Blind Users' Mental Model of Web Page Using Touch Screen Augmented With Audio Feedback

Ahmad Hisham Zainal Abidin[1, 2], Hong Xie[1], Kok Wai Wong[1]

[1] School of Information Technology
Murdoch University
90 South Street, Murdoch
Western Australia 6150
h.xie@murdoch.edu.au, k.wong@murdoch.edu.au

[2] College of Arts & Sciences
School of Multimedia and Technology Communication
Universiti Utara Malaysia
06010 Sintok, Kedah, Malaysia
hishamza@uum.edu.my

***Abstract*- The purpose of this study is to discover in detail the two-dimensional mental model created by the blind people using touch screen with audio feedback. It is hypothesized that by augmenting a touch screen display with audio feedback, blind user would be able to have two-dimensional perspective in their mental model. If their mental model of a web page is two-dimensional, the blind people should be able to gain an overview of the web pages accurately, which is important for navigation of the web page.**

This paper discusses an on-going study to elicit mental models from the blind users using diagrammatic representation. The study so far revealed that there is a significant difference between performances using screen-reader program and touch-screen display with audio feedback. However, the performance seems to be affected by page complexity.

I. INTRODUCTION

Most web pages are designed with some kind of two dimensional structure or layout. For the sighted people, these two dimensional structures are usually far more effective in conveying information. On the other hand, for screen-reader users, there is only one-dimensional string of fragments of the original content. The lack of the two dimensional information in the mental models of the screen reader users is the main obstacle for them to use the Internet effectively.

We believe that by accessing web pages using a touch screen display with audio feedback, a blind user would be able to gain two-dimensional perspective of the web page in his or her mental model. If the mental model of the page contains two-dimensional information, it would help the blind user gain a better orientation of the web page and have a sense of position in the web page. This is important for the user to navigate the web.

The remainder of the paper is structured as follows: Section II discusses mental model theories and the methodology to investigate user's mental model. Section III puts the work in context by discussing related researches conducted with the blind people. Section IV discusses the design of the study to generate mental model data. Section V presents the result of diagrammatic representation's study. Finally, Section VI presents the discussions and conclusions.

II. BACKGROUND LITERATURE

According to mental model theory, user performance is guided by the user's mental model [1]. A mental model in this context is defined as the conception that a human being developed internally to describe the location, function and structure of objects and phenomena in computer systems [2]. A mental model can incorporate information about temporal, spatial, causal, person and object related features of a particular event [3]. Mental model is not restricted to a specific modality [4]. This means that the mental model can be constructed not only from visual experiences but also from other senses including hearing and touching. The mental models will help to shape users behavior and approach to solving the problems and carrying out tasks.

Mental model has been studied initially by psychologists. The term mental model was first mentioned by Kenneth Craik in 1943. In his book "The Nature of Explanation", Craik claims that human mind constructs "small-scale models" of the reality that will be used to predict similar future events. The same terminology has been used by cognitive scientist as part of their effort to know how humans learn and make decisions.

In 1980s, Gentner and Stevens [5] made an effort to compile related studies in mental model. In the book edited by Gentner and Stevens, there are two earliest papers that discuss specifically about mental models of computer systems [6]. The first paper was by Norman [1] that introduced four types of representations which influence user-system interaction. The first type is the target system which is referring to the system that the person is using. The second is the conceptual model that is invented by the designers and engineers which is an appropriate model for the users to develop. To develop this model, it requires the inventor to engage in analysis and decision-making concerning the viable conceptualization model that will be presented to the users. The third is the mental models that people formulate internally through interaction with the target system. According to Norman, the user's mental models continue to modify in order to get workable results. Finally, the scientist's conceptualization, which is referring to the model of a model, describes the content and structure of the user's mental model as understood by the scientist.

Several observations by Norman revealed that mental models are incomplete and also unstable. People easily forget if they are not using the system for some time. Furthermore, there are no firm boundaries which make people easily get confused with similar operations. Nevertheless, mental models are application or problem specific and are not static [7]. Although there are many shortcomings, mental models are still important as a source of information, providing predictive and explanatory power in understanding the interaction [1].

In terms of modelling the mental models, Norman believes that using only verbal protocol is incomplete. He believes that what people believe and how people act are two different manners. Using verbal protocol, people usually compel to give a reason although they do not have one.

