

Blind User Visualization and Interaction in Touch Screen: A Designer Perspective

Abdelrahman H. Hussein

College of Computer Science and Engineering
University of Hail
Hail, Saudi Arabia

Abstract— In this paper, we describe how blind students views external system using an image map as a case study. We proposed two interaction techniques which allow blind students to discover different parts of the system by interacting with a touch screen interface. An evaluation of our techniques reveals that 1) building an internal visualization, interaction technique and metadata of the external structure plays a vital role 2) blind students prefer the system to be designed based upon their behavioural model to easily access and build the visualization on their own and 3) to be an exact replica of visualization, the metadata of the internal visualization is to be provided either through audio cue or domain expert (educator). Participants who used touch screen are novice users, but they have enough experience on desktop computers using screen readers. The implications of this study to answer the research questions are discussed.

Keywords- Blind; Visualization; Touch Screen; Accessibility; Usability; Image Map.

I. INTRODUCTION

Interaction with touch screen is a complex process for blind users. Understanding the internal visualization of cognitive activities is an important process for a designer to design external visualization. Scaife and Rogers (1996) pointed out that more emphasis should be provided by “the cognitive activities when interaction takes place with external visualization, the properties of the external and internal structures and their benefits with different visual representation” [14].

Under these circumstances, it is essential to understand internal visualization, their relation with external visualization and how physical activities help to bind these two visualizations. Without prior knowledge of these abstract concepts, it will be intricate for a novice designer to build external visualization. As a result, it is significant to make cognition a research agenda for building information visualization [8].

To address this research agenda, we propose the following research questions in accordance with mobile learning (m-learning): 1) What is meant by “internal visualization” for a blind user in terms of image map?; 2) For a given external visualization, how it can be related to internal visualization?; and, 3) How do physical activities relate external visualization

with internal visualization? This paper is derived to discuss the answer for the above mentioned research questions.

II. INTERNAL VISUALIZATION

The researchers suggest that blind people have an internal representation for the information they hear. Most current research is on geometrical shapes and mathematical symbols, and has not touched the core representation of “abstract” information. In particular, the information visualization does not vary much within the sighted user. On the other hand, visualization varies widely among blind users depending on the description of narrator, comprehension and prior experience. Analysing the mental model in the literature is the first step to investigate internal representation which may yield an effective theoretical concept.

1.1 Mental Models of Blind

A review of the relevant literature suggests that the mental model is an internal representation of real world phenomena, which is composed of many small-scale representations. Craik (1943) claims that mental model predicts future action by constructing small-scale models, in his book “The Nature of Explanation” [3]. The Rouse and Morris (1986) argue that the application of a specific mental model is necessary [11]. The emergence of mental models has led to a decline in the theoretical concept to HCI [12].

The literature also suggests that there are two potential and influential mental models: the Norman model and the Johnson-Laird model (see Fig. 1).

Both models suggest that there are four agents which influence the interaction between the system and user: 1) The target system which refers to the device the user is intended to use.; 2) A conceptual model developed by designers in which the system is developed; 3) The user model developed by the user through interaction with the system. This user model continues to update whenever the user is exposed to the system; and 4) The model for the user model, as understood by the scientist (Psychologist or Usability Expert) [10].

Norman revealed that mental models are not usually accurate and are highly “volatile”. Nevertheless, they are

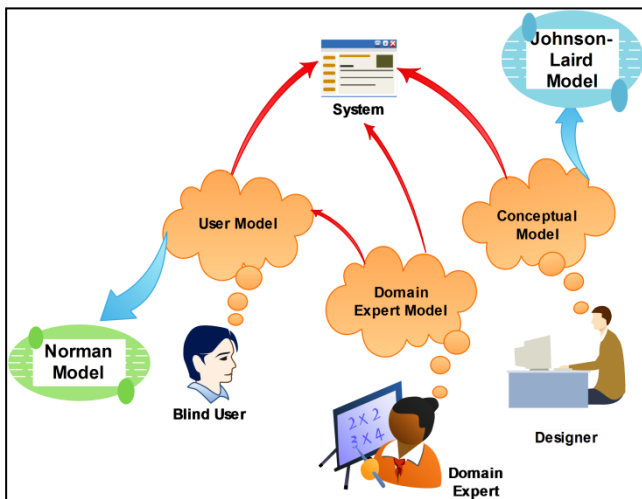


Fig. 1. Mental Model

“runnable” to serve certain rationale. Although there are many lapses, it provides prognostic and expounding power in understanding the interaction.

