and spheres has been re-analysed and the corrected results are presented below.
Re-analysis of virtual object data

Two sets of analyses were conducted. The first investigated whether the larger virtual objects were perceived as being larger, even though the response set given was not completely appropriate for this task. The second set of analyses investigated whether the virtual objects were perceived as larger when explored from the inside when compared to explored from the outside (the "Tardis Effect"). These analyses were conducted on the virtual cubes and spheres separately.

Virtual Cubes

Table 2 shows the number of participants (blind and sighted combined) who chose the various response sizes (1.0, 1.5, 2.0 and 2.5 cm) for the virtual cubes which were actually simulated sizes of 1.8, 2.7, 3.6 and 4.5 cm (and not the sizes corresponding to the response set, as had been supposed) after exploration from either the inside or outside of the virtual object.

<table>
<thead>
<tr>
<th>Actual size (cm)</th>
<th>Perceived size (cm)</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>2.5</th>
<th>χ^2</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exploration mode</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>Inside</td>
<td>2</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>6.72</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>Outside</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td>Inside</td>
<td>0</td>
<td>15</td>
<td>6</td>
<td>1</td>
<td>27.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Outside</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>6.00</td>
<td>n.s.</td>
</tr>
<tr>
<td>3.6</td>
<td>Inside</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>19</td>
<td>45.28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Outside</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>6</td>
<td>6.36</td>
<td>n.s.</td>
</tr>
<tr>
<td>4.5</td>
<td>Inside</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>17</td>
<td>35.09</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Outside</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes for table 2:
1. The outside presentation of a 1.8 cm (corrected size) cube was found that the object was found to be difficult to locate in the virtual space.
2. The inside presentation of a 4.5 cm (corrected size) cube was found to be too large for the virtual space: the probe could not touch all the sides and corners of the object.

There were significant deviations from a random distribution of responses in three of the six combinations of size and exploration mode presented (see Table 2). For example, for the largest simulated size (4.5 cm) explored from the outside, 17 out of 22 of the participants (77.3%) chose the largest response option (2.5 cm), even though this was not actually as large as the simulated cube, creating a distribution significantly skewed toward the upper sizes (χ^2 = 35.09, p < 0.001). Even for distributions where there was no significant deviation from a random distribution of responses these showed a substantial clustering of responses in the appropriate
places. For example, for the smallest simulated cube (1.8 cm), 17 participants chose either the 1.5 or 2.0 cm response, which were the two most accurate responses available, although this did not create a significant deviation from a random distribution.

The lack of a significant effect in 50% of cases may well be due to the relative coarseness of the response measure, the confusion over the appropriateness of the response set, the relatively small sample size and the need to combine blind and sighted participants to create an appropriate sample size for the analysis.

To investigate the Tardis effect, the distribution of responses was compared for the inside and outside exploration of the virtual cubes. This was only possible for virtual cubes of 2.7 and 3.6 cm. Friedman’s two way analyses of variance of ranks (Siegel and Castellan, 1988) on each size did not find any significant differences in size estimations between inside and outside exploration of the simulated cubes (for the 2.7 cm cubes: F = 4.80, N = 2, k = 4, n.s.; for the 3.6 cm cubes F = 5.21, N = 2, k = 4, n.s.).

**Virtual Spheres**

Table 3 shows the corresponding number of participants (blind and sighted combined) who chose the various response sizes (1.0, 1.5, 2.0 and 2.5 cm) for the virtual spheres which actually simulated sizes of 1.8, 2.7, 3.6 and 4.5 cm (and not the sizes corresponding to the response set, as had been supposed) after exploration from either the inside or outside of the virtual object.

<table>
<thead>
<tr>
<th>Actual size (cm)</th>
<th>Perceived size (cm)</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>2.5</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exploration mode</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td>Inside</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>7</td>
<td>9.63</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>Outside</td>
<td>16</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>31.10</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>3.6</td>
<td>Inside</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>12</td>
<td>18.73</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Outside</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>4</td>
<td>3.09</td>
<td>n.s.</td>
</tr>
<tr>
<td>4.5</td>
<td>Inside</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>20</td>
<td>51.45</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Outside</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>24.18</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Notes for table 3:
- The inside presentation of a 1.8 cm (corrected size) sphere, like the 1.0 cm cube, was found to be too small: the probe could hardly be moved at all inside the object which meant that it could not be explored.
- The outside presentation of a 1.8 cm (corrected size) sphere was found to be very difficult to locate in the virtual space.

