Learning Non-Visual Graphical Information using a Touch-Based Vibro-Audio Interface

Nicholas A. Giudice[1], HariPrasath Palani[1]
[1]Spatial Informatics Program:
School of Computing and Information Science,
University of Maine, 5711 Boardman Hall,
Orono, Maine, USA 04469
giudice@spatial.maine.edu,
hariprasath.palani@maine.edu

Eric Brenner[2], Kevin M. Kramer[2]
[2]Advanced Medical Electronics,
6901 East Fish Lake Road,
Suite 190, Maple Grove,
MN 55368
brenneraeric@gmail.com,
kkramer@ame-corp.com

ABSTRACT

This paper evaluates an inexpensive and intuitive approach for providing non-visual access to graphic material, called a vibro-audio interface. The system works by allowing users to freely explore graphical information on the touchscreen of a commercially available tablet and synchronously triggering vibration patterns and auditory information whenever an on-screen visual element is touched. Three studies were conducted that assessed legibility and comprehension of the relative relations and global structure of a bar graph (Exp 1), Pattern recognition via a letter identification task (Exp 2), and orientation discrimination of geometric shapes (Exp 3). Performance with the touch-based device was compared to the same tasks performed using standard hardcopy tactile graphics. Results showed similar error performance between modes for all measures, indicating that the vibro-audio interface is a viable multimodal solution for providing access to dynamic visual information and supporting accurate spatial learning and the development of mental representations of graphical material.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Auditory (non-speech) feedback, Evaluation/methodology; K.4.2 [**Social Issues**]: Assistive technologies for persons with disabilities

General Terms

Design, Experimentation, Human Factors

Keywords

Accessibility (blind and visually-impaired), assistive technology, information graphics, haptic cues, audio cues, android programming, graphs and diagrams.

1. INTRODUCTION

Gaining access to graphical information such as graphs, figures, maps, and images represents a major challenge for blind and visually impaired people. Access to printed material has largely been solved via screen reading software using text-to-speech, for example, programs such as JAWS for Windows (www.freedomscientific.com) or VoiceOver for the Mac and iOS-based portable devices (www.apple.com/accessibility/voiceover/). However, these programs do not have the ability to convey meaningful information about graphic and non-text-based material. Given the vast amount of information which is conveyed through visually-based representations, whether it is in the classroom, the boardroom, or the living room, blind people will continue to miss out on a major component of our information-driven culture unless new non-visual solutions providing access to graphical information are developed. Although this problem has been widely studied (see Section 3, Current Research), approaches for improving the accessibility of graphical information have not made much progress in reaching blind and low-vision users. As this demographic is estimated to number around 12 million people in the U.S. and 285 million people worldwide [21], the need for developing devices that are both usable and likely to be adopted is of growing societal importance. The path forward requires addressing the following limitations which have plagued progress in this domain: research and development projects all too often languish in research labs; the design of new hardware/software is frequently driven by engineering principles without solid theoretical knowledge of relevant perceptual and cognitive characteristics of the human end-user; the systems developed generally have a steep learning curve and rely on unintuitive sensory translation rules; many solutions necessitate purchase of expensive single-purpose hardware; assistive technology often is built-around non-portable devices; and there is an emphasis in the literature on describing technical design features and algorithms, rather than conducting empirical experiments and behavioral evaluations.

Our goal is to provide access to visually-based graphic material using an intuitive interface that provides dynamic information on a device which is inexpensive (i.e. is based on off-the-shelf commercial hardware vs. highly specialized adaptive equipment), is portable enough to be used in many contexts and environments, is multi-purposed (meaning that the underlying hardware can be used for other applications), and supports universal design principles (i.e., is highly customizable and includes many accessibility features in the native interface). To this end, this paper describes what we call a vibro-audio interface, used for conveying visual information via a commercial tablet, which satisfies these design criteria. We believe that the conjunction of considering these design factors from the onset, along with

ASSETS'12, October 22–24, 2012, Boulder, Colorado, USA.

conducting principled empirical investigations to evaluate and refine the perceptibility, usability, and acceptability of the interface, will not only ensure its efficacy in significantly improving the graphical information gap between blind persons and their sighted peers, but does so via a solution which is likely to be readily adopted. This approach avoids the engineering trap, which we argue is the reason that most assistive technology fails, i.e. development is driven by computational efficiency and often naïve assumptions of the designer without feedback of the functional utility of the technology or its ability to address the most critical needs of actual end-users.