The second paper that discusses mental models of computer systems was by Young [8]. In his paper, Young uses different terminology to describe mental model which is User's Conceptual Model (UCM). If Norman focuses on the users' own views of the world, Young focuses on the conceptual model of the system. Young defines UCM as representation or metaphor that users adopt to guide their actions and help them interpret the device's behavior. This means the users need to adopt certain metaphor or analogy to guide their actions and help them interpret the device's behavior. Similar to Norman, Young suggests that there are four agents of UCM, in which he believes that each agent has their own UCM; for User, Designer, Psychologist (Scientist) and the Device (System).

Other than these two papers, Sasse [6] in her PhD work has come out with comprehensive review about mental models that covers psychology, cognitive science and human-computer interaction disciplines.

Based on the literature, there are two ways to elicit the mental models. The first is by imposing appropriate models during training and trying to elicit users' mental model after they are exposed to the system. Another way is by just eliciting the mental model without prior training because some applications such as the Web is too complex to build a normative model for comparison [9].

Sasse evaluated different experimental scenarios for investigating user's mental model:

- **Task Scenario –** hands on
- **Teaching Back Scenario**
- **Advance Task Scenario** – verbal and hands on
- **Joint Exploration Scenario**

The first scenario which is Task Scenario requires users to perform a set of tasks using word processing application to derive performance data. The data which can be gathered under performance aspect are task completion, time to complete the task given and error. During the experiments, the users are prompted by the investigator to think aloud during the completion of tasks. Think aloud protocol encourages participants to say whatever they are looking at, thinking, doing, and feeling, as they perform their task. However, user performance alone cannot be taken as an indicator of a specific users' mental model [11].

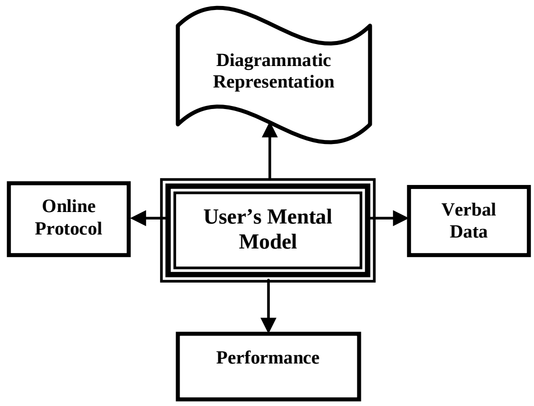

Figure 1. Data for Mental Model's Study

In the Teaching Back Scenario, the user is required to teach another person to use the application. This is to encourage users to verbalize their knowledge. However, this scenario is suitable if the users have been exposed to the system for a certain period prior to the experiment. In addition, the users seem to be dependent to the "learner" (investigator) when they face difficulties.

The third scenario is similar to the first one except that it requires the user to answer several questions related to the application prior to the hands on task. Then, the user has to describe about the application before the experiment ends.

Another experimental scenario is Joint Exploration. Using this method, the user is required to work together with another person and explore the application. It is found that users are more involved and communicative. The disadvantage of this scenario is the less experienced users will be more passive since they lack confidence and are still expecting for guidance. Other than that, this method requires a lot of time for transcribing and analysis since the data gathered can only be analysed from recordings.

Sasse evaluated the scenarios in terms of:

- resources required to implement a scenario;
- amount of data on users' models elicited;
- quality of data on users' models elicited;
- effort required to analyze users' models.

Sasse also has discussed three types of data collection for mental model studies:

- **Performance data -** typically assesses user's task completion, time taken to complete the tasks and number of errors
- **Online protocols** - record user and system actions
- **Verbal data** - requires users to verbalize during the interaction with the system

Sasse suggests alternative ways to elicit the mental models from the user including drawings, diagrammatic representations and schema-like representations. Schema-like representation is the extension description of the existing model including functions, components and properties of the extracted model. While the subjects participated in the study were sighted people, the current study will involve participants accessing the web pages without visual ability. The data gathering procedure should be modified to suit the participants. Obviously, it is not suitable to elicit mental model from the blind people using drawing. Figure 1 shows four types of data collection for mental model studies discussed above.