The Norman model focuses on the user model while Johnson-Laird (1983) focuses on the conceptual model of the system [6]. To put it more simply, Norman emphasizes the “behavioural” aspects of the user towards the system whereas Johnson-Laird defines the “Structural” aspect of the system.

Johnson-Laird views mental models as preserved entities and analogous of what they represent in the external visualization. The user can manipulate the system and configure his mental model with presupposing mental logic and formal rules. However, Norman believes that there are chances the user can pretend to behave with the system. As a consequence, users are forced to give a reason through verbal protocol, although they do not have one.

The Johnson-Laird and Norman models vary widely based on the emphasis and the nature of the problems being investigated. For instance, consider the premises “Tiger eats deer” and “Tiger eats rabbit”. The user can conceptualize two different models where the relations between the entities are present as below:

Table 1 – User mental model and their premises

Mental Model	Premises1	Premises2	Premises3
M1	Tiger	Deer	Rabbit
M2	Tiger	Rabbit	Deer

Both models are coherent with the premises; however, the user cannot predict whether deer, eats rabbit or vice versa by merely exploring their mental models. The user cannot apply logical rules to infer, since it does not store premises based on logical predicates such as eat (Tiger, Deer).

1.2 Significance and Application of Information Visualization

To understand the problem related to interface design, it is imperative for designers to understand: 1) how the blind user imagines the external visualization (Section 2.2.1); 2) which tools are used to stimulate these images (Section 2.2.2); 3) the role of data in image formation (Section 2.2.2); and, 4) how the data are mapped together to form an image (Section 2.2.3). To address these questions, we will discuss this in the following section.

1.2.1 Image or Model

Although the items are scattered throughout the web page, the blind user considers all the items in a web page to be a vertical list [9]. The desktop screen reader processes all pages, and produces output in a sequential order that can be navigated by tab or up/down keys.

Currently, touch screen technologies such as the iPhone and Android implement a static layout for the interaction [4]. With reference to this, two dimensional pages are collapsed to single dimensional to form a single horizontal list that contains a large set of items. It is burdened for the user to memorize the sequence of interest items. Furthermore, the relative position of item such as on the top or on the bottom is lost [1]. Although it has limitations, it is considered to be better than nothing.

Applying the interface design format such as size, position and color to the mental model, Johnson-Laird contends that mental models are essentially spatial representations and are more abstract. This inspection is in distinction with mental imagery. He emphasized that both the mental imagery and mental model can be used in logical analysis.

Such characterization makes it challenging to apply to the interface design, especially for blind users for two reasons: 1) The imagery varies widely between the sighted and blind user; and, 2) the designer cannot understand the imagery of the blind user. For instance, the designer can understand how the sighted user can imagine the size and shape of the checklist, but it is difficult to predict a blind user’s imagination about the same fact.

In interface design, the blind user can understand the spatial layout of the screen based on the training. Instead, it seems unreasonable to provide audio cues related the color and length of the widget. Our own experience of developing courseware for the blind user indicates that a blind user has to be provided with cues about the position of the widget in the touch screen device. By continuous exploration, the blind user becomes familiar with the position of the widget [5].

1.2.2 Mental Simulation

The mental model is abstract and cannot be evaluated directly. The review of literature suggests that there are two types to simulate the mental models. In the first type, training the user to use the system and simulate the mental model. In

the second type, the user is exposed to complex system without proper training and simulates the mental model [16]. Two factors are used to simulate the mental models: tools and data.