2. SYSTEM OVERVIEW

The vibro-audio interface was based on a Samsung Galaxy Tablet with a 7.0 inch touchscreen running Android OS version 3.2, Target version 13. Vibro-tactile information was generated from the tablet's embedded electromagnetic actuator, i.e., an off-balance motor, which was controlled by Immersion Corporation's embedded haptic player. The haptic effects, i.e., vibro-tactile stimuli, for the experimental application were based on the Universal Haptic Layer (UHL) developed by Immersion Corporation (www.immersion.com/products/motiv/index.html). The UHL is a JAR file containing all the classes, interfaces, and algorithms necessary to create dynamic haptic effects on Android devices. The UHL was installed as a plugin for the JAVA development platform (Eclipse) used to create the experimental code. This provided a set of pre-defined haptic effects which were incorporated into the android source code of the application. Auditory output was delivered from the device's onboard speakers. Users also received kinesthetic feedback as they moved their hand over the tablet's touchscreen. Any object, visual or non-visual, that was displayed on the tablet's screen was referenced to a fixed coordinate system and whenever an on-screen visual element was touched, pre-defined vibration patterns and auditory information could be synchronously triggered at that coordinate [see 17 for technical details on the interface]. Although there is only one vibration motor embedded in the device, the use of one finger provides a strong focal stimulus to the digit touching the screen, which is perceived as a tactile point or line as the finger is moved over the stimulus. It should be noted that other studies using touch-enabled devices have found that use of only one finger was sufficient for vibro-tactile line tracing [7, 16], and previous studies on exploration of haptic maps has shown little improvement in learning between conditions using one or multiple fingers [26]. Many stimulus variables could be manipulated and tested in this interface but in this paper, we used a fixed set of parameters established from earlier psychophysical studies in the lab that identified the vibro-tactile line width which is most conducive to line tracing and contour following and the vibratory patterns which best differentiate edges from vertices [17]. Thus, based on these findings, all lines were rendered with a width of 8.9 mm (0.35 inch), which corresponded to 60 pixels on the tablet's screen. This was also used as the minimum inter-line distance for all stimuli. Lines rendered in the vibro-audio mode were given a constant vibration, based on the UHL effect "Engine_100," which uses an infinite repeating loop at 250Hz with 100 percent power. The vertices, either at the end of a single line or at the intersection of two or more lines, were indicated by a pulsing vibration, as our previous research indicated that this cue was helpful for identifying changes in direction during line tracing and for finding the end of individual lines (e.g., the tops of the bars in our bar graphs). Pulses were given in a 60 x 60 pixel (0.35 x 0.35 inch) region encompassing the entire node at the vertex. As nodes at non-orthogonal vertices were not symmetric, the width of the pulsing region varied depending on the intersecting angle of the lines. The pulse signal was based on the UHL effect "Weapon_1," which uses a strong infinitely repeating wide pulse at a frequency of 10-20 milliseconds.

We believe that this interface provides a natural mapping of stimulus information to what is being perceived, while also employing a relatively large (7.0 inch) haptic workspace which can be quickly and easily updated in real-time. Another advantage of the touchscreen is that experimental scripts can be used to log the user's movement behavior and actions, which helps in identifying learning and exploration strategies (although this was not the principle goal of this paper). Assuming the experimental software was made available, this interface could be readily implemented on any off-the-shelf smart touch-based device with at least one embedded vibration motor, the UHL installed, and an audio output facility.

3. CURRENT RESEARCH

Much of the empirical research on accessible graphical displays, auditory or haptic, has focused on design guidelines and user preferences of the interface [12, 13], psychophysical factors characterizing optimal display properties to be implemented or the nature of the perceptual mapping employed [19], or interpretation and legibility of specific information being displayed [8]. These are all important aspects to consider when designing and evaluating a new display but the focus of the current paper addresses a different issue; namely, how accurately graphical information from our vibro-audio interface can be learned and represented in memory as a global spatial image. Earlier research in our lab has demonstrated that users have a favorable opinion of the vibro-audio interface we are using and we have already identified the core vibro-tactile parameters for presenting lines and vertices [17]. Thus, our interest here relates to evaluating whether use of this interface leads to development of an accurate spatial representation of the graphical information being conveyed. The logic is that if the vibro-audio interface is to be truly useful, learning must lead to an accurate representation in memory, similar to that derived from visual access, which supports subsequent mental transformations, computations, and behaviors. Our focus here is on spatial properties of the stimuli. Of note, most graphical information is based on spatial information, and a growing body of literature supports the notion that spatial information encoded from different input modalities can lead to common (amodal) representations in memory which function equivalently in supporting spatial behaviors [see 11 for review].

(Exp 1) (Exp 2) (Exp 3)

Figure 1. Example stimuli displayed on the touch-based device with the vibro-audio mode for the three experiments. Analog hardcopy tactile stimuli (not depicted) were used as a comparison in each experiment.

To address this issue, three experiments were conducted that assess comprehension of the relative relations and global structure between elements on a bar graph (Exp 1), pattern recognition via a letter identification task (Exp 2), and orientation recognition of complex geometric shapes on a shape discrimination task (Exp 3). Each experiment represents a different set of behaviors that rely on accessing an accurate spatial representation built up from learning common graphic material. They all compare two display mode conditions, one that employs the vibro-audio tablet interface at learning and another that employs hardcopy tactile stimuli produced by a graphics embosser (the gold standard for tactile output). In this paper, we are concentrating our discussion on these tasks. Thus, other applications using similar auditory, vibro-tactile, or multimodal displays, such as for navigation are not discussed [but see 6].

4. EXPERIMENT 1: GRAPH LEARNING

The ability to access visual representations of numeric data is critical in many educational and vocational contexts. Indeed, the lack of widely available non-visual rendering techniques has had detrimental effects for blind students on learning and conceptualizing graphs and interpreting patterns and trends of graphical data [20]. Over the years, there have been many research projects investigating the use of dynamic information displays providing access to various types of graphs and charts. These non-visual interfaces can be broadly classified into audio-based [20], haptic-based [24], language-based [4], or multimodal interfaces [25]. The greatest amount of work has been done with auditory graph displays utilizing different sonification techniques where changes in the visual data are mapped onto auditory parameters such as pitch, loudness, timbre, or tempo [19]. Various studies have demonstrated the efficacy of this approach for conveying meaningful information in sonified graphs to blind people [2, 20].

As discussed earlier, we believe that some form of tactual output is the best analog to visually rendered graphics and that a haptic-based display is the best choice for conveying visually oriented spatial data. Most of the research addressing haptic graphs beyond static hardcopy renderings has used force-feedback devices, such as the PHANToM from Sensable Technologies, or the Logitech WingMan force feedback mouse [24], or devices that use piezo-electric pins that dynamically move up and down [15]. The pros and cons of different haptic technologies goes beyond the scope of this paper [for review, see 12] but the main limitations of these displays relate to the earlier mentioned shortcomings of cost, portability, usability, and lack of multi-purpose application.

The advent and proliferation of smooth surfaced touchscreen based devices (e.g., smartphones and tablets) has opened the door to a new era of multimodal interfaces incorporating combinations of auditory, vibro-tactile, and kinesthetic cues. With these devices, hand and finger movements over the display provide position and orientation cues through kinesthesis and the presence of visual elements, such as lines and points, are delivered by an external synchronized cue (such as audio or vibration) when the user touches that element on the touchscreen. We differentiate these devices into two categories based on the perceptual cues provided: audio-kinesthetic interfaces, which couple text and sound cues with hand movement; and haptic-audio interfaces, which add vibro-tactile feedback. Examples of audio-kinesthetic interfaces include Timbremap, which uses sonification for representing complex indoor layouts on a touchscreen equipped smartphone [18] and the PLUMB project, which uses sonification to describe auditory graphs on a touch tablet [3]. Research with both projects supported efficacy of the devices, as users showed clear evidence for accurate perception of the experimental stimuli. Haptic-audio touchscreen-based interfaces differ from traditional hardcopy tactile stimuli and other electronic haptic devices as the cutaneous information being conveyed is purely through vibration on a smooth display surface, rather than the traditional method of feeling embossed lines or moving or vibrating pin arrays. Here, the vibration is generated by rotating electro-magneto vibration actuators which are either fixed internally in the device or fixed to the fingers of the users. An example of the former approach is TouchOver map, which showed that blindfolded-sighted participants could understand a road network through vibration and auditory labels when feeling a smartphone touchscreen, and then were able to accurately reproduce the map using vision while simultaneously exploring the now occluded display [16]. Similarly, the GraVVITAS project demonstrated that graphs, shapes, and maps could be understood by blind users when learned from a touch tablet with external vibrators affixed to the user's fingers [7]. Results from a similar project has shown promising results for apprehending tactile graphs and charts using a touchscreen and multiple piezo-electric motors to stimulate the finger [14]. From a technical standpoint, TouchOver map is most similar to the current research, although we are using a tablet which has twice the screen real estate as the smartphone employed in that project. The GraVVITAS project investigated similar stimuli as we do here, but it used external vibration motors and multiple fingers during exploration. By contrast, we are simply using the internal tablet vibration motor and one point of contact (the dominant finger) on the touchscreen. Importantly, none of these studies required development of an accurate spatial representation to perform the tasks, as is our goal here, and most did not use formal statistical procedures to analyze their data. Despite these differences and the preliminary nature of the research, we interpret the above findings, as well as those from earlier research in our lab with the vibro-audio interface implemented on a smartphone [17], as lending support for the utility of this interface in the current experiments.

4.1 Method

Twelve sighted participants (six males and six females, ages 18-35) were recruited for the study. Three additional blind participants (2 males and 1 female, ages 22-38) also participated. All three were congenitally blind and had no more than light perception. The etiology of blindness was Retinopathy of Prematurity for one participant and Leber's Congenital Amaurosis for the other two. All gave informed consent and were paid for their participation. The study took between 1.5 and 2 hours. Note that it is important to carefully consider whether blindfolded-sighted participants are a reasonable sample when generalizing to blind participants. We believe inclusion is justified here as we are testing the ability to learn and represent non-visual material which is equally accessible to both groups. In support, previous studies with auditory graphs [20] and tactile maps [5] found no differences between blind and blindfolded-sighted groups. Indeed, inclusion of non-representative users (e.g., blindfolded-sighted participants) is generally accepted in the preliminary efficacy testing of assistive technology [see 27 for discussion]. If anything, the performance of the blindfolded-sighted participants in the current experiments represents a conservative estimate of interface efficacy, as this group is likely to be less accustomed to using haptic cues as a primary mode of information gathering. Although our participant samples are too small to make valid statistical comparisons between groups, the similarity of performance observed between blindfolded-sighted and blind participants (as seen in the data figures corresponding to each experiment) provides support for the validity of our inclusion decision.

During the experiment, participants sat on an adjustable chair and adjusted the seat height such that they could comfortably interact with the experimental devices which rested on a 76.2 cm (30 inch) height table in front of them. During the learning phase of each experimental trial, participants wore a blindfold (Mindfold Inc., Tucson, AZ.). In the vibro-audio condition, they used a Samsung Galaxy Tab 7.0 Plus tablet, with a 17.78 cm (7.0 inch) touchscreen as the information display. Vibro-tactile feedback was generated when the user's finger touched the stimulus on the screen and auditory information was provided by tapping the vibrating region (see section 2). In the hardcopy braille conditions, tactile analogs of the same stimuli were produced on paper by a graphics embosser (ViewPlus Technologies, Emprint SpotDot). The paper was then mounted on a second Galaxy tablet such that auditory information could be given in real-time and the user's movement behavior could be tracked via its touchscreen as they felt the hardcopy stimuli (note that no vibro-tactile output was delivered in this condition). Exploration with both displays was done using only one finger (dominant) for all conditions.

4.2 Procedure

A within subjects design was used in the experiment, with participants learning and testing on three bar graphs in each of the two display mode conditions: hardcopy braille and vibro-audio (graph trials were randomized within display mode block, with block order counterbalanced between participants). Each display mode condition included a graph with *3*, *4*, and *5* bars (presentation order was randomized within graph set, with set order alternating between participants). Each bar was assigned a name, with set 1 based on food: *pizza, burger, salad, chocolate and ice cream*; and set 2 on fields of study: *biology, physics, chemistry, mathematics, and computer science*. The name was spoken as an audio message when the user tapped on the bar.

The experiment consisted of a practice, learning, and testing phase for each display mode condition, for a total of 10 trials. The first practice trial in each display mode was a demo trial where the experimenter explained the task, goal, and search strategies and the participant explored the stimuli with corrective feedback. In the second practice trial, participants were blindfolded and asked to perform the complete experimental learning and testing sequence. The experimenter evaluated their answers immediately to ensure that they understood the task correctly before continuing. During the learning phase, blindfolded participants were asked to explore the graph and to indicate when they believed that they had learned all of the material represented. They were instructed to learn as quickly and accurately as possible. They were told that the height of each bar represented *how many people liked the specific food category* (Set 1) or *how many people were enrolled in the class* (Set 2). After learning, the experimenter removed the device and the participant was allowed to lift their blindfold. The testing phase consisted of two tasks. In the spatial relation task, participants answered four questions about the graph they just learned. Two of the questions assessed spatial relations between bars. For instance, "*What is the relation between biology and physics?*" The answer required a directional response (e.g., biology *is left/right of physics*), and a height judgment (e.g., biology *is taller/shorter than physics*). The other two questions assessed participant's ability to comprehend the individual bar position in a global context. For instance, "*Which is the second highest bar?*" "*What is the middle bar?*" To reduce recall errors, the names of the bars were given in a list.

In the re-creation task, participants were asked to draw the graph on a template canvas of the same size as the display and to label each bar. Five equidistant textbox place holders were provided to indicate the possible bar positions. The only procedural differences for blind participants were that the questions were read aloud by the experimenter and the reproduction task was done with Lego™ pieces on a board (which provided the same position indicators). They labeled each bar by verbally indicating its name. All re-created graphs were analyzed in terms of whether individual bars had the correct label, position, and relative height in relation to the graph's global structure. From this design, we can evaluate several measures as a function of display mode condition. These include learning time from the learning phase, relative height accuracy, relative directional accuracy, relative position accuracy (i.e., individual bar position in relation to the global context), re-creation accuracy, and bar labeling accuracy.

4.3 Results

The most important outcome of this experiment, as shown in Figure 2, is the similarity of performance across all measures for both display modes (braille or tablet) and participant groups (blindfolded-sighted and blind).

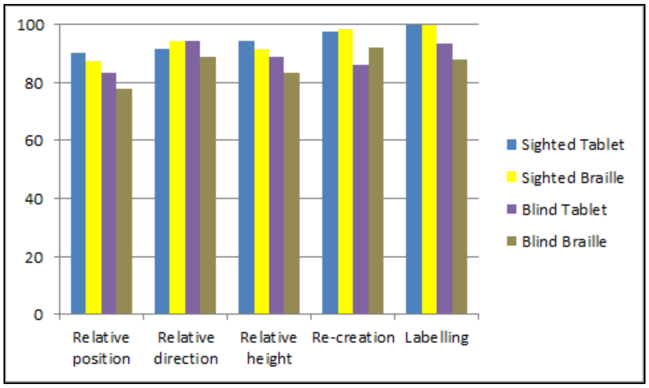

Figure 2. Accuracies on test measures as a function of display mode and subject group.

Corroborating what can be seen in the figure, results of paired-sample t-Tests between the two display modes(tablet and braille) were highly in-significant for all measures ($\alpha = 0.05$ was used for all statistical tests): relative height accuracy ($t(35) = -0.329$, $p = 0.744$); relative directional accuracy ($t(35) = -0.329$, $p = 0.744$); relative positional accuracy ($t(35) = -0.828$, $p = 0.413$); and re-creation accuracy, ($t(35) = -1.000$, $p = 0.324$).

Repeated measures ANOVAs were also conducted on each variable to assess if there were effects of the number of bars (e.g., 3, 4, or 5) between the two display modes; no statistically significant differences were found (all ps >0.05).

What is evident from these data is that use of a vibro-audio interface on a touch-enabled device supports accurate learning of relative relations and global structure of a bar graph. Importantly, the similarity of performance with this interface compared to that observed after learning with traditional hardcopy tactile output suggests the building up and accessing of functionally equivalent spatial representations between display modes. Superior performance for the hardcopy tactile mode was observed in learning time, ($t(35) = -4.924$, $p < 0.001$). This makes sense, as it is easier to find and track the line using the embossed brailled stimuli. Despite these differences, the more important findings of this experiment are the striking similarity in output performance between display modes for both participant groups. We interpret these results as providing compelling evidence that once learned, the representations built up from use of the vibro-audio display supported the same level of spatial behaviors as those built up from hardcopy tactile stimuli.

From these results, it can be seen that in general, both blindfolded-sighted and blind subjects yielded higher accuracy values with the re-creation task than with the spatial relations task. This result may be due to re-creation being done sequentially, whereas performance on the spatial relation questions required making judgments about bars that often required non-contiguous and non-sequential judgments. Also, as seen in Figure 2, participants average accuracy with the tablet mode for measures of positional accuracy, relative direction, relative height, and labeling were numerically higher than in the braille mode. Although not statistically different, this trend suggests that the interface leads to development of a spatial representation of the graph which is at least as good, if not better than from hardcopy stimuli.

5. EXPERIMENT 2: LETTER RECOGNITION

This experiment used the same vibro-audio tablet interface and hardcopy tactile stimuli as Exp 1 but now for recognizing patterns based on capital letters from the English alphabet. Letters represent complex but well known shapes and require participants to trace the contour of the stimuli and build up a global representation of its shape in order to correctly name the letter. This task has been used effectively in the past with different vibro-tactile stimuli [9] as well as visual apprehension with a limited field of view [22, Exp 3]. To our knowledge, non-visual letter recognition has not been studied with vibro-tactile touchscreen devices but early research with systems that converted visual information from camera input into vibro-tactile output were shown to support letter recognition via a 20 x 20 array of vibro-tactile stimulators on the back [10]. A device called the Optacon, which used an array of 144 electro-tactile stimulators felt by the finger, even proved useful for real-time letter recognition and limited reading [1].

Figure 3. Subject tracing stimuli displayed on the touch-based device with the vibro-audio mode.

Although we are using letters as the stimuli in this experiment, our goal here is to compare pattern recognition performance between the vibro-audio interface and hardcopy braille and not to test the efficacy of this interface for reading printed letters, although it could in theory be used in this capacity.

5.1 Method

The same participants, apparatus, and two display modes were used here as in Exp 1. The within subjects design also followed the same procedure of two practice trials and three experimental trials per display mode (counterbalanced). The task in this experiment was for blindfolded participants to explore the stimuli (one of six randomly presented letters) and to name the letter as soon as it was recognized. The six letters used during the experimental trials included: *D*, *F*, *M*, *P*, *T*, and *W* (with *N* and *C* used in the practice conditions). The letters were selected such that each display mode condition included three unique patterns including a letter with straight lines (*F* or *T*), a letter with curves (*D* or *P*), and a letter with slanted lines (*W* or *M*). A pulsing vibration was provided at each vertex in the vibro-tactile condition. No audio cues were used in this experiment. If the letter was mis-identified, a second learning period was allowed following the same procedure. Incorrect identification on the second learning phase was considered a miss.

Total learning time, number of learning iterations, and pattern recognition accuracy were evaluated as a function of display mode condition.

5.2 Results

As shown in Figure 3, the letter recognition performance for blind participants was done without error in both display modes. However, for sighted subjects, the ~89% letter recognition accuracy performance with the vibro-tactile interface was significantly worse than the 100% accuracy observed in the hardcopy braille mode, as assessed by a paired samples t-Test, (t(35) = 2.092, p = 0.044). This difference is likely due to the impoverished orientation cues available in the tablet mode, which made it harder to detect line orientation, especially if the line was slanted or curved. Although the pulsing vibration at the vertices helped in determining an intersection or end node, there were no orientation cues to assist with non-rectilinear stimuli, which is apparently particularly challenging in the vibro-audio interface. In the braille condition, the embossed lines make it easier to detect line orientation and to follow the lines when they change direction (something that pilot studies in the lab have indicated is challenging with the vibro-audio interface).

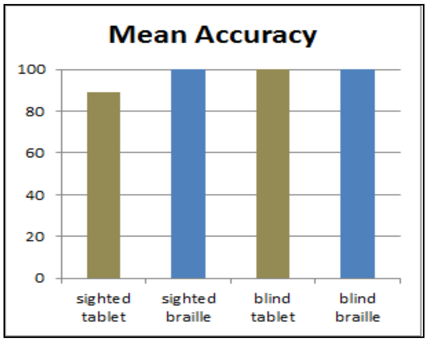

Figure 4. Letter recognition accuracy as a function of display mode and subject group.

The mean exploration iterations (sighted tablet: 1.2, sighted braille: 1.0, blind tablet: 1.2, blind braille: 1.1) for both the modes are greater than 1 iteration, which suggests that even in the braille modes participants made errors in their first recognition attempt. Also, the letters with symmetric patterns contributed to the wrong interpretation. For example, the W was often interpreted as V, U, or M. This occurred because subjects often traced only half (or part) of the object and then guessed that it was U or V. However, when traced fully, subjects tended to count the number of lines and to use this as a strategy to narrow the possible letter alternatives. Finally, as in the previous experiment, a significant difference was observed in learning time, ($t(35) = -6.137$, $p < 0.01$). As expected, this manifested with tablet conditions being slower than braille conditions, as discussed earlier.

6. EXPERIMENT 3: ORIENTATION DISCRIMINATION

This experiment investigated the ability to learn and represent the orientation of irregular shapes, consisting of four-sided polygons which were misaligned with the display's intrinsic frame of reference. After learning, participants had to match the learning stimulus with four alternatives based on the same shape presented at four different orientations. Other research has shown that touchscreen devices with external vibration actuators are beneficial in supporting recognition of shapes and patterns [7, 23]. The importance of this experiment is that it not only requires learning a complex shape, as has been previously investigated, but that the representation built up from learning was sufficiently robust to recall and discriminate the target shape in the presence of geometrically identical alternatives.

6.1 Method

The same participants, apparatus, and two display modes were used here as in the previous experiments. The within subjects design also followed the same procedure of two practice trials and three experimental trials per display mode (counterbalanced). The task in this experiment was for the blindfolded participant to explore the shape during a learning phase and to stop once they felt that they were familiar with its global geometry and orientation. Three distinct shapes were used in each display condition (counterbalanced). Only the bounding contour of the shape was rendered and none were readily namable polygons (see figure 1). No audio cues were used in this experiment and the pulsing vibration was provided at each vertex in the vibro-tactile condition. During learning, participants were asked to imagine the vertices, length of the sides, and the orientation of the shape on the display. Once participants indicated that the shape was learned, the experimenter removed the device and placed an A4 size paper containing the same shape in four different orientations. The shapes were numbered from 1 to 4 in a column.

Figure 5. Alternatives for the example shape displayed in Figure 1.

Participants removed their blindfold and marked the alternative which matched the orientation of the shape previously learned. Blind participants performed the same task but made their comparison based on a sheet with 3D cut-outs of the four alternative shapes (all stimuli were size-matched). Measures analyzed included time to learn and orientation accuracy.

6.2 Results

No reliable differences were observed in the paired sample t-Tests conducted between the two display modes for orientation accuracy ($t(35) = 0.298$, $p = 0.768$). These results suggest that learning with the tablet mode was functionally equivalent to learning with the braille mode for apprehending shapes and for identifying the reference shape from geometrically identical alternatives.

However, as is shown in Figure 6, the orientation performance with the tablet mode yielded lower numeric means,~83% and ~77% mean accuracy, contrasting with ~86% and ~88% accuracy for the braille mode for sighted and blind subjects respectively.

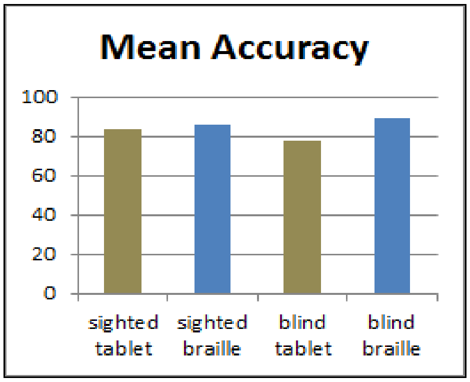

Figure 6. Orientation accuracy as a function of display mode and subject group.

Similar to Exp 2, subjects self-reported difficulty in identifying the slanting lines as they felt that the perceptual cues from the vibro-audio interface were not as "sharp" as with the hardcopy braille stimuli and that it was hard to monitor their hand trajectory when it was not moving in register with one of the intrinsic axes of the device. As with the previous experiments, learning time with the braille stimuli was significantly faster than with the vibro-audio interface, ($t(35) = -7.170$, $p < 0.001$).

7. GENERAL DISCUSSION

Three experiments were conducted that assessed the ability to learn and represent three types of graphical information using a newly developed vibro-audio interface [17] and compared performance with traditional hardcopy tactile representations of the same graphics. Overall, our results provided strong support for the efficacy of the vibro-audio interface for accurately perceiving and learning the experimental stimuli and in building up accurate mental representations supporting various spatial operations for both blindfolded-sighted and blind participants. Importantly, error performance did not reliably differ between display modes on any of the measures tested, demonstrating that the vibro-audio interface provides a comparable level of access to graphical material as is possible from a traditional hardcopy medium. These findings are important as this interface provides dynamic and readily implemented information, whereas hardcopy material is static and requires expensive, highly specialized equipment to produce. In addition, as the vibro-audio interface is based on inexpensive, multi-purpose, and commercially-available hardware, it represents a viable alternative to the expense and complexity of existing auditory and haptic solutions which have various shortcomings, as described earlier. Although all participants reported familiarity with touchscreen devices, they had never experienced vibro-audio graphical stimuli as were evaluated here. Indeed, the highly similar results between modes observed across experiments are quite remarkable given the absence of significant training with the vibro-audio interface. This performance occurred with a combined practice period of around 30 minutes, compared to shape and letter recognition proficiency with other assistive technology, (e.g., the optacon) which took well over 100 hours [6]. We interpret this robust finding across all of the experimental conditions and between both blindfolded-sighted and blind participant groups as showing the intuitiveness of the interface. In support, post-experiment debriefing revealed that all participants liked the vibro-audio interface and that the blind participants expressed interest in adopting it as a primary graphics display if it were further developed. Providing increased familiarity and some small modifications to the interface may well improve learning time and some of the behavioral ambiguities we observed, as discussed below.

Although all subjects performed quite well in the Exp 1 graph conditions, their strategy of moving perpendicularly between the tops of the bars (i.e., to gauge their relative heights) was sometimes challenging in the tablet condition as they had trouble moving laterally, often deviating upward during their trace. This behavior was not observed in the hardcopy braille condition, as the lines provided a better fixed reference on the paper. These results, along with the challenges observed in the tablet condition for following slanted and curved lines in the letter and shape recognition experiments, suggest the need for developing a secondary cue to assist with contour tracing and for staying oriented when exploring non-rectilinear stimuli. A related phenomenon is that slight orientations in the stimuli (10 to 20 degrees) were perceived as a straight line in the tablet conditions. This problem could be resolved in the future by using auditory information to indicate deviation from a given line orientation.

In Exp 2, letter recognition performance was influenced by the similarity of the pattern in the tablet condition. That is, letters such as “D” and “P” were interpreted as the same since they have a line and a curve in common. Since these pattern errors were only observed in the first learning attempt, and correct recognition was near perfect after the second learning iteration, this problem is likely due to lack of familiarity with the tablet interface than to actual challenges interpreting the information conveyed. Thus, with the addition of new auditory cues to complement the vibro-tactile information, and more training with the interface, it is likely that many of these challenges would be ameliorated. Even so, it is remarkable how well the tablet device faired compared to the tried and true hardcopy tactile output. Not only was performance with the vibro-audio interface nearly equivalent on most conditions, it was actually better on some, even though this interface was completely new to our participants.

The time taken to learn was significantly different between the braille and tablet modes for all conditions. Although the learning time with the tablet was approximately four times greater than the time taken in the braille conditions, this was not unexpected owing to differences in the way information is conveyed and extracted between modes. As discussed earlier, adding additional complementing cues and allowing greater experience with the vibro-audio interface is predicted to narrow this gap. Future experiments need to investigate which cues might aid in this process. We believe that with additional technical advancements and usability evaluations, the tablet could be improved to support even better performance and provide access to a far broader range of graphics than were tested here with the initial prototype interface.

Future studies will focus on enhancing the ability of the interface such that it can automatically convert existing visual graphics into an accessible form suitable for presentation as vibro-audio graphics. Research will also include identifying the optimal vibrating pattern for different spatial objects and the use of piezo-electric actuators instead of mechanical vibrators. Piezo-electric actuators will enable high definition haptic effects. For instance, Immersion’s high definition embedded player is touted to deliver powerful and crisp effects with: (1) a wide haptic band-with (50 - 350 Hz) affording increased range, strength, and precision, (2) superior effects isolation, as the actuator can be mounted to vibrate only the touchscreen, (3) instantaneous, low latency touch feedback which reduces haptic lag time below human perception, and (4) quiet piezo haptics, that reduce noise below the audible range(http://ir.immersion.com/releasedetail.cfm?ReleaseID=444761)]. We also plan to further evaluate the usability of this interface for other tasks, such as for non-visual map learning and for assisting spatial learning, navigation, and cognitive map development.

8. ACKNOWLEDGMENTS

We thank Advanced Medical Electronics Corporation for providing the required software and Tablet devices in these experiments, and Bill Whalen for assistance with the manuscript. We acknowledge support from NIDRR grant H133S100049 and NSF grant CDI-0835689 on this project.

9. REFERENCES

[1] BLISS, J.C., KATCHER, M.H., ROGERS, C.H., and SHEPARD, R.P., 1970. Optical-to-tactile image conversion for the blind. *IEEE Transactions on Man-Machine Systems 11*, 58-65.

[2] BREWSTER, S.A., 2002. Visualization tool for blind people using multiple modalities. *Disability and Rehabilitation Technology 24*, 11-12, 613-621.

[3] COHEN, R.F., YU, R., MEACHAM, A., and SKAFF, J., 2005. Plumb: displaying graphs to the blind using an active auditory interface. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility (ASSETS05)* ACM Press, New York, NY, 182-183.

[4] FERRES, L., LINDGAARD, G., and SUMEGI, L., 2010. Evaluating a tool for improving accessibility to charts and graphs. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility* ACM, New York, NY, Orlando, Fl, 83-90.

[5] GIUDICE, N.A., BETTY, M.R., and LOOMIS, J.M., 2011. Functional equivalence of spatial images from touch and vision: Evidence from spatial updating in blind and sighted individuals. *Journal of Experimental Psychology: Learning, Memory and Cognition 37*, 3, 621-634.

[6] GIUDICE, N.A. and LEGGE, G.E., 2008. Blind navigation and the role of technology. In The *Engineering Handbook of Smart Technology for Aging, Disability, and Independence,* A. HELAL, M. MOKHTARI and B. ABDULRAZAK (Eds). John Wiley & Sons, 479-500.

[7] GONCU, C. and MARRIOTT, K., 2011. GraVVITAS: Generic Multi-touch Presentation of Accessible

Graphics. In *Proceedings of INTERACT2011* Springer, Berlin / Heidelberg 30-48.

[8] HOGGAN, E., BREWSTER, S.A., and JOHNSTON, J., 2008. Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens. In *Proceedings of ACM CHI2008* ACM Press, Addison Wesley, Florence, Italy, 1573-1582.

[9] LOOMIS, J.M., 1974. Tactile letter recognition under different modes of stimulus presentation. *Attention, perception, & psychophysics 16*, 2, 401-408.

[10] LOOMIS, J.M. and APKARIAN-STIELAU, P., 1976. A lateral masking effect in tactile and blurred visual letter recognition. *Attention, Perception & Psychophysics 20*, 221-226.

[11] LOOMIS, J.M., KLATZKY, R.L., and GIUDICE, N.A., in press. Representing 3D space in working memory: Spatial images from vision, touch, hearing, and language. In *Multisensory Imagery:Theory & Applications*, S. LACEY and R. LAWSON (Eds). Springer, New York.

[12] MACLEAN, K.E., 2008. Haptic Interaction Design for Everyday Interfaces. In *Reviews of Human Factors and Ergonomics*, M. CARSWELL (Ed). Human Factors and Ergonomics Society, Santa Monica, CA, 149-194.

[13] NEES, M.A. and WALKER, B.N., 2009. Auditory Interfaces and Sonification. In *The Universal Access Handbook*, C. STEPHANIDIS (Ed). CRC Press, New York, 507-521.

[14] PETIT, G., DUFRESNE, A., LEVESQUE, V., HAYWARD, V., and TRUDEAU, N., 2008. Refreshable tactile graphics applied to schoolbook illustrations for students with visual impairment. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility (Assets08)*, 89-96.

[15] PIETRZAK, T., CROSSAN, A., BREWSTER, S.A., MARTIN, B., and ISABELLE, P., 2009. Exploring Geometric Shapes with Touch. In *Proceedings of INTERACT2009*, 145-148.

[16] POPPINGA, B., MAGNUSSON, C., PIELOT, M., and RASSMUS-GRÖHN, K., 2011. TouchOver map: audio-tactile exploration of interactive maps. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices* ACM, Stockholm, Sweden, 545–550.

[17] RAJA, M.K., 2011. The development and validation of a new smartphone based non-visual spatial interface for learning indoor layouts. Unpublished Masters thesis: *Spatial Information Science and Engineering*, University of Maine, Orono (advisor: N.A. Giudice).

[18] SU, J., ROSENZWEIG, A., GOEL, A., DE LARA, D., and TRUONG, K.N., 2010. Timbremap: enabling the visually-impaired to use maps on touch-enabled devices. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices* ACM, 17-26.

[19] WALKER, B.N., 2002. Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied 8*, 211-221.

[20] WALKER, B.N. and MAUNEY, L., 2010. Universal Design of Auditory Graphs: A Comparison of Sonification Mappings for Visually Impaired and Sighted Listeners. *ACM Transactions on Accessible Computing 2*, 3, Article 12, 16 pages.

[21] WORLD HEALTH ORGANIZATION, 2011. Visual impairment and blindness Fact Sheet, Retrieved 2012 from: http://www.who.int/mediacentre/factsheets/fs282/en/

[22] WU, B., KLATZKY, R.L., and STETTEN, G.D., 2012. Mental visualization of objects from cross-sectional images. *Cognition 123*, 1, 33-49.

[23] YATANI, K. and TRUONG, K.N., 2009. SemFeel: a user interface with semantic tactile feedback for mobile touch-screen devices. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, 111-120.

[24] YU, W. and BREWSTER, S.A., 2002. Comparing Two Haptic Interfaces for Multimodal Graph Rendering. In *proceedings of IEEE VR2002, 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Florida.

[25] YU, W. and BREWSTER, S.A., 2002. Evaluation of multimodal graphs for blind people. *Universal Access in* the Information Society 2, 2, 105-124,.

[26] GIUDICE, N.A., MASON, S.J., & LEGGE, G.E., 2002. The relation of vision and touch: Spatial learning of small-scale layouts. Journal of Vision, 2(7), 522A.

[27] SEARS, A. AND HANSON, V., 2011. Representing users in accessibility research. Proceedings of the 2011 annual conference on Human factors in computing systems.