There were significant deviations from a random distribution of responses in five of the six combinations of size and exploration mode presented (see Table 3). As with the cubes, the distributions were generally skewed in the directions expected by the
sizes of the objects in the response set. Thus, for the largest virtual spheres (4.5 cm), for both the inside and outside exploration modes, the distributions are strongly and significantly skewed towards the largest response size (2.5 cm), even though this was not actually as large as the virtual object. One notable exception was the 2.7 cm sphere explored from the outside, in which the distribution was strongly and significantly skewed toward the smallest response (1.0 cm), even though it would have been more accurate to give the largest response (2.5 cm).

To investigate the Tardis effect, the distribution of responses was compared for the inside and outside exploration of the virtual spheres. This was possible for virtual spheres of 2.7, 3.6 and 4.5 cm. Friedman’s two way analyses of variance of ranks (Siegel and Castellan, 1988) on each size did not find any significant differences in size estimations between inside and outside exploration of the simulated cubes (for the 2.7 cm spheres: F = 0.16, N = 2, k = 4, n.s.; for the 3.6 cm spheres F = 2.55, N = 2, k = 4, n.s; for the 4.5 spheres: F = 5.55, N = 2, k = 4, n.s.).

Discussion

The re-analysed data show that participants did generally perceive the differences in the sizes of the virtual objects, in spite of the coarseness and inaccuracy of the response set. The effect was considerably stronger in the spheres than in the cubes. It is not clear why such a difference should occur.

However, the re-analysis of the data failed to show the Tardis effect found in the original analysis of the data. This raises the question of whether this effect was somehow an artefact of the original inaccurate analysis.

This question is addressed by considering the data collected as part of an undergraduate project (Bruns, 1998) jointly supervised by Helen Petrie and the researcher. In this project, a number of improvements to the methodology used in the original study of the size estimation were made. Most importantly, the response system was improved: instead of a multiple choice from visual/tactile representations of the objects, as used in the original study reported here, a small plastic ruler was used. Two small cardboard sleeves slid over the ruler, allowing the participant to create a length between the sleeves of any length they chose. Participants could again assess this length either tactually, by feeling the sleeves and the length of the ruler between the sleeves, and for sighted participants, visually, by looking at the ruler. This allowed participants to make a much more precise, and interval-based, estimate of the size of the virtual objects.

In this study, 10 sighted and 7 blind participants estimated the sizes of virtual cubes of 2, 3, 4 and 5 cm and virtual spheres of 2, 3, 4 and 5 cm, explored from both inside and outside the objects. The same apparatus, the IE 3000 was used, and this experiment was conducted after BT had re-calibrated the apparatus.

Figures a and b (below) show the size estimates for the virtual cubes and spheres respectively, in relation to the simulated size of the objects, in the inside and outside exploration conditions. A significant difference was found in size estimation between different sizes of virtual objects (F3,45 = 61.00, p <.000), with the larger objects being perceived as larger. Indeed, Figures a and b show that the sizes track the simulated
sizes appropriately (increasing in size) except in one case, the largest size of sphere (5 cm) when explored from the outside.

Interestingly, all virtual objects were underestimated in size in comparison to the simulated sizes, but there was a significant difference in the magnitude of the underestimation between inside and outside exploration ($F_{1,15} = 30.90, p < .000$). When explored from the outside, the objects were perceived as significantly smaller than when perceived from the inside. Thus the Tardis Effect was very clearly found.
Indeed, Table 4 shows that the level of underestimation was extremely similar for the cubes and the spheres, and the difference between the inside and outside exploration modes was very substantial — an average of 19.2% underestimation when exploring from the inside and an average 43.6% underestimation when exploring from the outside.

**Table 4: Underestimation of virtual objects (percentages) in inside and outside exploration mode**

<table>
<thead>
<tr>
<th>Exploration mode</th>
<th>Inside</th>
<th>Outside</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cube</td>
<td>19.5%</td>
<td>43.5%</td>
</tr>
<tr>
<td>Sphere</td>
<td>18.8</td>
<td>43.8</td>
</tr>
<tr>
<td>Mean</td>
<td>19.2</td>
<td>43.6</td>
</tr>
</tbody>
</table>

*References for Addendum.*
