Comparing Natural Language and Vibro-Audio Modalities for Inclusive STEM Learning with Blind and Low Vision Users

Justin R. Brown
VEMI Lab, The University of Maine, Orono, ME, USA
justin.r.brown@maine.edu

Stacy A. Doore
INSITE Lab, Department of Computer Science, Colby College, Waterville, ME, USA
sadoore@colby.edu

Justin K. Dimmel
IMRE Lab, School of Learning and Teaching, The University of Maine, Orono, ME, USA
justin.dimmel@maine.edu

Norbert Giudice
VEMI Lab, The University of Maine, Orono, ME, USA
norbertgiudice@gmail.com

Nicholas A. Giudice
VEMI Lab, School of Computing and Information Science, The University of Maine, Orono, ME, USA
nicholas.giudice@maine.edu

ABSTRACT

Data representations continue to be produced in predominantly visual forms within STEM disciplines. The disparity in access to these graphical representations between students who are blind or have low vision (BLV) and their sighted peers is exacerbated as the adoption of digital screens become more prevalent in educational settings. Standard accessibility solutions rely heavily on natural language processing, e.g., screen readers, for non-visual information access. But can other non-visual modalities, like touch, be effective for learning graphical content rendered on touchscreens? To investigate this question, we conducted a user study with a multimodal touchscreen learning system to assess the effectiveness of two non-visual graphical presentation modalities: 1) a vibro-audio condition, which used the device's embedded vibration motor plus an auditory content overview (a spatial and multimodal technique), and 2) a natural language condition that provided a complete description of the content (a cognitively mediated technique). BLV participants (N = 19) were presented with the learning system and asked to answer multiple-choice questions about three different graph types using both presentation modalities. Findings showed that the two presentation modalities were functionally equivalent for learning the graphical information presented, suggesting that for these stimuli, the presentation modality did not have a significant effect on participant graph learning accuracy. However, the type of graph being learned did have a reliable effect. Moreover, a majority of the participants stated a preference towards the natural language condition and, on average, learned graphs faster than with the vibro-audio condition. The similarity found between the two learning modalities is interpreted as supporting user learning preferences while providing redundancy in the information being communicated. This approach layers the various types of information found in graphical representations (text, numerical, spatial) for individuals with accessible learning needs, providing more control, independence, and responsive tools to optimize their own educational materials.

CCS CONCEPTS

• **Human-centered computing** → Accessibility; Accessibility systems and tools; Human computer interaction (HCI); HCI design and evaluation methods; Usability testing; • **Applied computing** → Education; Interactive learning environments; • **Social and professional topics** → User characteristics; People with disabilities.

KEYWORDS

Information access, Vibro-audio, Natural language, STEM graphs, Blindness and low vision

ACM Reference Format:
Justin R. Brown, Stacy A. Doore, Justin K. Dimmel, Norbert Giudice, and Nicholas A. Giudice. 2023. Comparing Natural Language and Vibro-Audio Modalities for Inclusive STEM Learning with Blind and Low Vision Users. In *The 25th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '23), October 22–25, 2023, New York, NY, USA*. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3597638.3608429

ASSETS '23, October 22–25, 2023, New York, NY, USA

ACM ISBN 979-8-4007-0220-4/23/10…$15.00
https://doi.org/10.1145/3597638.3608429

1 INTRODUCTION

The ability to access information through graphical charts, diagrams, graphs, and plots is a critical skill for success in science education, business transactions, and informed citizenship as it allows for the communication of information quickly and effectively [18, 31, 68]. For sighted individuals, vision is used in an immediate and unconscious manner for understanding these data representations. For instance, a complex table can be transformed into a color-coded chart that summarizes business expenses into meaningful categories that are accessed through vision to interpret a business' expenditures. The skills used to interpret these data visualizations, like the chart described, are developed starting in early elementary school curricula [13, 19] and reinforced throughout primary and secondary education [54]. Indeed, reliance on graphical content and numeric representations in the classroom is sharply increasing, with graphics found on almost every page in an evaluation of 5-8 grade math textbooks [20] and another study estimating

1.3 figures per page in high school biology texts [55]. This trend is extremely problematic for BLV learners, given the dearth of alternative modes of access to this information. The impact of inaccessible content, leading to an insufficient understanding of graphical information, can have life-long consequences for the 7.5 million (2.5%) Americans who are blind or have low vision. It also has immediate detrimental impacts on the near 550,000 children with severe vision difficulty under the age of 18 [65], many of whom are students in classes that rely on this inaccessible content.

There are alternative avenues for BLV students to learn foundational graphical concepts alongside their sighted peers. These include tactile graphics on embossed paper with Braille [54], individual instruction from a support aid [56], and using a variety of other assistive technologies (see literature review below). While these methods have shown to be beneficial, they also have some significant challenges that limit their efficacy and implementation. Most notably, this includes being time intensive to produce readily available and accessible STEM (Science, Technology, Engineering, and Mathematics) graphical equivalents by highly specialized transcriptionists. Estimates for transcribing a full course textbook are between 4 months and up to a full academic year, which is not practical when a student needs the textbook content immediately to keep pace with the rest of the class instruction [73]. It is also often prohibitively expensive for many small school districts to hire these transcription specialists, as converting printed STEM textbooks into an accessible Braille and tactile graphics format is hugely expensive. Although such numbers are hard to find, as transcription services generally require a formal request to provide a quote, we have been informally told by multiple teachers of the visually impaired (TVIs) and school districts that these conversion costs can be up to $30,000 per book, which is in alignment with a 2017 news piece suggesting that converting only five chapters of a science book may cost $15,000 [40]. Finally, the process does not promote students' feelings of independent learning and agency [22, 24, 54] or full inclusion in the educational environment.

As the digital age continues to produce vast amounts of data, high quality data visualizations have become the primary way to communicate data, identify trends, and summarize significant patterns [33]. The lack of accessible data science and graphical materials for BVI learners at all levels presents severe challenges for creating inclusive STEM courses [6, 42]. If BLV learners are not given the appropriate tools to learn these valuable data science concepts, many of them may be discouraged to pursue STEM related careers. NSF STEM participation data does not break down science participation by disability type [8], however, providing equitable opportunities is only possible through designing accessible STEM training tools and learning materials to support future labor contexts and career training for people of all abilities and ages. Given the high levels of under/unemployment for BLV adults ages 18-64 (~50%) [65], providing greater access to data focused training resources is a particularly pressing societal need for improving career opportunities and earned income potential. Developing a low-cost multimodal diagram system to assist BLV students in engaging and interpreting data through graphical representations is one of the goals of this research. Furthermore, our research is motivated by personal learning environment frameworks [14] that are customizable, equitable, and further increase student success.