III. BLIND PEOPLE'S MENTAL MODEL

It is found that at the moment there is still little work on mental model for blind people. Study by Murphy et al. [5] was conducted to understand the challenges faced by the blind users when accessing the Web using assistive technology. In terms of mental model, the study found that blind people attempt to remember the page structure when they scan for the desired information using screen-reader program. The study describes the mental model of the screen-reader users as a "vertical list", and they perceive all web pages to be in a single column-like structure. This happens because screen-reader program processes pages and produces output to the blind people in a sequential order, from top left of the page to the bottom right of the page.

As a result, the screen-reader users will try to remember the sequence of the interested items. It is a burden to mental load if screen-reader users have to memorize a large set of information and options. Furthermore, for wide-ranging and complex websites, the "vertical list" will be so extensive that it is impossible to be remembered. In addition, by collapsing the two–dimensional web page into a single list of items, many useful navigational hints would be lost, for example, the position of an item relative to other items (eg, on the left, on the right). Due to the above factors, the screen-reader users spent more time and needed more efforts to perform a task on the Internet [4] in comparison to their sighted counterparts [15].

Another work on the mental model of the blind was by Takagi et al. [3]. The focus of their study was to observe the interaction and strategies adopted by the blind when accessing online shopping web sites. Takagi et al. observed that the mental model used by screen-reader users to access the online shopping web sites is a vertical list. They found that the blind users had their own strategy to speed up the searching process. They also found that the screen-reader users focused on landing in the main content area by using gambling scanning or exhaustive scanning.

Kurniawan and Sutcliffe [16] investigated blind people's mental models in dealing with a new Windows-based screen-reader program through interviews. Online protocol (video recording) was also used to observe participant's interaction. One of the important remarks from the study is that the participants did not bother forming a mental model of a complex diagram/structure or a desktop layout because the screen-reader does not have the capability to describe it to them.

IV. THE STUDY

The purpose of this study is to investigate differences of the mental models created by the blind people from a two-dimensional web page using two different means: one using a touch screen display with audio feedback and the other using a screen-reader only. We believe that by accessing web pages using a touch screen display with audio feedback, a blind user would be able to have two-dimensional perspective of the web page in his or her mental model. If the mental model of a page contains two-dimensional information, it would help the blind user gain a better orientation of the web page. This is important for the user to navigate the web.

Task scenario experiments were carried out which involve blind participants accessing the experimental pages using a screen-reader only and the other experiment using a touch screen display augmented with audio feedback. The participants were asked to perform specific tasks. Think-aloud protocol with modified version dedicated for screen-reader users was adopted in this study [10]. The participants' interactions with the experimental pages were observed and recorded for later analysis.

From the literature, it is found that previous researches on mental model of the blind people focus on performance data, verbal protocol and on-line protocol. However, this study extends the previous works with another protocol which is diagrammatic representation. Foam blocks with different types of surfaces (rough and smooth) were used to represent different web elements (Figure 2). However, only two elements were asked; headings (rough surface blocks) and data (smooth surface blocks).

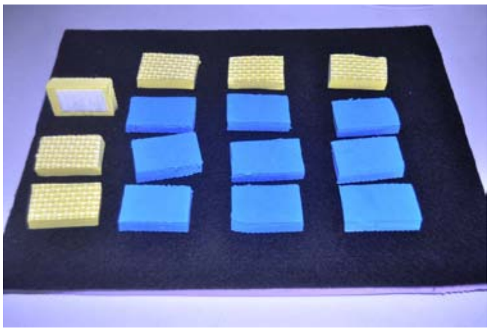

Figure 2. Diagrammatic Representation using Foam Blocks

This project is assisted by The Blind Citizen Australia (BCA) in Western Australia, the Association for the Blind WA and also by Malaysia Association for the Blind (MAB). Eleven blind people participated in this experiment (4 females, 7 males). However, this study has to exclude a low vision impairment participant (6% remaining vision on her left eye) because she chose to use her magnifying pendant during the experiment. The number of participants is small, however considering the difficulty in recruiting blind participants and compared to similar studies, the size of reasonable. For instance, Takagi et al. reported that most experiments conducted with blind participants usually have only four to six people [11]. Mobility problems and specific criteria needed for particular experiment were some of the reasons. Call for participation in this study was distributed through the blind association's mailing list. Interested participants have to contact the researchers by email or phone to discuss suitable date and time. The researchers focused on recruiting blind people with some basic skill using JAWS screen-reader. The participants' ability to use computer ranged from fair to expert.