Tools: Different traditional tools can be applied to simulate the mental model for blind users [13]: 1) **Task-based scenario:** It requires users to perform a series of tasks to achieve the target. Generally, task acquisition time, number of clicks and error rate of the task performance are gathered. The user is prompted to think aloud during the experiment. It helps the investigator to extract the reason for performed errors; 2) **Verbal and Hands on Scenario:** It includes a task-based scenario, but it requires the user to respond to a query pertaining to the task performed with the system. Finally, toward the end of the experiment, the user has to explain the system; 3) **Tutoring Scenario:** The user is required to tutor about the system to another person. It is suitable if the user has enough experience to handle the system. Furthermore, the user has to depend on the learner. However, it encourages the users to articulate their knowledge; and 4) **Exploration Scenario:** Under this scenario, the user is involved with another person and investigates the application. In this way, the users were communicative and more involved. The main impediment to this type is a highly experienced user who dominates the session and less skilled users who will be controlled due to deficiencies in exposure. Additionally, it consumes much time to gather the data which is extracted through recordings.

Data: Mental models do not possess any data. The information about the external visualization is stored in the form of data which we termed as external data. The external data is used to simulate the mental model. The external data may be constructive data or passive data. Constructive data is data that is transferred from external visualization of the user in the form of audio cues to stimulate the next course of action. On the other hand, passive data is the acknowledgment sent by the system to users in the form of haptic feedback. Along these lines, data are transmitted to humans through the hands and ears of the human body. The transmission of data from external visualization of the user simulates the mental model by analyzing the data, selecting the task and choosing appropriate interaction techniques to accomplish the task. For instance, data named "Enter your name" invokes the user to search for the required information (name) in the internal data, select the data, and deliver the data using a voice synthesizer interaction technique. Finally, item level information is stored in the form of data. The information may be individual, such as "Click here" or may be aggregate information such as "error code".

1.2.3 Mental Map

The mental map is the internal representation of data in the mental model, the available data are mapped to each based on interaction with the system. The mental map is primarily related to item-level information. According to the schematic and semantic level, data are aggregated and the relevant task

is performed to construct and simulate the mental model for different visualizations to get a feel of interface design. Thus, the mental model is more abstract than the mental maps.

III. DYNAMICS BETWEEN MENTAL MODEL AND EXTERNAL VISUALIZATION

The specification for e-assessment has to coexist with standard HCI usability (content usable, easy to use, ease to learn, intuitive). The usability requirement for both blind and sighted users is needed on assessment resources for each assessment type and assessment method.

As discussed earlier, the blind user who uses the latest technologies, such as the iPhone and Android, views the external visualization as a horizontal list. Whatever the layout design is, a blind user navigates in the form of a queue keeping the layout static. The navigation is achieved through flung gesture, which supports both back and forward directions.

In Fig. 2 the dynamic interaction that we devised, the blind user views the external visualization as dots in Braille located at fixed locations in the touch screen interaction. For each blind user, each dot acts as a stack in which items pile on top of each other. In this manner, the blind user can navigate through the items in the stack by click gesture. Since many stacks can be placed on the screen, many items can be accommodated which can be reached easily as compared to a static layout. The only predicament the blind user faces will be identification of the required stack. Navigation is supported by external data, such as audio feedback.

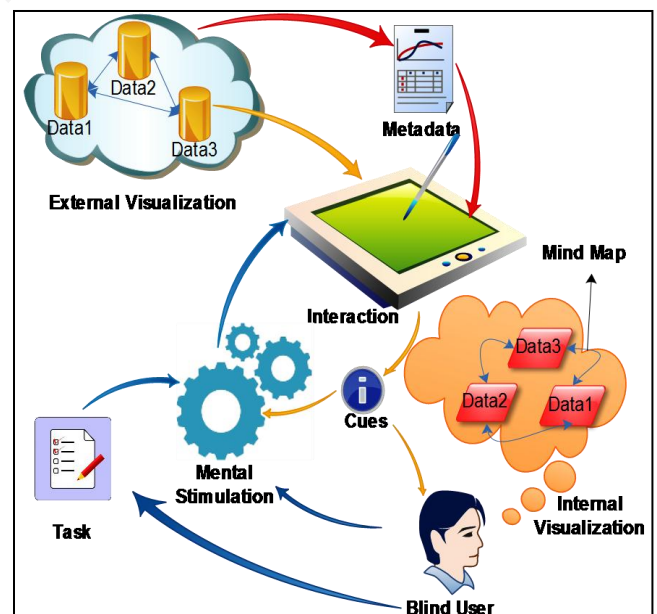


Fig. 2 – Dynamics between the mental models and external visualization.