The user study we describe in this paper evaluates the usability of two viable non-visual methods of accessing information from graphs presented through the multimodal diagram system developed for an earlier study [17] by nineteen BLV participants. The two methods used within this study are: Natural Language descriptions (NL) and haptic presentation using vibration cues with an auditory overview of the graphic, called Vibro-Audio exploration (VA). The auditory overview serves as a quick summary to guide and orient the user to the graphic and improve their exploration behavior. Including this overview was informed by previous research that indicates the importance of contextual cues and overviews for effective active exploration [16, 32]. The VA condition allows for perceptual interaction through touch and may prove more accurate for conveying the spatial information of the tested graphics [23]. The NL modality, which provides descriptions that are more cognitively mediated, is a more familiar approach owing to the prevalence of screen readers employing text-to-speech engines, but lacks direct spatial perception. The diagram system is built on an iOS-based smartphone, which is a platform well adopted by BLV users [27] and can be navigated non-visually with Apple Inc.'s native VoiceOver feature [4]. The development of the diagram system is beyond the scope of this paper and a full description can be found in Doore et al., 2023 [17]. Our focus in this study is on comparing the effectiveness of the two non-visual graphical presentation modalities being tested.

2 RELATED LITERATURE

Inclusive educational environments are now the most common learning setting with over 80% of BLV students attending local public schools and spending most of their time in inclusive lessons with their sighted peers [2, 29]. This move for inclusion is important but also presents significant challenges in school for BLV students who may not be able to fully participate in spontaneous whole class discussions and small group work centered around interpreting inaccessible graphical information and making connections between the data and larger concepts. This means BLV students are missing out on important social, communication, and behavioral skill building that are a critical part of conceptual learning and the inclusive classroom experience [57, 67]. Limited availability of accessible materials can also put teachers in a situation where phrasing, structure, or terminology does not correspond with best pedagogical practices for teaching graphical concepts. This results in visual graphical information interpretation being difficult at best, and impossible at worst, leading to BLV students consistently falling behind standard grade level academic achievement [7, 61]. While there have been many approaches to bridging the gaps in accessible graphical content in the classroom, most of these solutions rely on a translation of the visual (i.e. spatial) information using either natural language descriptions or haptic renderings.

2.1 Natural Language Literature

The most notable piece of assistive technology that utilizes natural language processing is a screen reader. Screen readers work well to access digital text and can access the alt text on digital figures, like on webpages, to make them available for consumption without vision. A digital graphical rendering may be more conducive for sharing information quickly and in a greater expanse, i.e., the internet, but it does not make it more accessible [32, 44]. With graphical content, screen readers are limited by the alt text that is provided; this text can be highly variable and may provide an insufficient description. A problem that is often encountered by BLV individuals is the absence of this critical alt text on digital content. In a recent 2022 paper, Jung et al. [30] found that of the nearly 7500 figures submitted to two scientific publications (CHI and ASSETS), only 40% of these figures had accompanying alt text. A majority of these alt text descriptions more or less summarized the figure's information, but failed to provide the type of chart or graph it was describing. Take note that these poor numbers are the result of content that has been submitted to usability/accessibility-related conferences. The general use and effective implementation of alt text within knowledge dissemination is inevitably worse for non-accessibility focused contexts.

With this problem in mind, many researchers have looked toward developing systems, programs, and guidelines to automatically generate alt text or summaries [10, 21]. Most current approaches involve a combination of computer vision and natural-language processing methods of chart classification, text recognition and extraction, chart element decoding, and data extraction [1, 51] and the testing of these systems have focused on line graphs and bar charts [21]. In some cases, the system can take in the visual image and then generate structured descriptions based on chart type templates. In other cases, the chart images (bitmap, png, jpeg, etc.) are converted to dynamic representations that can interact with a screen reader, providing some level of choice of low-level (low detail) vs. fine-grained (high detail) information [10]. Low level descriptions convey the general structure and pertinent statistics within a graphic and in fine-grained level descriptions the individual values of the data may be presented along with notable trends and contextual knowledge relevant to the purpose of the graphic. A recently proposed four-level model for semantic content used in natural language descriptions of accessible graphs suggests using a framework for evaluating generative alt text systems for accessible content and the amount of flexibility in the type of information that the user can choose to have presented [39].

While many of these recent approaches for automated accessible descriptions offer promising advances on the technical side of the problem, there is some concern that many of the current solutions do not include end-user input or hands-on experience from the target group of BLV individuals who rely on accessible images [11, 43]. This lack of inclusion in the design process may impact the quality and type of information that is needed to adequately interpret graphical figures without the use of vision. For example, these descriptions often lose contextual and spatial information that may convey critical insights for creating an accurate mental model of the graphic necessary for subsequent interpretation and recreation (i.e., axes, legends, nesting of figures, size, chart element configuration, and ordering). Our goal in this experiment is to investigate the efficacy of accessible NL descriptions that provide both high/low level detail within a structured description and also convey spatial information in the graph as well as the numerical data. The rationale for this focus on conveying spatial information is to create NL descriptions (cognitive processing) that are ideally situated to work together with other sensory channels in multimodal devices to create improved conceptualization of both the graph content and the graph context. There are many other sensory channels that might help to communicate graph information in a non-visual format, however, we turn to the use of haptics (perceptual processing) to convey graphical information, as we think this is the most promising and prevalent compared to NL descriptions.

2.2 Haptic/Tactile Literature

The teaching of graphical information is often focused on the visualization of a dataset and the numeric representation's organization and spatial structure (e.g., categories of a Venn diagram, clusters in a scatter plot, or comparing the total values in a series of bar charts) [37]. As spatial perception and reasoning can be efficiently supported through touch, learning aids and assistive technologies that translate visual information into non-visual haptic/tactile information using object attributes (e.g., vibrations, textures, points, edges, or holes) have been used extensively in a large body of research (for discussion, see [24, 34, 35]). Strategies for accessing information include: preliminary scanning and exploration, count organizing, and space partitioning to track elements by using one hand to indicate a partition and the other to explore the rest of the diagram. However, the use of touch requires more planning, working memory, and concentration than vision, so it is often a slower method and increases the cognitive processing load for BLV learners [3, 5, 41]. Therefore, graphical teaching using haptic/tactile learning materials need to be clearly structured to emphasize salient features representing the numerical and spatial content. Recent studies suggest that well-designed haptic vibration representations are effective in accurately perceiving and following lines and graphs, graph shapes, and graph patterns [24, 25, 28, 48].

While there are other approaches for rendering tactile graphics, they require expensive specialized equipment to create force feedback or pin array interfaces [45]. In some cases, such systems have been used to provide combined tactile and audio access to graphical information such as charts, graphs, and maps [12, 36, 62, 70, 70]. There are also a number of mobile applications utilizing commercially available smart devices such as a phone or tablet to augment the capabilities of traditional tactile graphics. These multimodal systems augment embossed tactile graphics by placing them on top of the tablet's screen and adding supplemental auditory information to the tactile overlay [63, 66, 71].

Our study uses a mobile application, based on embedded haptics instead of overlays, to compare haptic representations (perceptual processing) using smartphone haptic engines along with a structured NL overview (cognitive processing) to better understand how layering multisensory information may help in the interpretation of graphical content and spatial context over NL descriptions alone.

3 MULTIMODAL DIAGRAM SYSTEM

There are many benefits of adopting a multimodal design approach over the unimodal UIs (NL or haptic) used in most accessible technologies. Multimodal interfaces (and their affordances) maximize the brain's inherent ability for multisensory information processing, have the greatest potential to support the most people across ability/disability, age, and learning style, and are most consistent with how we actually experience and interpret the world through multiple channels of information (see [23, 64] for review). In other words, multimodal UIs represent an excellent inclusive design approach for all but are particularly important for BVI accessibility when learning graphically-oriented STEM content, as is studied here, which often combines semantic and spatial information in a single image [17].

The Multimodal Diagram System (MDS v 2.0) described in this paper, is an adaptation and extension of the touchscreen system developed by Giudice et al. [24], with the notable difference in OS platform, changing from an Android touchscreen tablet to an iOS touchscreen smartphone, as was initially done in a study with the MDS by Doore and colleagues [17]. This change in device was motivated by the large BLV user base (over 70%) that prefer Apple iOS devices [69] and the native accessibility features that come standard in all iOS devices. The use of this touchscreen technology leverages built-in haptic engine hardware (e.g., vibration motors) and affords portability, cost-effectiveness, and customizability.

4 STUDY PURPOSE

When developing new interfaces and systems, receiving feedback from stakeholders and potential users is a critical step in the iterative design process. Collecting information on the users' performances and their attitudes about the multimodal interface and non-visual graphical access methods helps to identify the next steps in the system development. This feedback can be analyzed for possible interactions between the type of graph experienced and the modality used to access the presented information. In this study, we explored the viability of two non-visual graphical presentation modalities with a group of participants with blindness and low vision who were recruited to help with the testing and feedback process. Using an Apple iPhone 12 Pro Max, we compared the effectiveness of vibro-audio exploration to natural language descriptions when accessing three types of informational graphics (line graphs, pie charts, Venn diagrams). We chose these three types of graphs because they are some of the most frequent diagrams found in popular press, textbooks, and accessible diagram studies [32, 59].

Line graphs represent a series of data points connected by straight line segments and ordered points. The line can be either straight or curved depending on the data being presented. Line graphs are one of the simplest forms of representing quantitative data between two variables. They are often used to represent a trend in temporal data where the ordering of x-axis data are drawn chronologically.

Pie charts are circular diagrams used to compare parts of the whole, dividing the whole circle into sectors that are equal in proportion to the quantity represented. Studies have shown that proportions in pie charts were judged more accurately than with several other basic diagram forms for simple evaluation of proportions and had a distinct advantage when more complicated part whole judgements were required [58, 60].

Venn Diagrams are simple closed curves drawn on a plane to represent basic relationships (similarities and differences) between sets. They are classified as a type of Euler diagram that has two or more data sets and emphasize explicitly the nesting between sets. They can be particularly helpful in investigating how individuals arrive at an inference [53] and have been used extensively in previous research on strategies that individuals (both sighted and BLV) employ when solving reasoning problems about sets of data [52, 72].

In addition to these three types of graphs being commonplace, the communicated information and concepts between the graphs are unique and different from each other. For example, line graphs communicate data trends and data point comparisons, pie charts communicate proportionality and value estimation of regions, and Venn diagrams communicate the concept of set relationships. Importantly, the skills needed to learn, interpret, and reason on these graphs are consistent and are not dependent on perceived complexity. This is why we chose to design simple graphs for this study that can be used to clearly gauge the information access ability of the presentation modalities rather than using pre-made graphics from a STEM textbook.

5 METHODS

5.1 Participants

Nineteen participants (7 female; 12 male; 19 - 69 years; SD = 14.63) were recruited from the 2022 NFB (National Federation of the Blind) annual convention and from The Carroll Center for the Blind (Boston, MA) through email and phone. Participants were all legally blind with vision ranging from no usable vision to light perception and minimal functional vision. It was noted that every participant had some previous experience with the graph types tested in the user study; many of them stating they learned these through school or work. A complete summary of participant demographics can be found in Table 1A in Appendix A. This study was accepted and approved by the Institutional Review Board at The University of Maine.

5.2 Presentation Modalities and Stimuli

The two presentation modalities used in the study were natural language descriptions (NL) and vibro-audio exploration with an auditory overview (VA). NL uses speech processing to describe the graph in its entirety to the user through the Voiceover feature in iOS. The description is activated with a single tap on the touchscreen. The system then reads the description of the graph, starting with a more general overview (the same overview used in the VA condition) and then adds more detail, generally describing features left to right as they are visually rendered. The descriptions were developed based on several factors: 1) previous research informed by an analysis of graphic descriptions provided by secondary and post-secondary math instructors (i.e., experts) for learners who could not see the diagram [16]; 2) current models for descriptions of graphics that provide leveled information [32, 39]; and 3) ensuring that the NL description mirrored the same type of information that was provided in the VA condition, as carefully matching the information between conditions is critical for valid multimodal comparisons [38]. For example, a Venn diagram in NL (Figure 1) would be described as:

> "This is a Venn diagram that uses three circles and two intersections to represent relationships among three sets of numbers. The circle on the left is labeled A and

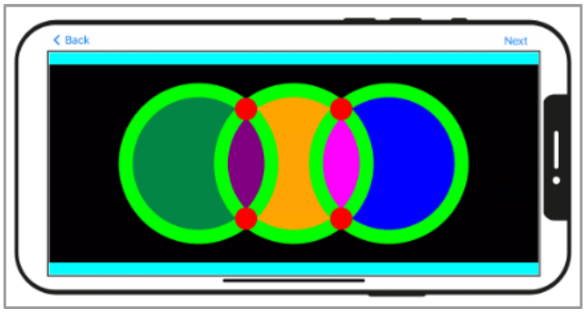


Figure 1: A Venn diagram with three circles as it would appear on the iPhone during the user study.

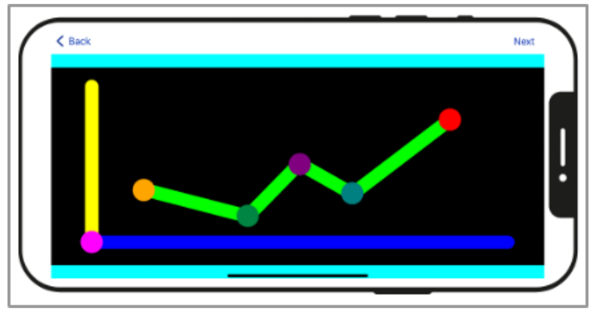


Figure 2: A line graph with five data points as it would appear on the iPhone during the user study.

> represents even numbers. The circle in the middle is labeled B and represents multiples of 3 that are less than 100. The circle on the right is labeled C and represents odd numbers."

This NL technique gives the user a complete description of the graph all at once (overview), similar to reading the alt text on well described graphics. It also provides the spatial configuration of the individual elements, their location in the screen space (on the left/right, in the middle), and the data associated with each element. The description also employs instructional language that is key to understanding the spatial and set relationships between the graph elements (intersections). In the Venn diagram described above, the intersection of circles A and B, as well as the intersection of circles B and C, represent the relationships of even multiples of 3 and odd multiples of 3, respectively. Although circle B on its own does not contain any numbers because all multiples of 3 are either even or odd, including this empty circle in the NL description is important for forming the set relationships (intersections) with the other circles. The NL description could contain additional information (e.g., high contrast colors for individuals with residual vision), however, in order to test the efficacy of each condition in this study, the description needed to be limited to what all of the participants could perceive in the VA condition and nothing more. Appendix B contains the descriptions for all of the graphs used in the study.

The VA utilizes vibration patterns and short natural language cues that are activated when a user's fingertip reaches specified regions on the touchscreen device. When the screen is tapped once, the system provides a brief natural language overview to provide context and help guide the user's search behavior. As shown in Figure 2, an overview of the practice line graph used in the study would be, "This is a line graph with 5 data points and 4 segments." When the user places their finger on a point on the line graph, the system produces a unique pulsing vibration pattern (to convey a point rather than a line) and says the data point's label and coordinates (e.g., Point D; 10, 2). When the user moves to a line segment, the system tells the user which line segment they are on, such as, "line segment AB." The user will continue to feel steady vibrations while they are on the line and cease to feel them when they are off the line, enabling them to trace the line segment to the next point. This VA technique leverages multimodal and spatial components to give users the ability to target specific points of interest on a graph and to quickly access and perceive the information, producing an interactive equivalent to traditional embossed tactile images with Braille labels. The cues used here were based on a recent set of design guidelines set forth for optimizing the presentation of vibrotactile graphical content [26, 46]. A complete list of descriptions for both presentation modalities for all three graph types can be found in Appendix B.

Each participant experienced four graphs per graph type, thus experiencing a total of twelve graphs throughout the study. For each graph type, two graphs were used for the participant to practice with, one for each presentation modality (VA, NL). The other two graphs were used in the experimental condition, one for each modality. Only the graphs within the experimental conditions were used for the formal analysis (six total graphs, one for each modality and graph type combination).

The questions presented to the participants targeted different facets of graph complexity to measure comprehension. The first two questions were criterion questions used to assess general understanding of the presentation modalities and to ensure a common baseline for all participants before moving on to the more complicated test questions. If either question was answered incorrectly, the application would instruct the participant to try and learn the graph again. The remaining four test questions were the only questions used in the subsequent analysis. Through pilot testing, we ensured that every question within the experiment could be answered with the information given to the participants through each modality. Each question was multiple choice, with four or five answer options that the participant would navigate and select using the VoiceOver gestures described to them. Participants were allowed to move back and forth between a question and the graph (i.e., answers were not reliant on memory of the learned graph) but once an answer was submitted, they could not go back to that question. A complete list of questions can be found in Table 1C in Appendix C.

5.3 Protocol

A within-subjects design was employed for this study, which explored the two non-visual graphical presentation modalities, vibrotactile with auditory overview (VA) and natural language descriptions (NL). The three graph types used were Venn diagrams, pie charts, and line graphs. All participants received the graphs in a fixed order (informed by pilot testing based on complexity): 1) Venn diagrams; 2) pie charts; 3) line graphs. The experiment was counterbalanced such that each modality was experienced first by half of the participants (e.g. NL then VA, versus VA then NL).

Each graph had six accompanying multiple-choice questions. The first two questions were used as criterion questions while the remaining four questions were of increasing difficulty based on the degree of interpretation and inference that was required to answer the question (refer to Appendix C). An Apple iPhone 12 Pro Max was used to facilitate the experiment; we developed a multimodal mobile application to display the graphs, initiate the non-visual graphical access method, and record participant question responses (stored on a secure server). Apple Inc.'s VoiceOver [4] was used to navigate the application.

5.4 Procedure

A single researcher proctored the study. Before starting, participants were asked to fill out a consent form, intake survey, and could ask questions about the study if they had any. Once complete, the researcher gave the participant control of the iPhone with the experimental app loaded, VoiceOver enabled for navigation, and began the experiment which took approximately 1 hour. All participants wore a Mindfold brand eye covering as a control for any potential impacts of residual vision between participants.

Each participant began with a practice Venn diagram and used either the vibro-audio or natural language modality to learn how the presentation modality worked with that graph type. After learning the presentation modality, addressing any questions, and completing the practice, the participant progressed to the experimental trial for that modality with a new, more complex Venn diagram. The experimental trial was completed once the participant correctly answered the two criterion questions and then the subsequent four test questions relating to the graph. The four test questions were randomized between participants and were not repeatable.

After the first experimental Venn diagram was completed, the participant learned how to use the remaining presentation modality with another practice Venn diagram. After the practice Venn diagram, the participant would interpret a new Venn diagram similar in complexity to the experimental graph used with the first modality and follow the same procedure as above; answering the two criterion questions, followed by the four experimental questions. This concluded the Venn diagram trial and the participant then repeated this process with the pie charts and the line graphs to complete the study. Following the study, the researcher uploaded the participant's experimental data to the secure server, issued a post-study survey to collect qualitative feedback about their experience with the system, then compensated the participant for their time and input.

There were two independent variables (IVs) in this study: IV1 = non-visual presentation modality (e.g., the NL condition vs. the VA condition), and IV2 = graph type (e.g., Venn diagrams, pie charts, and line graphs). There were 3 primary dependent variables (DVs) measured for the study: DV1 = accuracy in answering the test questions, DV2 = total time needed to complete an experimental graph, DV3 = user satisfaction with system performance, individual features, and efficacy; perception of difficulty of tasks based on modality and graph type. A two-way repeated measures ANOVA based on an alpha of 0.05 was used to compare the within-subjects measures (accuracy and time) between each presentation modality and graph type.

6 RESULTS

Nineteen BLV participants participated in the study. However, only eighteen participants produced data that were usable in the statistical analysis as one participant did not pass a majority of the criterion questions. From the remaining eighteen participants who passed the criterion questions, five participants needed a second attempt to submit a correct answer (one participant for a Venn diagram, one participant for a pie chart, and three participants for the line graph). A two-way repeated measures ANOVA was used to analyze the participants' total accuracy based on discrete correctness of the multiple choice test questions for each presentation modality and graph type pairing (24 total questions). The average total accuracy for this group of participants across all modalities and graphs was 65.5% (SE = 5.14%). A breakdown of the following accuracy and timing results can be found in Table 1 below.

6.1 Accuracy

Within the vibro-audio condition, participants had an average accuracy of 63.89% (SE = 5.80) across all graph types. By graph, participants averaged 84.72% (SD = 21.39%) on Venn diagrams, 47.22% (SD = 30.78%) on pie charts, and 59.72% (SD = 36.52%) on line graphs; these mean accuracies for both conditions are shown visually in the bar graph, Figure 3. Within the natural language condition, participants scored, on average, 67.13% (SE = 4.97%) correct across all graphs. By graph, participants averaged 86.11% (SD = 21.39%) on Venn diagrams, 54.17% (SD = 24.63%) on pie charts, and 61.11% (SD = 35.59%) on line graphs. Although the natural language modality had a slightly higher numeric average than the vibro-audio modality (a difference of 3.24%), the ANOVA test did not reveal a statistically significant effect of presentation modality on participant accuracy (F(1,17) = 0.959, p = 0.341, η_p^2 = 0.053). In fact, when comparing the graphs between the two modalities, there were no significant differences between any of the graphs as a function of presentation modality: Venn diagrams (F(1,17) = 0.106, p = 0.749, η_p^2 = 0.006), pie charts (F(1,17) = 2.033, p = 0.172, η_p^2 = 0.107), or the line graphs (F(1,17) = 0.027, p = 0.871, η_p^2 = 0.002). There was, however, a significant main effect of graph type on participant accuracy (F(2,16) = 15.338, p < 0.001, η_p^2 = 0.657). Put simply, the graph types in both conditions showed similar participant accuracies when compared to their counterparts but the accuracies between the graphs, irrespective of presentation modality, were significantly different. This may be in part due to the participants' familiarity and previous exposure with the three graph types or that some graphical concepts are more difficult to learn, regardless of the presentation modality used. In the post-study survey, all participants mentioned having prior experience with the graphs in the user study; fifteen participants experienced these within an academic setting and nine participants experienced these in their past or current occupations. In summary, results from the two presentation modalities tested demonstrated that both are comparable in accessing the graphical information and enabling the participants to answer the questions with similar accuracy.

Table 1: Reports of mean accuracy and time for Venn diagrams, pie charts, and line graphs with vibro-audio and natural language modalities

Modality	Measure	Venn Diagram	Pie Chart	Line Graph	Overall Mean
Vibro-Audio	Accuracy	84.72%	47.22%	59.72%	63.89%
	Time (min:ss)	5:07	5:47	8:34	6:29
Natural Language	Accuracy	86.11%	54.17%	61.11%	67.13%
	Time (min:ss)	4:37	5:30	7:05	5:44

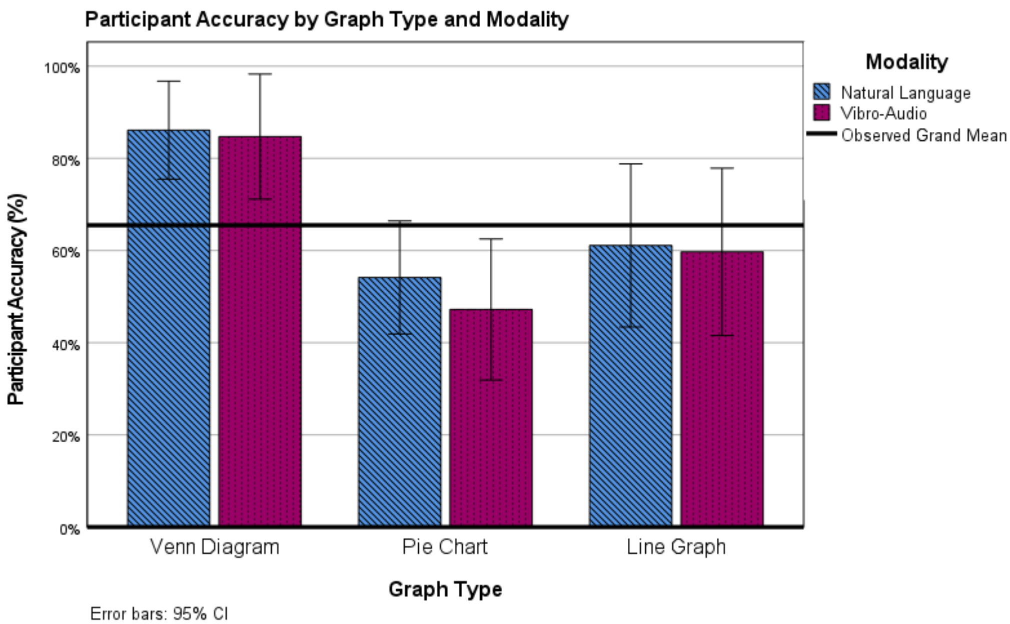


Figure 3: This is a clustered bar graph showing participant accuracy by graph and modality

6.2 Timing

Timing data were also collected to examine the speed at which the participants could learn the graphs and answer the associated questions. We were curious if the modality used to interpret the graphs would affect the participants' learning time. In the vibro-audio condition, the average time to complete a graph (given as minutes:seconds) was 6:29 (SE = 22.8s). By graph, participants averaged 5:07 (SD = 1:47) on Venn diagrams, 5:47 (SD = 1:35) on pie charts, and 8:34 (SD = 3:50) on line graphs. In the natural language condition, the average time to complete a graph was 5:44 (SE = 27.9s). By graph, participants averaged 4:37 (SD = 1:37) on Venn diagrams, 5:30 (SD = 2:56) on pie charts, and 7:05 (SD = 3:04) on line graphs; these results are shown visually in the bar graph, Figure 4. From the collected timing data, it was found that modality had a significant main effect on participant completion time (F(1,17) = 4.986, p = 0.039, η_p^2 = 0.227). Put simply, performance in the NL modality demonstrated faster average completion time than in the VA modality. The type of graph also had a significant main effect on the participants' time (F(2,16) = 7.21, p = 0.006, η_p^2 = 0.474). This means that the graphs, regardless of the modality used to learn them, had significantly different completion times among participants. It is clear that Venn diagrams were the fastest, followed by pie charts, and then line graphs with the slowest times. These findings may suggest that graphs that are more conceptually complex, like the line graph, take participants longer to learn, interpret, and rationalize. Furthermore, using a mode of access that is more familiar (i.e., natural language) may explain the decreased average completion time of the participants.

6.3 Post-Study Survey

After participants completed the experiment, they were asked a series of questions on a 7-point Likert scale pertaining to the difficulty of learning the three graphs with both modalities (1 = not difficult at all, 7 = extremely difficult). Figure 5 below displays the participants' ratings within a clustered bar graph. For the vibro-audio modality, participants rated Venn diagrams to be the easiest (M = 2.06, SD = 0.80), followed by pie charts (M = 4.17, SD = 1.47), and then line graphs (M = 5.44, SD = 1.58). For the natural language modality, participants had the same trend, rating Venn diagrams to be the easiest (M = 1.83, SD = 1.34), followed by pie charts (M = 3.0, SD = 1.41), and then line graphs (M = 5.5, SD = 0.86).

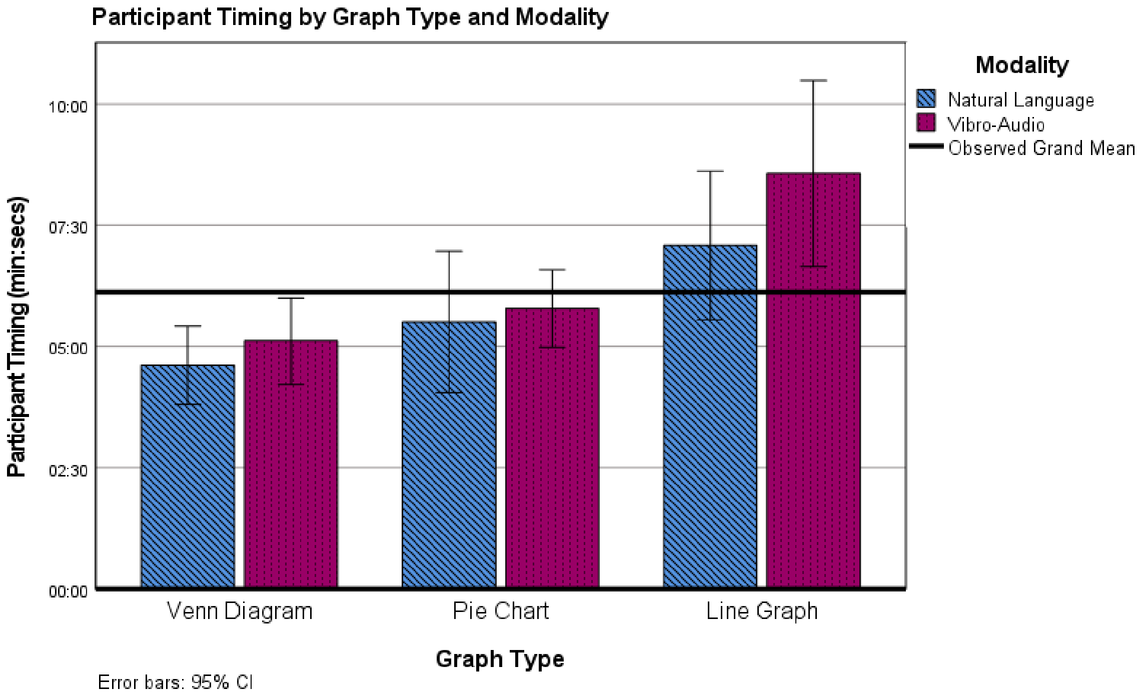


Figure 4: This is a clustered bar graph showing participant timing by graph type and modality

The clarity of information provided by each presentation modality as well as the instructions on using the learning system were also rated by participants on a 7-point scale (1 = strongly disagree, 7 = strongly agree). The clarity of the natural language modality (M = 6.06, SD = 1.00) obtained a stronger agreement amongst participants than the clarity of the vibro-audio modality (M = 4.83, SD = 1.43). The instructions on how to use the learning system were very clear for participants (M = 6.44, SD = 0.86). With respect to participants' overall preference for presentation modality, two-thirds of the participants (N =12) preferred the natural language modality over the vibro-audio modality.

7 DISCUSSION

This research was motivated by the development of a low-cost multimodal diagram system for assisting BLV individuals in accessing and learning graphical content commonly used during educational and vocational instruction. This issue is of growing societal concern given the ubiquity of inaccessible graphical STEM content that exists at every level of education and workplace training. Our findings comparing two non-visual graphical presentation modalities implemented into our diagram system contribute to the growing body of literature investigating the efficacy of multimodal solutions for providing BLV individuals greater access to this information, thereby increasing opportunities and successful entry into the STEM workforce.

The similarity of findings observed in our study (i.e. statistically insignificant differences on participant accuracy between the two presentation modalities) provide preliminary support that vibro-audio and natural language techniques are able to support functionally equivalent behavior when learning basic, yet commonly used, STEM graphs and diagrams. However, the type of graph experienced had a significant effect on the participants' accuracy and learning time with both modalities. Although we asked participants about their prior experience with the three types of graphs after they concluded the study, our questions were not sufficiently specific to further analyze the impact of prior graph experience (or the modality used with previous exposure) on current performance. It is noteworthy that all nineteen participants stated that they had experience with these graphs, whether it was in primary, secondary, and post-secondary school or through their past or current occupations. Future studies will further probe this issue, as it will facilitate additional analyses about whether the best performers were those with the most experience with the tested graph types or how prior experience may have influenced performance with a particular presentation modality.

7.1 Natural Language Condition

Despite the similarity in empirical performance, most participants (N = 12) preferred graphs conveyed through natural language. However, several participants (N = 5) said that they would like both presentation modalities as an option — supporting the benefits of a multimodal design offering users the ability to change how information was presented based on their preference. Utilizing the built-in features of the iPhone that run our learning system, it would not be a challenge to further develop the system to switch learning

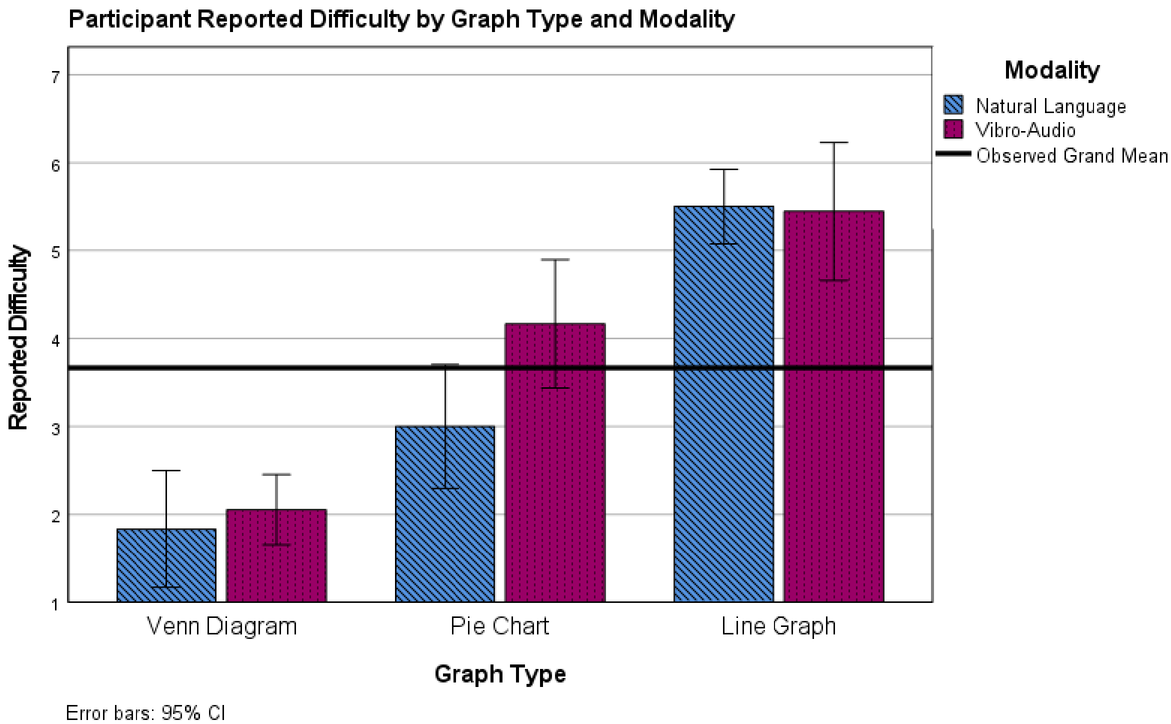


Figure 5: This is a clustered bar graph showing participants reported difficulty by graph type and modality. 1 = Not difficult at all and 7 = Extremely difficult

methods with a designated gesture on the touchscreen. This would ultimately give the end user more control over the learning style that works best for them as a function of the information being consumed. Such flexibility in accessible content is missing in many assistive technologies but is actually very important as it provides a critical inclusive design focus on giving BLV users options for customizing and combining different content levels based on their individual needs and preferences rather than the designers default decisions.

The NL descriptions used here were based on current guidelines for creating accessible diagrammatic information [9, 15, 32] and required substantial time for the researchers to curate. The NL descriptions started with a brief overview of the category of the diagram that was identical to the haptics NL overview. The full description was carefully designed to mirror the information that could be conveyed haptically. Ensuring this information-matched comparison is critical for evaluating the efficacy of these two non-visual interface solutions [38]. Lunguard and Satyanarayan (2022) [39] have made the argument that natural language serves as both an interface to interpret diagrams as well as an interface to the actual data. There is certain semantic content that may be more efficiently communicated with natural language (e.g., values, statistics, explanations) over other diagram information that is inherently spatial in nature (e.g., clusters, proportions, sets), which may be more easily understood through haptic channels [11, 50].

7.2 Vibro-Audio Condition

The vibro-audio rendering was praised by participants for providing them rapid access to specific information. However, some participants commented that this technique needed further development to increase its usability. This sentiment may be driven by lack of familiarity. NL information is commonly used (e.g., screen readers) while vibrotactile feedback to access information (graphical or otherwise), as was used in this study, is a new interaction style. While many of our participants had experience with using traditional embossed tactile graphics, which are generally explored using two hands on hardcopy media, our experimental system is quite different as it is based on touch-based vibrotactile cues explored using only one finger. Multi-finger and two-handed exploration of the graphs is not possible using this method because 1) the haptic effects are driven by one point of contact with the screen based on the single vibration motor in the device and 2) multi-finger gestures are used to control the system interface behavior. Given the novelty of this interface, the results showing comparable learning accuracy performance to the natural language gold standard for graph accessibility speak to the efficacy and future potential of this solution. We imagine that with more experience and time to become accustomed to the vibro-audio condition, its potential advantages for directly conveying spatial content may be more evident.

In future studies, a tutorial could be implemented to introduce the vibro-audio interface and ensure that participants fully understood how this interaction style can be used to best extract information and learn the structure of the graphs. We also feel that future studies would benefit from inclusion of a reconstruction task. This type of task best reflects the accuracy of a participant's mental representation built up from learning, as successful graph recreation requires accurate representation of spatial relations. It is possible that the spatial advantages of our VA interface vs a purely NL-based interface may be more evident, especially when learning complex graphical content, when comparing performance on a graph recreation task.

7.3 Limitations

Each presentation modality, as it is currently implemented, has its own advantages and limitations that may serve or hinder a user's success depending on their preference. Natural language uses a communication pathway that is familiar to participants and may be preferred for information extraction. However, the descriptions can be verbose and lengthy which can make it difficult to target specific information without repetition, especially for complex graphical content. The NL descriptions are read in their entirety once activated and can be stopped at any time. However, once stopped, the description resets. If a piece of information is not understood by the user, they would (in the current system) need to replay the entire description. This can be frustrating, especially if the description takes a substantial amount of time to read and the user is trying to get to a specific piece of the description. This was noted by multiple participants (N = 10) in the post study survey. It was suggested that the natural language descriptions should be separated by sentence, so that users could employ a gesture to skip through the description to reach the target information, e.g., swiping between functionally-relevant description chunks. Further development will be needed to successfully implement this constructive feedback.

Vibro-audio learning techniques provide direct access to spatial information but have limitations imposed by the current technology. The relatively small size of the iPhone screen constrains the size and detail of the graphs that can be rendered within a single viewport. This was important, as we wanted the entire image to fit within one display. Besides screen size, this limitation is based on the perceptual resolution of vibrotactile stimuli, which has lower acuity than traditional embossed tactile stimuli. Palani and colleagues have conducted extensive usability and psychophysical work defining the appropriate parameters for vibrotactile rendering and found that using vibro-tactile stimuli on a touchscreen, as was done here, requires lines to be rendered with a minimum width of 4mm [46]. Thus, any use of this approach for multisensory renderings employing vibro-tactile stimuli, as we did here with the vibro-audio condition, must account for these factors when considering the display. An iOS device with a larger screen would, in theory, be the simplest option to increase the detail of the graphs without sacrificing line width. Unfortunately, none of the larger form factor iPads to date have vibration motors. These limitations must be considered against the many benefits of using the iPhone, e.g., being incredibly portable, multisensory, and a platform that most BLV users already own.

Note that rendering larger graphics beyond the viewport of current vibration-equipped phones is possible but would require using standard image manipulation techniques like panning and zooming. These are common spatial operations for visual image manipulations but not frequently used by BLV people in nonvisual context. However, there have been some promising techniques developed for performing non-visual panning and zooming with maps using a similar vibro-audio interface as was tested here [47, 49]. Future research could extend these techniques to support other types of nonvisual and multisensory graphics, as were studied here. In the meantime, our efforts are focused on improving the system as it currently runs on the iPhone.

8 CONCLUSION

The user study elaborated in this paper examined the implementation of two graphical presentation modalities as part of a multimodal diagram system and compared their success at conveying graphical information to members of the BLV community. Venn diagrams, pie charts, and line graphs were used as stimuli to test the usability of the system and explore the effectiveness of two nonvisual interaction modalities as viable access solutions for common types of graphical content. It was found that both vibro-audio and natural language presentation modalities were equally viable for learning the graphical information presented. However, natural language was preferred by more of the participants and was slightly faster than using vibro-audio rendering. Given that exploration/learning using vibro-audio information was a new experience for most of the BLV participants, it is likely that with more practice this approach could serve as a viable general-purpose tool for tasks that require learning, interpreting, or representing complex graphical information, especially tasks involving the interpretation of spatial concepts. These findings also put an emphasis on designing multimodal systems that can be inclusive, adaptive, and support different users' learning preferences. Based on our results showing overall success for accessing graphical content non-visually with the NL and VA approaches, we believe that employing multiple channels of access in future user interfaces will ultimately provide the best performance and support a greater audience than standard unimodal interfaces. This multimodal diagram system, leveraging two interaction modalities in one, low-cost platform, is but another example of a practical solution for replacing the time-exhaustive and expensive tactile graphics used as the standard in current BLV instruction. With further development of the system and refinements to the interaction modalities, we predict that information accessibility, system usability, and graphical interpretation will significantly increase for inclusive STEM access in educational and vocational contexts.

ACKNOWLEDGMENTS

We greatly appreciate the participants involved in the user study for their time and valuable contributions to our research. This work is supported by the National Science Foundation under Grants DRL 2048394 and IIS 1822800.

REFERENCES

[1] Md Zubair Ibne Alam, Shehnaz Islam, and Enamul Hoque. 2023. SeeChart: Enabling Accessible Visualizations Through Interactive Natural Language Interface For People with Visual Impairments. In *Proceedings of the 28th International Conference on Intelligent User Interfaces* (IUI '23), Association for Computing Machinery, New York, NY, USA, 46–64. DOI:https://doi.org/10.1145/3581641.3584099

[2] American Printing House for the Blind (APH). 2019. *Annual Report 2019*. American Printing House for the Blind, Inc., Louisville, KY. Retrieved from https://www.aph.org/annual-reports/
[3] Yiannoula Andreou and Steve McCall. 2010. Using the voice of the child who is blind as a tool for exploring spatial perception. *British Journal of Visual Impairment* 28, 2 (May 2010), 113–129. DOI:https://doi.org/10.1177/0264619609360285
[4] Apple Inc. 2022. VoiceOver. *Accessibility - Vision*. Retrieved from https://www.apple.com/accessibility/vision/
[5] Soledad Ballesteros, Dolores Bardisa, Susanna Millar, and Jose M. Reales. 2005. The haptic test battery: A new instrument to test tactual abilities in blind and visually impaired and sighted children. *British Journal of Visual Impairment* 23, 1 (January 2005), 11–24. DOI:https://doi.org/10.1177/0264619605051717
[6] Bernhard Beck-Winchatz and Mark A. Riccobono. 2008. Advancing participation of blind students in Science, Technology, Engineering, and Math. *Advances in Space Research* 42, 11 (December 2008), 1855–1858. DOI:https://doi.org/10.1016/j.asr.2007.05.080
[7] Edward C. Bell and Arielle M. Silverman. 2019. Access to Math and Science Content for Youth Who Are Blind or Visually Impaired. *Journal of Blindness Innovation and Research* 9, 1 (2019). DOI:https://doi.org/10.5241/9-152
[8] Brianna Blaser and Richard E. Ladner. 2020. Why is Data on Disability so Hard to Collect and Understand? In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, IEEE, 1–8. DOI:https://doi.org/10.1109/respect49803.2020.9272466
[9] Braille Authority of North America and Canadian Braille Authority. 2010. Guidelines and Standards for Tactile Graphics. Retrieved from https://www.brailleauthority.org/tg/web-manual/index.html
[10] Sandra Carberry, Stephanie Elzer Schwartz, Kathleen Mccoy, Seniz Demir, Peng Wu, Charles Greenbacker, Daniel Chester, Edward Schwartz, David Oliver, and Priscilla Moraes. 2013. Access to multimodal articles for individuals with sight impairments. *ACM Trans. Interact. Intell. Syst.* 2, 4 (January 2013), 21:1-21:49. DOI:https://doi.org/10.1145/2395123.2395126
[11] Jinho Choi, Sanghun Jung, Deok Gun Park, Jaegul Choo, and Niklas Elmqvist. 2019. Visualizing for the Non-Visual: Enabling the Visually Impaired to Use Visualization. *Computer Graphics Forum* 38, 3 (June 2019), 249–260. DOI:https://doi.org/10.1111/cgf.13686
[12] Andrew Crossan and Stephen Brewster. 2008. Multimodal Trajectory Playback for Teaching Shape Information and Trajectories to Visually Impaired Computer Users. *ACM Transactions on Accessible Computing* 1, 2 (October 2008), 1–34. DOI:https://doi.org/10.1145/1408760.1408766
[13] Frances R. Curcio. 1987. Comprehension of Mathematical Relationships Expressed in Graphs. *Journal for Research in Mathematics Education* 18, 5 (1987), 382–393. DOI:https://doi.org/10.2307/749086
[14] Nada Dabbagh and Anastasia Kitsantas. 2012. Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *The Internet and Higher Education* 15, 1 (January 2012), 3–8. DOI:https://doi.org/10.1016/j.iheduc.2011.06.002
[15] DIAGRAM Center. 2019. Image Description Guidelines. *DIAGRAM Center*. Retrieved from http://diagramcenter.org/table-of-contents-2.html/
[16] Stacy A. Doore, Justin K. Dimmel, Rose Xi, and Nicholas A. Giudice. 2021. Embedding expert knowledge: A case study on developing an accessible diagrammatic interface. In *Proceedings of the forty-third annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Philadelphia, PA, 1759–1763.
[17] Stacy A. Doore, Justin Dimmel, Toni M. Kaplan, Benjamin A. Guenther, and Nicholas A. Giudice. 2023. Multimodality as universality: Designing inclusive accessibility to graphical information. *Frontiers in Education* 8, (April 2023). DOI:https://doi.org/10.3389/feduc.2023.1071759
[18] Frances Dowen, Karishma Sidhu, Elizabeth Broadbent, and Helen Pilmore. 2017. Communicating projected survival with treatments for chronic kidney disease: patient comprehension and perspectives on visual aids. *BMC Med Inform Decis Mak* 17, (September 2017), 137. DOI:https://doi.org/10.1186/s12911-017-0536-z
[19] Carolien Duijzer, Marja Van den Heuvel-Panhuizen, Michiel Veldhuis, and Michiel Doorman. 2019. Supporting Primary School Students' Reasoning about Motion Graphs through Physical Experiences. *ZDM* 51, 6 (2019), 899. DOI:https://doi.org/10.1007/s11858-019-01072-6
[20] Robert Wall Emerson and Dawn Anderson. 2018. What Mathematical Images Are in a Typical Mathematics Textbook? Implications for Students with Visual Impairments. *Journal of Visual Impairment & Blindness* 112, 1 (January 2018), 20–32. DOI:https://doi.org/10.1177/0145482x1811200103
[21] Leo Ferres, Gitte Lindgaard, Livia Sumegi, and Bruce Tsuji. 2013. Evaluating a Tool for Improving Accessibility to Charts and Graphs. *ACM Trans. Comput.-Hum. Interact.* 20, 5 (November 2013), 28:1-28:32. DOI:https://doi.org/10.1145/2533682.2533683
[22] Giovanni Fusco and Valerie S. Morash. 2015. The Tactile Graphics Helper: Providing Audio Clarification for Tactile Graphics Using Machine Vision. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (ASSETS '15), Association for Computing Machinery, New York, NY, USA, 97–106. DOI:https://doi.org/10.1145/2700648.2809868
[23] Nicholas A. Giudice. 2018. Navigating without vision: principles of blind spatial cognition. *Handbook of Behavioral and Cognitive Geography* (2018), 260–288. DOI:https://doi.org/10.4337/9781784717544.00024
[24] Nicholas A. Giudice, Hari Prasath Palani, Eric Brenner, and Kevin M. Kramer. 2012. Learning non-visual graphical information using a touch-based vibro-audio interface. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '12*, ACM Press, Boulder, Colorado, USA, 103. DOI:https://doi.org/10.1145/2384916.2384935
[25] Cagatay Goncu and Kim Marriott. 2011. GraVVITAS: Generic Multi-touch Presentation of Accessible Graphics. *Human-Computer Interaction – INTERACT 2011* (2011), 30–48. DOI:https://doi.org/10.1007/978-3-642-23774-4_5
[26] Jenna L. Gorlewicz, Jennifer L. Tennison, P. Merlin Uesbeck, Margaret E. Richard, Hari P. Palani, Andreas Stefik, Derrick W. Smith, and Nicholas A. Giudice. 2020. Design Guidelines and Recommendations for Multimodal, Touchscreen-based Graphics. *ACM Transactions on Accessible Computing* 13, 3 (August 2020), 1–30. DOI:https://doi.org/10.1145/3403933
[27] Nora Griffin-Shirley, Devender R. