Human-Computer Interaction -- INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 631-638

How Blind Users' Mental Models Affect Their Perceived Usability of an Unfamiliar Screen Reader

Sri Hastuti Kurniawan, Alistair G. Sutcliffe, Paul L. Blenkhorn

Department of Computation, UMIST, P.O. Box 88, Manchester M60 1QD, U.K.

{s.kurniawan, a.sutcliffe, p.blenkhorn}@co.umist.ac.uk

Abstract: This study investigates blind users' mental models of Windows environment and their strategies in coping with new desktops and applications. The relationship between users' mental model and their perceived usability problems when using an unfamiliar screen reader is also reported. Blind users in this study possess a functional or structural mental model or a combination of thereof. They also have a rich and highly procedural strategy for coping with a new/unfamiliar Windows environment and application. Users' established mental model developed from using their familiar screen readers was found to contribute to what users perceived as usability problems when using an unfamiliar screen reader that does not work in a similar way to their familiar screen readers or does not have desired features.

Keywords: blind user, mental model, screen reader

1 Introduction

For blind and visually impaired people, computers are considered as one of the keys to the occupational and social integration (Bornemann-Jeske, 1996). However, despite a general understanding that blind and visually impaired people interact with computers differently from sighted users, little research has investigated blind users' mental models[1] and how such models may contribute to usability problems.

The majority of usability studies of computer aids for the blind and visually impaired users have been carried out with normally sighted users (Barnicle, 2000), despite the findings that there are limitations of MS-Windows software conditioned by blindness (Bornemann-Jeske, 1996). A Web study found that even when aided with screen readers and screen magnifiers, users with disabilities were three times worse in terms of usability measurements compared to normally sighted users (Nielsen, 2001).

One of the most important computer interfaces for blind people is a screen reader. Windows environment (throughout this paper, this term is used to describe an interaction style with independent and potentially concurrent task specific areas of a computer screen) and most Windows-based applications claim to be accessible for blind users using shortcut keys and screen readers. However, even though blind users are able to use Windows environment and applications (sometimes in a limited way), the level of complexity of this software makes it unrealistic to expect an average blind user to develop an adequate mental model without some assistance from sighted persons (Landau, 1999).

Mental models are important because they help users predict commands and interpret system actions (Norman, 1988). Unfortunately, blind users can only acquire mental models through audio-mediated experience and this is not always supported in Windows environment.

The present study investigates whether there is a relationship between blind users' mental model of Windows environment, application and screen readers and their perceived usability problems when using an unfamiliar screen reader. This study aims to answer three research questions:

1. How do blind computer users interact with Windows environment?

This question is decomposed into two sub questions:

  a. What mental model of Windows environment do blind users possess?
  b. What strategies do blind users employ for interacting with a new/unfamiliar system?

[1] Mental models are human's internal conceptual models of how things work, or how things would behave in a particular situation.

2. What problems do blind users face when learning to use a new/unfamiliar screen reader?
3. Do blind users have a consistent mental model for different screen readers?

Participant	CE	BH	PH	SS	RS
Gender	M	M	M	F	M
Age (yrs)	54	31	44	47	36
Computer use (yrs)	13	21	24	31	12
Job	Unemployed (IT background)	Education officer	Second line IT support	Student	Lecturer in Computer Science
Screen readers used	JAWS	HAL, JAWS, Window Eyes, Window bridge, LookOut	HAL	HAL, JAWS, Window Eyes, Outspoken, IBM Screen reader	JAWS, IBM Screen reader, Vocaleyes
Blindness	Some vision	Total	Total	Almost total	Total
Blind since	1966	birth	birth	birth	1989

Table 1: Participants' demographic and computer experience data.

2 Method

2.1 Participants

Five blind computer users who accessed computers using screen readers on a regular basis were recruited to participate in the study. One of the requirements of this study is that the participants must have no or very little experience with the tested screen reader. All of them were familiar with Windows environment and word processors or text editors. Participants' demographic data are listed in Table 1.

2.2 Procedures

The study consists of two sessions. The first session was performed either in participants' home or at the experimenter's laboratory. During the session, only the experimenters (two or more persons) and the participant were present.

At the beginning of the first session, participants' demographics and computer experience data were collected through an interview. Limited instruction was given on the most important functions of the tested screen reader. Users were then given some time to familiarize themselves with the Windows environment and the screen reader in the experimenter's laptop.

Participants were reminded that there was no criterion for failing the task or time limit to finish the task. The participants were also asked to think out aloud when using the tested screen reader, either to give comments, feedback or complaints.

The main session consists of word-editing tasks. The tasks in the present study utilized features that users most likely had used before.

The instructions in the main session were read out by an experimenter. They are as follows:

1. Open a Microsoft Word document called LookOut_manual.doc (this was the electronic manual of tested screen reader).
2. Read the document using the screen reader.
3. Cut a paragraph and paste it to a new document.
4. Type several words of your choice.
5. Create a header with a new font style and size.
6. Save the new document in My Documents folder as [participant's name].doc.

At the end of the session, the participants were interviewed to investigate their mental models of Windows environment and the screen reader and to investigate the usability problems in more depth. The questions asked at the end of the main session were as follows:

1. What is your mental model of the Windows environment, desktop or application?
2. What is your strategy in coping with a new application or an unfamiliar computer?
3. What is your strategy for operating your own computer or a familiar computer environment?
4. What problems bother you about this screen reader?
5. What problems do you have with your current screen reader?
6. What differences did you find between this screen reader and your current screen reader?
7. What features do you think a good screen reader should have?

The whole first session (from pre-session questionnaire to post-session interview) took around one and a half hours.

In the follow-up session, participants were interviewed on the phone to get a sense of the structure of their mental models by asking them to spontaneously list up to 20 Microsoft Word functions. However, they were allowed to stop at anytime. Only four participants from the first session were involved in this session.

2.3 Stimulus

The experiments were performed with a Pentium-based laptop equipped with an external PC keyboard running MS-Windows ME. The word processor program used was Microsoft Word 2000 and the screen reader used was LookOut version 1.0. LookOut is a Windows-based screen reader that is highly customizable due to its built-in script language and uses musical tones as indications of cursor location (Lazarro, 2001). LookOut was used because most participants had used HAL or JAWS; hence their mental model and strategy may be affected by their familiarity to these screen readers.

The experiments were videoed with camera directed at the screen and participants' hands. The video records also incorporated participants' comments as well as the outputs of the speech synthesizer used by the LookOut screen reader (the Microsoft English ASR version 5 Engine).

2.4 Data Collection

Data were collected from experimenter's observation (either directly or by replaying the videotapes) and the comments given by the participants (either during the interaction session or in the post-session interviews). The time participants took to complete the tasks were not measured, because participants often paused during the session to make comments or to answer experimenter's questions.

Perceived usability problems were recorded in real time during the session or in post-session interviews. Later observation of videotapes also revealed some additional problems. Mental models and strategies were identified through post-session interviews and observations of user behaviour.

3 Results and Analysis

3.1 Mental Models of Windows environment

Users' mental models of Windows environment were identified through a straightforward question "What is your mental model of the Windows environment, desktop or application?"

Three participants visualized the Windows environment as a one- or two-dimensional arrangement of icons or symbols associated with applications. Each application is represented by a physical location on the desktop. There were some individual differences in the mental models that participants employed. One participant described the desktop as a place where he can move from one object to object without imagining any geometrical structure. Another participant pictured a strict column and row array of icons. These were categorized as *structural* mental models.

Users with structural mental model might run into problems when the applications are not arranged in a pattern that matches their mental models. Some participants in our study had problem finding a file because they expect a one-dimensional array (hence only used the up and down arrow keys) while the directory was arranged as a two-dimensional array. One user suggested that a good screen reader should always inform users the layout of a desktop.

One participant associated each application in the computer with a user-defined shortcut key (e.g. CTRL+MW = Microsoft word). In other words, their mental model of the Windows environment is a series of functions and commands. Therefore, this mental model is called *functional* mental model.

Users with functional mental model would be less dependent on how the applications were arranged on the desktop because each application/ function would be associated with a shortcut key. Difficulty might arise when these users use an unfamiliar desktop with different shortcut key-to-application mapping. This problem can be alleviated through informing users about available applications on the desktop and their associated shortcut keys and providing wizard to redefine the shortcut key associated with each application to better match users' mental model.

One participant associated part of the Windows environment with shortcut keys and part of it with a structure. For example, to get to Microsoft Word, it would be pressing ALT+P (for program). Microsoft Word is the fourth item from the top of the list. This mental model is simply called *hybrid* model.

Participants were also asked about how they view the relationship between the desktop, MS-Windows applications and the screen reader. Some participants thought of all three as an integral whole, hence they did not have separate mental models for each. Others thought of them as three different entities. It can be argued that participants with structural mental model would think of these three items as separate entities, with the MS-Windows being the placeholder of the icons (which represents the applications) and the screen reader being the vehicle that describes where the icons were in relation to the placeholder. Participants with functional mental model, on the other hand, would be more likely to think of the system as a collection of functions with shortcut keys.

Even though the participants were able to describe their mental models, some said that they did not bother forming mental model of anything the screen reader were not able to describe to them, e.g. a complex diagram/structure. This represents a challenge to software developers to build an application that can describe a complex diagram/structure or a desktop layout to blind users.

To understand the structure of blind users' mental model, a follow up session where the participants of the first session were contacted by phone was performed. There were only four participants in this session. In this follow up session, they were asked to free recall up to 20 word processing functions spontaneously. The participants were not time limited. Table 2 lists the top ten functions in their original order.

	CE	BH	PH	RS
Recalled	20	11	10	10
1	Open file	Cut & paste	Type text	Open file
2	Save	Delete	Select text	Create new file
3	Format paragrap h	Save	Format paragrap h	Close file
4	Format font	Bold	Spell check	Print
5	Spell check	Centre	Search text	Find text
6	Select text	Indent	Move text	Replace text
7	Copy & paste	Jump to	Copy text	Select text
8	Delete text	Search text	Format table	Word count
9	Type text	Spell check	Number	Format style
10	Navigate through text	Replace text	Create table	Bullet

Table 2: Top-ten Word functions recalled.

The functions that the participants can spontaneously list are assumed to be the functions they most often used and most are related to text manipulation (format, cut/copy and paste). The order the functions were listed by our participants is assumed to represent the way they structure these items in their minds. Blind users in our study seemed to have functional mapping rather than structural mapping of menu items (i.e., the items were listed in the order of their functions rather than their physical locations on Microsoft Word). This is true even for participants with structural mental model. Unsurprisingly, the participants did not mention more complicated functions (such as macro or customize).

Although the participants did not seem to have any problem to recall 10-20 functions, one common comment that they made was that it was difficult to list those functions without contextual use. A study by RNIB provided the explanation about why it was difficult to free-recall items without context: sight provides the memory with useful prompts and reminders. Memory is based on the previous receipt of information, and the ability to activate memory from perceived cues. Blind people may need specially created 'prompts' to jog their memory (RNIB, 2002).

3.2 Strategies in Coping with Unfamiliar Windows environment or Applications

Participants adopted somewhat similar strategies when faced with new Windows environment or application. The flow diagram of their common strategy is depicted in Figure 1. This diagram is based on observation of users' action in the experiment as well as users' answers to the post-session interview.

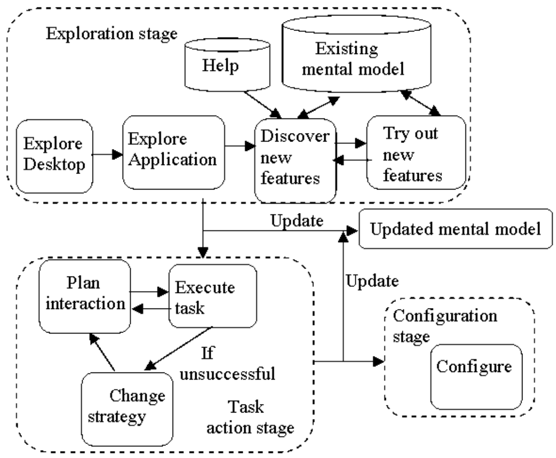

Figure 1: The blind users' mental model of Windows environment.

Blind users in general adopted three stages of strategy in dealing with new system: exploration, task-action and configuration. In the exploration stage, users first explored the desktop to see what applications were available. Depending on their need, they would then start exploring the necessary application. When the application was new, users iterated between discovering and trying out new functions. In this loop, users utilized either system's help or association with existing mental model to help memorizing the features. The result of the exploration stage is an updated mental model. In the next task-action stage, users created another loop of planning how to interact with the system and executing the task. When the planning did not achieve the action they desired, they changed strategy. For example, in this study when participants were asked to create a new file, some pressed the CTRL+N key. However, this key is used by LookOut as a bypass key. In one case, when a participant realized that pressing CTRL+N did not create a new file, he went to the "File" menu and picked up "New" item from the menu. The last stage was the configuration stage, which users only performed after they were comfortable with the application. In this stage they configured the application, e.g. by redefining shortcut keys.

Individual differences were observed in user's strategies. Two participants have a highly procedural strategy, while others are much quicker in changing strategies when their current strategy failed. Three spent more time in performing trial and error experiments, while others immediately asked the experimenter when they got stuck with a problem. One user blamed oneself when the system did not do what he expected, while another user quickly blamed the system.

In terms of asking for help, some participants prefer to seek help from other users from a mailing list for blind users, while others seek help from sighted colleagues. Most participants mentioned that they would only seek outside help as the last resort. Another variation was that some participants prefer Braille manual while others were content with online help and virtual cursor mode as manual.

Some participants also mentioned that very occasionally they face a totally new Windows environment of application. Therefore, there were some expectations that the "new" system would be similar or can be configured with carried-over knowledge of previous system.

3.3 Perceived Usability Problems

The relationship between blind users' mental model of Windows environment, application and screen readers and their perceived usability problems is analysed by observing the difficulties that participants encountered when using LookOut and recording their verbal reports.

Problem	CE	BH	PH	SS	RS
1. Interference with participants' existing knowledge					
a. Interference with standard MS-Windows/ Word shortcut keys	2	2	2	2	2
b. Features missing compared to their current screen reader	2	2	1	1	1
c. Different method compared to their current screen reader	1	2	1	1	1
d. Interference with current screen reader's shortcut keys	1	2	1	1	1
2. Feedback-related problems					
a. No indication when finished performing a task	2	2	2	2	2
b. Poor response time / slow feedback	1	-	1	-	1
c. Window's title not read	-	-	-	-	2
d. First item in a dialog box / menu / task bar not read	-	-	2	-	2
e. Highlighted icon not read	-	-	1	-	1
f. Buttons / keys not read	-	-	1	1	-
g. Items in palette /dialog box not read	-	1	1	1	-
3. Speech- or sound-related problems					
a. Broken/sharp junctions in speech	-	-	-	-	1
b. No punctuation	-	-	-	-	1
4. Memory-related problem					
a. Shortcut key to action mapping hard to remember	2	2	-	-	-

Table 3: Perceived usability problems

The breakdown of the problems by participant is given in Table 3. We did not count the frequency of the comments because of two reasons. First, all participants were experienced computer users; therefore, in most cases they managed to avoid getting into the same problem more than once. Second, because of the short session and limited tasks most problems only appeared once. The number under each participant describes the severity of the problem. A severity rating of 1 means participants encountered and made a remark about the problem but did not stop performing the task and 2 means participants stopped the task specifically to complain about the problem.

3.3.1 Interference with User's existing knowledge

Reviewing the types of usability problems, the most severe type is the interference between the way Lookout works and user's established mental model based on their familiar screen reader (all participants noted it, and many with severity ratings of 2). This finding highlighted the problem suggested by Redish (1994), that it is very difficult to change an established mental model.

The feature participants complained about the most was interference between LookOut shortcut keys and Microsoft Word's keys (e.g. CTRL+N for creating new file in Word and for transferring shortcut key assignments between MS-Windows application and screen reader in LookOut). Because users already had established mental models based on their familiar screen reader, when LookOut did not perform a task in the same way or there were features in their familiar screen reader that they thought were not available in LookOut, participants considered this as a usability problem. The most common example was the shortcut keys: participants often consider CTRL+numeric keypad (which JAWS uses) as the standard method of assigning shortcut keys.

3.3.2 Feedback-related problems

Feedback-related problems were not so severely rated as problems of interference with established mental model. However, one feedback-related problem that bothered all participants was the lack of feedback when the screen reader finished performing a task. Our users have to rely on the audio feedback so they thought the screen reader was still finishing the task (probably because the computer processor was slow, reason one participant) and consequently repeated the task (causing the task to be executed twice) or blamed themselves because they thought they have not performed the right action.

Because participants of the present study were expert users of screen readers, slow feedback from the speech synthesizer was considered a nuisance. Some commented that slow speech might be more useful for novice users; hence controlling verbosity was suggested. Others could not tolerate the slow responsiveness of the application. One participant thought that too many features executed at the same time (e.g. playing the musical tone when moving the cursor up and down and reading the sentences in that particular line) causing the system to slow down.

3.3.3 Speech or sound-related problems

Only one participant complained about the quality of the sound/speech. As this particular participant remarked, he used his computer for programming; therefore, it was very important to hear every word clearly, including commas and periods. It should be noted that the quality of the speech is not a feature of the screen reader. It is a feature of the speech synthesizer being used, which is independent of the screen reader.

3.3.4 Memory-related problems

Some participants complained that the shortcut keys used by LookOut were hard to associate with the actions. This is understandable as they were not used to the application. However, it might also be caused by interference with existing knowledge. That is, participants expect a certain shortcut key to represent a certain action and consider it a usability problem when the combination did not meet their expectation.

Some of the aforementioned problems were intermittent (e.g. in some occasions LookOut didn't read the content of a font palette box or document type dialog box). Some others were caused by incompatibility of LookOut with more advanced features of MS-Windows application or Microsoft Word's configuration problem.

Users' tolerance toward usability problems varied widely. While some users would find an excuse for it ("It is fine because this product is perhaps the least expensive screen reader in the market"), some expressed a strong opinion ("I hate this interference"). Although there was no study suggesting that the degree of tolerance is an indicator of willingness to use a product, it might be induced that the stronger the intolerance is, the less likely a user would use the product in the long run.

Some participants suggested that to thoroughly investigate usability problems, participants should be given more exposure to LookOut. Since often usability testing was performed on one snapshot session rather than longitudinally, it might not capture problems that can be highlighted if users were given more opportunity and time to explore and use the products.

In summary, this study found that when users already had an established mental model derived from their familiar screen readers, users perceived as usability problems any features that work differently from their familiar screen readers. This study is not intended to perform a usability evaluation on, or to improve, the tested screen reader.

3.4 Desired Features of a Screen Reader

The participants were asked to describe what features they would like a screen reader to have. The following is the list of features suggested by participants. The star next to the feature indicated that the feature had already been implemented in LookOut.

1. Useful shortcut keys:
    a) Shortcut key for reading the whole file with a skip function to stop the reading.
    b) Shortcut key for selecting the whole paragraph.
    c) Shortcut key for reading font's style*.
2. Sound-related suggestions:
    a) The ability to turn off musical tone*.
    b) The ability to turn off any sound without unloading the program*.
3. Shortcut key-to-action mapping:
    a) Systematic shortcut key combinations (both semantically and in terms of their physical locations).
    b) The use shortcut keys that don't interfere with MS-Windows' shortcut keys.
    c) Multiple use of the same key for similar functions (e.g. CTRL+4 once to read a character, twice to read a word and three times to read a sentence).
4. Setting flexibility:
    a) Ability to perform global and application-dependent settings (including redefining shortcut keys)*.
    b) Ability to configure for various levels of expertise (i.e. expert and novice modes)*.
5. Information/features to optimize task:
    a) Ability to get information about visual layout of a document (justification, font style, spacing, etc)*.
    b) Ability to provide screen layout description of the desktop.
    c) Providing the content material before the action (e.g. "…[information]…selected" rather than "selected… [information]…")*.
    d) The ability to alter between reading the whole highlighted all words or some words.
    e) Font style's change announcement to prevent users from accidentally altering font size, type or CAPS.
    f) Screen review mode (the ability of reviewing the desktop or the application and telling them what menus were available and what menu items were greyed out)*.
    g) Virtual cursor mode (a mode to visit menu items and physically try out the shortcut keys without actually executing them).
    h) Shorter verification questions (e.g. "Shutdown?" rather than "Do you really want to shut down?").
6. Reliability and product quality:
    a) Automatic loading after MS-Windows application crashes*.
    b) Product upgrade and technical support*.

3.4.1 Useful shortcut keys

As one participant mentioned, blind users are keyboard-based users as compared to sighted users who are mostly mouse-based. Therefore, to be able to finish their tasks more efficiently, shortcut keys are important for blind users. For example, when participants were asked to explore a document to familiarize themselves with how the screen reader works, some of them noted that shortcut keys to read the whole document will be a useful feature.

Similarly, because the tasks in the main session includes changing font style and cutting and pasting the whole paragraph, participants remarked that shortcut keys for cutting/copying a paragraph would be useful.

3.4.2 Sound-related suggestions

For blind computer users, audio provides most if not all of the information they needed to perform a computer task. Therefore, it might be helpful to create a rich sound/speech for various types of information they might need, sometimes several sounds simultaneously if necessary. However, as one participant remarked, when the computer talked or made sound, blind users tend to stop whatever they were doing and listened to the computer. Hence, it is important to provide the audio information only when needed.

Some participants prefer less musical tones for navigation and capital letters. As some participants suggested, it is important to control the verbosity so users can adjust the amount of audio information to their need.

One user also suggested a feature of temporarily disabling the speech without unloading the program. This is necessary because blind users sometimes work in an environment where sighted users might be present, and when sighted user needs to borrow a computer that belongs to blind users for a little while, this feature will be very useful.

3.4.3 *Setting flexibility*

A screen reader has to cope with many different applications and levels of user's expertise and experience. Therefore, it is important to ensure that a screen reader has global and application-dependent settings as well as novice and expert modes.

3.4.4 *Information/features to optimize tasks*

Suggestions in this category ranged from features that will help users learn faster (virtual cursor mode and screen review mode) to a feature that would save users from listening to too many words (e.g. shorter verification question).

Those extra features would enable blind users to have shorter learning curve (because for blind users, there is always a learning curve when faced with something even slightly different from previous application, one participant argued). Those features would also enable shorter task execution because sometimes (expert) users do not need to listen to a complete message/instruction.

Although the suggestions cover a broad range of features, the essence of those suggestions was to enable blind users to compete with sighted users in the information technology market, as one participant emphasized.

3.4.5 *Reliability and product quality*

As with any other product, reliability and product quality were a concern for users. Some participants mentioned the importance of ensuring that users received product upgrade and technical support.

Because, as one user mentioned, MS-Windows application and screen reader were inseparable for blind users, some users find it useful that the screen reader automatically uploads when the MS-Windows was reset after a crash.

4 CONCLUSIONS

Holst (1999) speculated two alternative mental models of blind users:

1. Users do not know what the screen looks like. They usually memorize a large set of options to achieve a given effect.
2. Users have a good cognitive or mental map of the screen.

The present study indicates that our users' mental models are a combination of these two models. However, some of our users possess slightly different models. Each model has its requirements that need to be accommodated by developers of screen readers.

In general, from the present study, there are three main messages for developers of screen readers. Firstly, it is important to perform usability evaluation with prospective users in longitudinal fashion rather than in a snapshot fashion because the number of usability problems highlighted is related to the duration and frequency of use. Secondly, user's established mental model may affect their perceived usability problems when they interact with new system. However, as our users suggested, it is possible to alter their established mental model given enough exposure to the system. The practical implication of this is that it is necessary to provide training for new users to help them re-establish their mental model to match the new system. Finally, thirdly, the study collects several features that some blind users considered essential in a good screen reader (as reported in 3.4) that designers of screen readers can use as a guideline.

5 REFERENCES

Barnicle, K. (2000). Usability Testing with Screen Reading Technology in a Windows environment. *Proceedings of Conference on Universal Usability CUU'00,* ACM Press, pp. 102-109.

Bornemann-Jeske, B. (1996). Usability tests on computer access devices for the blind and visually impaired, in J. Klaus, E. Auff, W. Kremser, W.L. Zagler (Eds.) *Interdisciplinary Aspects on Computers Helping People with Special Needs, Proceedings of the 5th International Conference ICCHP'96*, OCG Press, pp. 139-147.

Holst, A. (1999). Making the Windows environment Touchable - Feedback that Really Makes a Difference! *Proceedings of 14th Annual International Conference: "Technology and Persons with Disabilities"* Available at: http://www.jsrd.or.jp/dinf_us/csun_99/session0108.html.

Landau, S. (1999). Tactile Graphics and Strategies for Non-Visual Seeing. *Threshold 19*, 78-82.

Lazarro, J. (2001). ZoomText. *Byte, April*. Available at: http://www.byte.com/documents/s=591/byt20010404s0001/index4.htm.

Nielsen, J. (2001). Beyond Accessibility: Treating Users with Disabilities as People. *Alertbox, November 11.* Available at: http://useit.com/alertbox/20011111.html.

Norman, D.A. (1988). *The psychology of everyday things*. Basic Books, New York, NY, USA.

RNIB. (2002). Understanding and using sight: issues for work with people with severe disabilities. Available at: http://www.rnib.org.uk/multdis/ under.htm.

Redish, E. F. (1994). Implications of Cognitive Studies for Teaching Physics. *American Journal of Physics, 62*, 796-803.