From this point of view, the blind user does not have to hold the exact external visualization in their heads [2]. Since the internal visualization (mental model) does not contain item specific information about the data, internal visualization is simulated by using carefully designed external data in the form of audio cues.

However, in certain cases, the internal visualization should be an exact replica of external visualization. For instance, in the image map, it is imperative to understand the position of each section. In addition, there are limitations in visualizing the parts as either stack or queue. Thus, to understand the external visualization or to expand the internal visualization, focus should be on interaction and not on the brain. In specific, the events performed by different parts of the body, such as hand and ear, which interact with external visualization to construct, manipulate and stimulate the mental model have to be understood.

IV. INTERACTION

The interaction technique is the technique used by the user to communicate with external visualization. It is characterized by 1) initiating physical action, and 2) following alterations in the visualization state. Tufte (1997) formulates around 57 interaction techniques to be used in the environment [15]. However, most of the interaction is ignored in HCI research since there are no appropriate features to be implemented with it. Thus, many features are aggregated to a single group.

Generally, interaction in touch screen technologies is classified into tactile and non-tactile interaction. Whereas clicking the button and holding the device is considered as tactile interaction, listening to audio is considered as non-tactile interaction. It can also be generalized based on characterization of visualization (static or dynamic) and based on modality (such as hand or ear interaction). HCI research reveals that human events are often ignored, which results in usability problems.

The main aspects of human events to consider regarding user interaction can be:

Project: Project is the special kind of action performed due to mental stimulation accompanied with keenness by shrinking the ear to concentrate. Krish (2009) observed that people give special attention to things which are augmented and projected [7]. The blind user is not able to project visual structure if the structure is visually projected by color encoding or circling. The project is a highly complicated task to project verbally through audio cues. It primarily depends on the intelligence of Text-to-Speech synthesizer and comprehension by the blind user. Although metrics are available to determine the intelligence and comprehension, context awareness is vital to understand the projection.

Select: The external representation is generated based on an algorithm in information visualization. The blind user creates a mental model of location while tutoring the system and subsequently modifies the model based on frequent

exposure and prior experience. The audio cues inform target name followed by the hand movements to reach the target.

Precision: An unexpected state of external visualization occurs if the interaction is not precise. Spatial factors such as cell padding, cell spacing and other related factors have an effect on the precision of hitting the target. The torque force and angle of inclination and recurrence of interaction also plays a vital role in hitting the target precisely.

Coupling: Occasionally, coupling of many actions gives rise to new events. It primarily occurs in multimodal interaction. For instance, a single click informs the target name through audio cues. On the other hand, if the blind user accidentally presses single click twice (equivalent to double click) this will invoke another event to open the item. Thus, the user has to get confirmation before proceeding to the next level to avert the unexpected event's happening.

Investigation: Investigation is a daunting task, especially for blind users if they don't have enough information at hand or any aural cue to reach the desired page. In this case, the blind user applies their own strategy in an ad hoc way to get the required result. When the exploration is new, the blind user uses an iteration process between discovered and new items. While in the investigation process, the blind user uses system help, domain expert or a sighted user and an existing related mental model if any, then the result will be the formulation of the new state of visualization. When the investigation does not achieve the required result, they alter the strategy.

Configuration: The blind user saves the state of visualization for later retrieval; whenever physical interaction takes place, the visualization is retrieved and applied. If a new feature of visualization is found, the new features are configured with the existing state to obtain an updated state of visualization. For instance, a user might identify a new way to reach a starting page.

V. PARTS OF BRAIN – A CASE STUDY

After a brief discussion about visualization, we present a case study to understand how an external visual space is understood by the blind user using a touch screen device. As part of our courseware, an image map is created for exploring the parts of the brain. Our techniques do not require alterations to the underlying touch screen hardware. These techniques are entirely software based. Our design aims to improve the accessibility of existing touch screen hardware for blind users.

We formulate two techniques, namely the Touch technique, and the Control technique. The touch technique is based on the Johnson-Laird model giving more emphasis the conceptual model, where the blind user has adapted to understand the system behaviour. In the control technique, more emphasis is given to user behaviour (Norman Model) in order to facilitate easy accessibility of the system. The techniques which we devised are discussed in detail now.

5.1 Touch based technique

According to these techniques a blind user has to press the surface of a touch screen either from left to right or vice versa or in a zig-zag manner. If the blind user touches the surface and the target is located, then the audio will inform the user of the name of the target (Fig. 3). If the target is not located, then no audio feedback will be received. Long tapping provides audio information about the target. Since we tested in small screen smart phone, if numbers of targets are more, then it is easy to locate more targets. If the target is less in the count and scattered wider around the space, then it is difficult to locate the target and time consuming. This technique needs a lot of patience and memorization of target location when used for the first time. Memorizing the target location will enable blind users to reach the target directly next time.

5.2 Control based technique

The control based technique is based on linearization of items while preserving the original layout of the screen. Blind users, on each tap over the widget placed on the bottom of the screen, will be provided with audio feedback about the name of the target (Fig. 4). A long tap will provide the audio information about the target as in the above techniques.

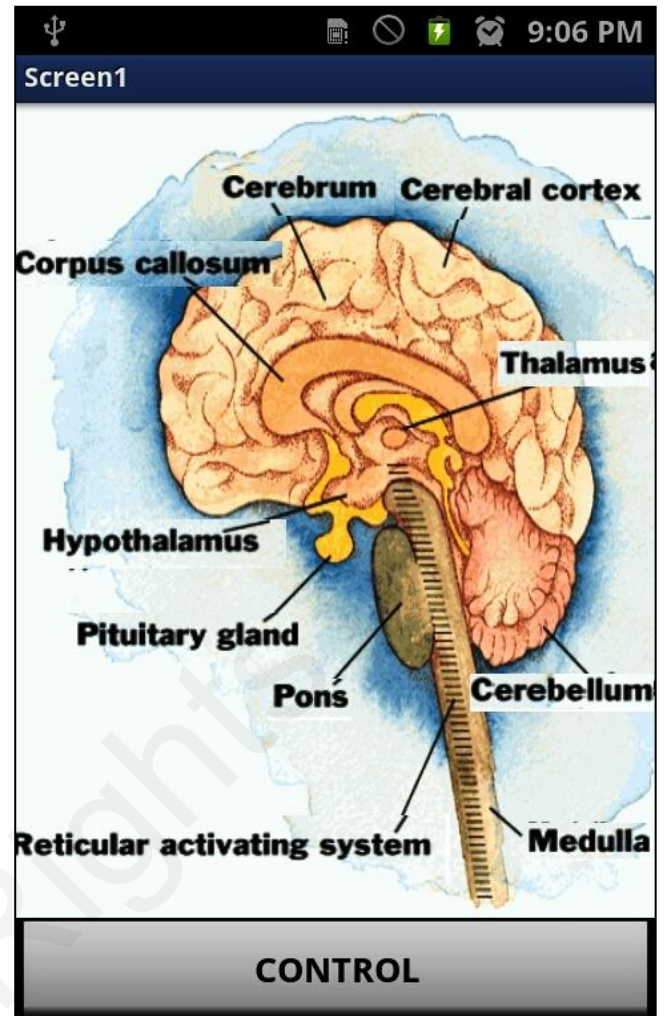


Fig. 4- Control technique

VI. EVALUATION

6.1 Participants

We recruited 10 blind computer users (3 male, 5 female), with an average age of 50.2 (SD=12.4). No participants used a touch screen before. However, all of the participants are use a mobile using screen reader regularly in their daily life.

6.2 Apparatus

The study was conducted using Samsung Galaxy S2 smart phone based on Android. No hardware modification was made. We requested the blind users to identify the single widget above the home button in touch and control based techniques.

6.3 Procedure

The prototype is developed for each technique. The participants are tested to locate the item and understand the information about the item for each technique. Our prototype

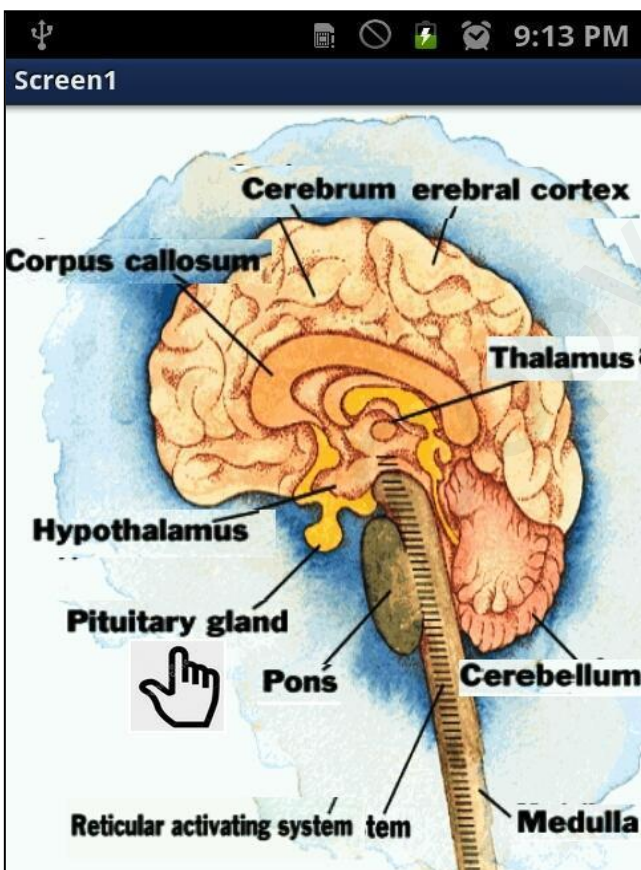


Fig. 3 – Touch technique

has an image map about parts of the brain. It has 9 items such as cerebrum, cerebellum, medulla, Pons and so on. The participants were given the target item and asked to reach the target item. While testing, we observe the following for each technique.

Task Acquisition Time: Time taken by the participant to reach the target was measured in seconds. The time starts when the participant performs the first tap and ends when he reaches the target item.

Stroke Count: Stroke count is the measure of the number of taps the participants performed to reach the target item.

After performing each technique, participants rated the technique using a Likert-scale questionnaire. After all techniques had been tested, participants ranked the techniques in order of preference. Questionnaires were administered verbally, and the experimenter recorded the answers.

VII. RESULTS

The user evaluation is conducted for the Quiz Touch prototype to evaluate accessibility and usability features.

Most complexity in dealing with touch screen by a blind user is in finding targets on the screen. A sighted user can quickly identify a target in a visual interface which uses empty spaces to design individual and group targets. Locating targets on the touch screen requires blind users to touch the empty area when they may not know where they are touching.

Each participant performed averagely 10 trials for each technique. The participant performance for each of the techniques and participant feedback is provided below. Our observation during the trials of each technique is also provided.

7.1 Descriptive statistics

We examined the target acquisition time and stroke count to reach the target for both the techniques.

Stroke Count: On average, blind participants used 21.17 strokes in touch technique to reach the target which is higher than the control technique for which they used 4.25 strokes on average to reach the target (Fig. 5). In addition, the maximum strokes performed by the blind user to reach the target were 60 in the touch technique and 9 in the control technique.

Target Acquisition Time: We analyze how much time the blind users required to reach the target. The descriptive statistics reveal that on average, 9.7 seconds are required to reach the target for the control technique. While the mean time for the touch technique is 43.91 seconds. For the control technique, the maximum duration to reach the target is 16 seconds and minimum is 2 seconds. The maximum task completion time in the touch technique is 125 seconds and minimum is 4 seconds.

7.2 Technique comparison

The dependent variables are checked for normality by using Shapiro Wilkson (W) test based on techniques. The data are not normalized for touch techniques using duration ($W(12) = 0.278, P < 0.05$) or control technique using stroke count ($W(24) = 0.191, P < 0.05$). Hence the Kruskal Wallis H test was performed for not normalized data to find the significance of each technique on the dependent variable.

From the result, it was concluded that there was a statistically significant difference between the duration of the techniques ($H(1) = 9.6134, P < 0.05$). It can be further concluded that the duration of touch (Mean rank=26.17) was more than the duration of control technique (Mean rank=14.67).

A statistically significant difference was also found between the techniques for a stroke count ($H(2) = 9.943, P < 0.05$). The stroke count for touch (Mean rank=26.21) was more than the stroke count of the control technique (Mean rank=14.65).

7.3 Feedback Analysis

The participants completed a questionnaire about the two techniques following the experiment. The participants show their compliance with each technique using a 5-point Likert scale (1=Strongly Disagree, 5=Strongly Agree) using a series of statements.

According to the Friedman Test, there was a statistically significant difference between the control and touch techniques, $\chi^2(1) = 45.302, p < 0.005$. Pair wise comparison using Wilcoxon on signed rank test found that the control technique is more preferred than the touch technique ($Z = -6.232, P < 0.05$).

Significant results were also found for the following measures: easy to use ($\chi^2(1) = 5, P < 0.05$), easy to learn ($\chi^2(1) = -2.60, P < 0.05$), familiar ($\chi^2(1) = -2.232, P < 0.05$), easy to navigate ($\chi^2(1) = -2.041, P < 0.05$) and intuitive ($\chi^2(1) = -2.06, P < 0.05$).

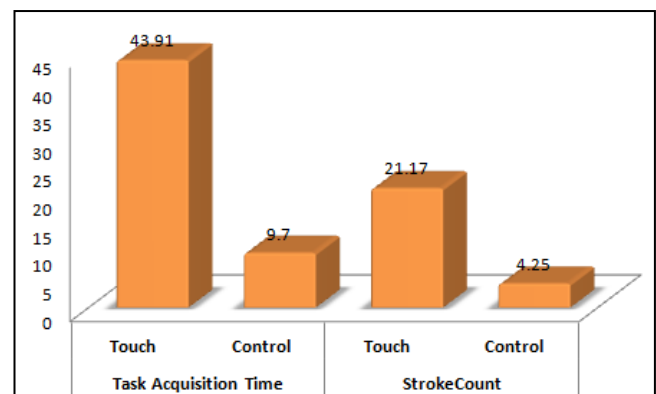


Fig. 5- Descriptive Statistics

VIII. DISCUSSION

A qualitative difference between the touch and control techniques with touch screen device is observed. The primary difference between the two techniques was how the blind user visualizes the external visualization. In the touch technique, users are required to scan the entire surface to locate the target. This method was somewhat slow and time consuming. Considerable efforts and patience are required to accomplish the task. Our observation of this technique reveals that blind users were frustrated during attempts to find the target.

On the other hand, in the control technique, blind users were able to navigate items in a linear fashion. This allowed the users to iterate the items quickly. In addition, the interaction with the system is minimized.

Although participants were faster overall with the control technique, they are not able to visualize the system. They are able to extract the external data through audio cues. However, the actual mind map between the data is missing. Using the control technique, the mind map is linear. On the other hand, using the touch technique, the blind user is able to visualize the external structure.

Note that we adopted “not” describing the actual location of each part. One important consideration when comparing the duration and stroke count is the feedback from the user. The feedback reveals that the control technique is more favoured irrespective of not achieving the visualization. In other words, the blind user prefers more behavioral aspects than conceptual aspects. Fortunately, they feel the visualization can be achieved through the domain expert (educators).

In the end, we can conclude that interaction plays a vital role in building internal visualization. The blind user expects easy interaction with the system. At the same time, the blind user prefers to build up their own mental model based on the external data. In order to maintain equilibrium with both the Norman and Johnson-Laird models, enough metadata has to be provided through external data either in the form of audio cues or through lecturing by educators. The metadata includes the actual location, size and shape of the external structure.

As a result, the external data are just converted to internal data and it is not of much utility in building the internal visualization equivalent to the external visualization. Specifically, the metadata about the external structure helps the blind users to build the internal visualization. By doing this, both behavioral and conceptual models can be kept in tandem.

IX. CONCLUSION

In this paper, we inspect the characteristics of internal representation relevant to an external system. We also investigate the interaction techniques useful to building internal visualizations when external visualization is based on the image map. Our research explores the solution for the questions mentioned in the first section.

What is meant by “internal visualization” for a blind user in terms of image map?. We distinguish the mental model as behavioural, structural and internal. In addition, the mental

model preservation data and mapping between the data or exact replicas of external visualization relies on metadata from external structures sent via audio cues.

For a given external visualization, how can it be related to internal visualization? The relationship between external and internal depends on the interaction technique. The interaction should be simple and easily accessible to the external system. The interaction may be based on the behavioural model developed by the user for easy access or a conceptual model developed by the designer to build the system and the blind user has to adapt to the system.

How do physical activities relate external visualization with internal visualization? It is necessary for external visualization to augment with internal visualization so that they form a blended system. Internal-external blending is performed in terms of six purposes: project, select, precision, coupling, investigation and configuration.

By addressing the research questions, we set up a framework to design an image map using touch screen that merges external visualization, internal visualization, interaction and analytical process. We trust this framework can direct and notify future actions on the design, evaluation and comprehension of the image map in courseware.

REFERENCES

- [1] Abidin, A. H. Z., Xie, H. & Wong, K. W., Blind users' mental model of web page using touch screen augmented with audio feedback. Kuala Lumpur, International Conference on Computer & Information Science (ICCIS), pp.1046 – 1051,2012.
- [2] Brooks, R. A., Intelligence without reason. In: COMPUTERS AND THOUGHT, IJCAI-91. Morgan Kaufmann, pp.569-595, 1991.
- [3] Craik, K., The Nature of Explanation. Cambridge, Cambridge University Press, (1943).
- [4] Fakrudeen, M., Ali, M., Yousef, S. & Hussein, A. H., Analysing the Mental Model of Blind Users in Mobile Touch Screen Devices for Usability, pp. 837-842, (2013a).
- [5] Fakrudeen, M., Yousef, S., Hussein, A. H. & Ali, M., Presentation design of e-assessments for blind users using touch screen devices. Pitesti, pp. 1-6, (2013b).
- [6] Johnson-Laird, P. N., Mental models: towards a cognitive science of language, inference, and consciousness. Cambridge, MA, USA: Harvard University Press,1983.
- [7] Kirsh, D., Projection, problem space and anchoring. In Proceedings of the 31st Annual Conference of the Cognitive Science Society, pp. 2310–2315, 2009.
- [8] Liu, Z., Nersessian, N. & Stasko, J., Distributed Cognition

as a Theoretical Framework for Information Visualization. IEEE Transactions on Visualization and Computer Graphics, vol 14, pp.1173-1180, (2008).

[9] Murphy, E. et al., An empirical investigation into the difficulties experienced by visually impaired Internet users. Universal Access in the Information Society, Springer, Vol 7(1-2), pp.79-91, 2008.

[10] Norman, D., Some Observations on Mental Models. New Jersey: Lawrence Erlbaum Associates, 1983.

[11] Rouse, W. B. & Morris, N. M., On Looking Into the Black Box: Prospects and Limits in the Search for Mental Models. Psychological bulletin, VOL 100(3), pp. 349–363, 1986.

[12] Saei, S. N. S. M., Sulaiman, S. & Hasbullah, H., Mental model of blind users to assist designers in system development. IEEE CONFERENCE PUBLICATIONS, pp.1-5, (2010).

[13] Sasse, M., Eliciting and Describing Users' Models of Computer Systems, PhD, University of Birmingham, (1997).

[14] Scaife, M. & Rogers, Y., External cognition: how do graphical representations work?. International Journal of Human-Computer Studies, VOL. 45(2), pp. 185 – 213, Aug. 1996

[15] Tufte, E. R., Visual Explanations: Images and Quantities, Evidence and Narrative. 4th printing with revisions edition ed, Graphics Press, 1997.

[16] Zhang, Y., The influence of mental models on undergraduate students' searching behavior on the Web. Information Processing & Management, VOL. 44, pp.1330–1345, 2008