The study followed the experiment protocol as below:

1) Explanation about the study.
2) Consent to participate. All consents were given orally by the participant and audio-recorded for future reference.
3) The participants were trained about the procedure of the experiment using two training pages for about 10-15 minutes each. This is to make sure that any issues or doubt should be resolved before the real experiment started.
4) The experiment started by asking the participant to explore the experimental pages using JAWS screen-reader. There are altogether 6 experimental pages with different layout complexity used in the study. Layout complexity was determined objectively, based on whether the layout has even or uneven table size. Layout with even table size is associated with lower complexity and vice versa.
5) After exploring each page, the participants were asked to briefly describe about the page.
6) Then, the participants were asked to construct the layout of the table using diagrammatic representation.
7) After the construction of diagrammatic representation, the participants were asked to answer a question.
8) Spatial Ability Test.
9) Repeat steps 4 - 7 above but now using touch screen display augmented with audio feedback. This study chooses an Apple iPad as the experiment's apparatus. VoiceOver program is available as screen-reading software for Apple iPad. The participants were trained on how to use the device prior the real experiment. Two types of gestures were trained; "gliding" and flicking.
10) Demographic survey.

This study adopted within-subject crossover design. This means that each user performed in both experiments. To control learning effect, half of the participants accessed the experimental pages using touch-screen with audio feedback first, and then the screen-reader only and the other half of the participants conducted the two experiments in the opposite order.

Each session normally lasted 2-3 hours and the whole session was recorded using video and audio recorder.

V. RESULTS

None of the participants have any experience in using Apple iPad but few of them have used Apple iPod or iPhone before. The experiment was the first time for all of them to use the Apple iPad. On the other hand, all participants were familiar with JAWS screen-reader program in their daily life when using personal computer. In terms of their ability in using Assistive Technology, three participants are described as expert, while the other seven are described as good.

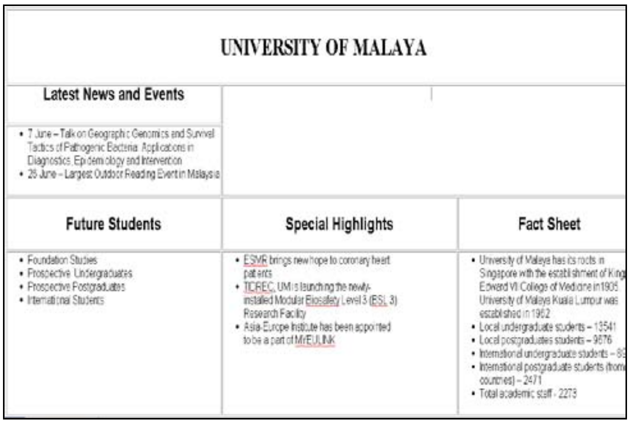

Figure 3. Page 3

A. *Diagrammatic representation for complex layout*

Figure 3 shows Page 3, an experimental page with higher complexity. Figure 4(a) shows all diagrammatic representations constructed by the participants for Page 3. Table 1 below shows the percentage of correct arrangements of headings and cells for Page 3. The average percentage of correct arrangements of headings and cells for Page 3 using JAWS is 38.85 and slightly lower for the iPad which is 38.70.

A Wilcoxon Signed-Ranks Test shows no significant differences between the performance when using screen-reader program and when using touch-screen display with audio feedback to access Page 3 (Sig. value = 1.000).

TABLE 1
PERCENTAGE OF THE CORRECT ARRANGEMENTS OF HEADINGS AND CELLS FOR PAGE 3

	Page 3	
	JAWS	**iPad**
	33.3	100.0
	55.5	0.0
	77.7	100.0
	66.6	44.4
	11.1	0.0
	33.3	0.0
	33.3	22.2
	44.4	88.8
	33.3	33.3
	11.1	0.0
AVERAGE	38.85	38.70
STD DEV	23.54	42.62

B. *Diagrammatic representation for simple layout*

Figure 5 shows Page 6, an experimental page with lower complexity. Figure 4(b) shows all diagrammatic representations constructed by the participants for Page 6. Table 2 shows the percentage of correct arrangements of headings and cells for Page 6 (P6). The average percentage of correct arrangements of headings and cells for Page 6 using JAWS is 43.50 and using the iPad the participants were able to construct better arrangements with an average of 66.47.

A Wilcoxon Signed-Ranks Test for data in Table 2 shows significant differences between the performance using screen-reader program and touch-screen display with audio feedback to access Page 6. (Sig. value = 0.041).

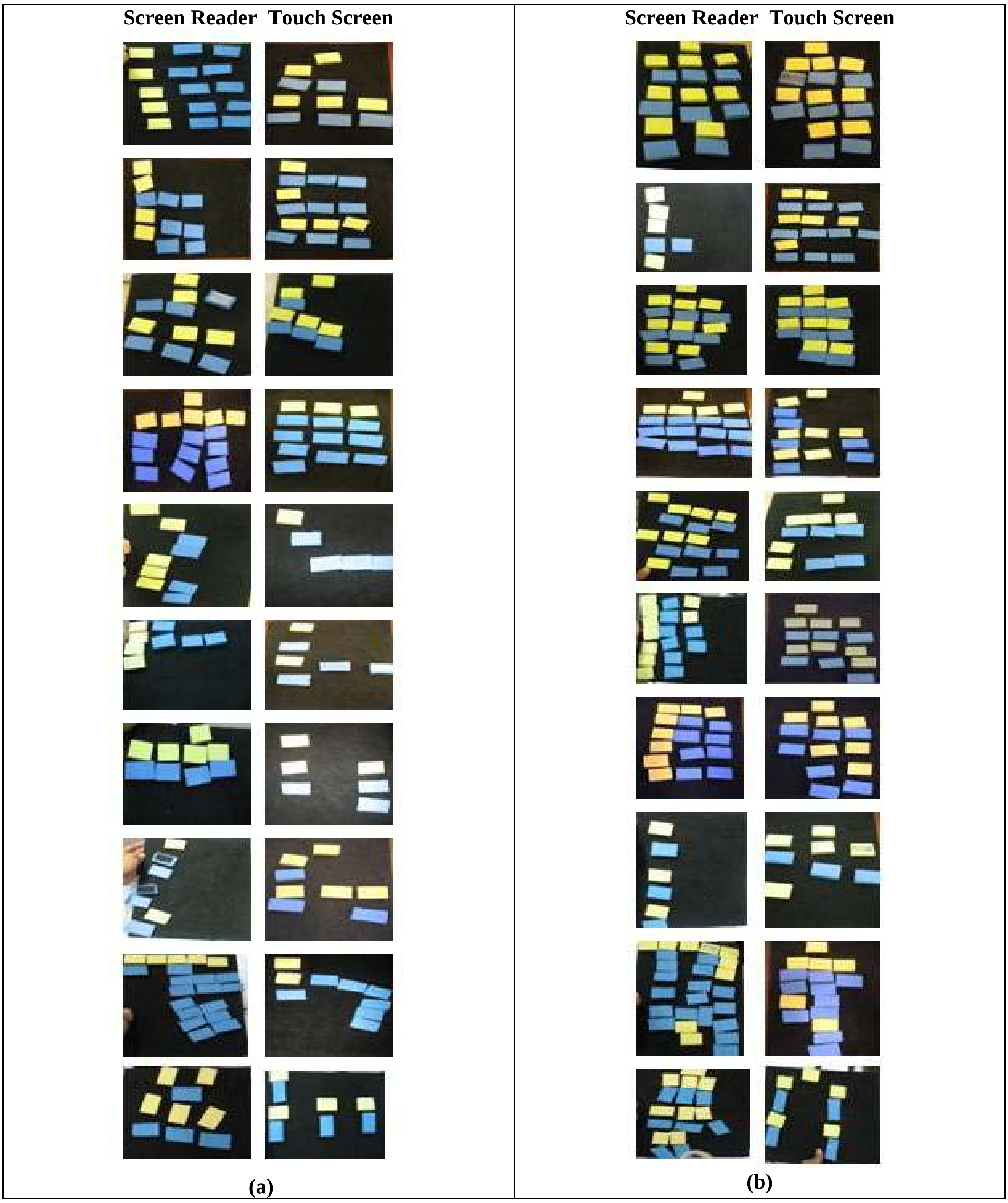

Figure 4. Diagrammatic Representations

VI. DISCUSSIONS AND CONCLUSIONS

From the diagrammatic representations, the study confirmed previous findings that the mental models of the screen-reader users are in a single column-like structure. However, expert and more experienced screen-reader users were able to imagine the layout in two-dimensional perspective although it was not 100% correct.

The study revealed that there is a significant difference between performances using screen-reader program and touch-screen display with audio feedback for simple layout page. Using touch screen display with audio feedback, the blind users were able to get higher percentage of the correct arrangements of headings and cells as compared to using screen-reader program.

Figure 5. Simple Experimental Page (P6)

TABLE 2
PERCENTAGE OF THE CORRECT ARRANGEMENTS OF HEADINGS AND CELLS FOR PAGE 6

	Page 6	
	JAWS	**iPad**
	88.2	100
	5.8	76.5
	88.2	100
	58.8	52.9
	35.3	41.2
	23.5	70.6
	29.4	58.8
	0.0	47.1
	17.6	64.7
	88.2	52.9
AVERAGE	43.50	66.47
STD DEV	34.76	20.57

On the other hand, for complex layout, there is no significant difference between performances using screen-reader program and touch-screen display with audio feedback.

This study has some limitations. The training sessions were too short for the participants to really familiarize the new gestures using VoiceOver program in the Apple iPad. On the other hand, all participants were familiar with JAWS screen-reader program and actively use the technology in their daily life. Moreover, the participants were encouraged to utilize “gliding” gesture to be used together with flicking gesture when using touch screen display with audio feedback. This was totally a new concept to all of them and may affect their performance in using the technology. The future work should consider longer training session to really expose participants to the technology.

As a conclusion, using touch screen display with audio feedback, the blind people’s mental model of a page is in two-dimensional perspective. Because of that, they are able to gain the overview of the web page accurately. This may help them to navigate the page effectively. However, the accuracy of the overview seems to be affected by page complexity.

REFERENCES

[1] D. A. Norman, *Some Observations on Mental Models*. Hillsdale, New Jersey, London: Lawrence Erlbaum Associates, 1983.

[2] D. H. Jonassen, and Henning, P., "Mental models: knowledge in the head and knowledge in the world," in *International Conference on Learning Sciences*, Evanston, Illinois, 1996, pp. 433 - 438.

[3] M. L. Noordzij, Zuidhoek, S., and Postma, A., "The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions," *Cognition*, vol. 100, pp. 321 - 342, 2006.

[4] V. Gottschling, "Visual Imagery, Mental Models, and Reasoning," *Mental Models and the Mind*, vol. 138, pp. 211 - 235, 2006.

[5] D. Gentner, and Stevens, A.L., Ed., *Mental Models*. Hillsdale, New Jersey, London: Lawrence Erlbaum Associates, 1983.

[6] M. A. Sasse, "Eliciting and Describing Users’ Models of Computer Systems," PhD, Computer Science, University of Birmingham, 1997.

[7] H. Vieritz, Pfeiffer, O, and Jeschke, S., "BeLearning: Designing Accessible eLearning Applications," in *37th ASEE/IEEE Frontiers in Education Conference*, Milwaukee, WI, 2007, p. 6 pages.

[8] R. M. Young, *Surrogates and Mappings: Two Kinds of Conceptual Models for Interactive Devices*. Hillsdale, New Jersey, London: Lawrence Erlbaum Associates, 1983.

[9] Y. Zhang, "The influence of mental models on undergraduate students' searching behavior on the Web," *Information Processing and Management*, vol. 44, pp. 1330-1345, 2007.

[10] P. Strain, Shaikh, A.D., and Boardman, R., "Thinking but not seeing: Think-aloud for non-sighted users," presented at the CHI2007, San Jose, California, USA, 2007.

[11] H. S. Takagi, S. Fukuda, K. and Asakawa, C., "Analysis of Navigability of Web Applications for Improving Blind Usability," *ACM Transaction on Computer-Human Interaction*, vol. 14, p. Article 13 (36), September 2006 2007.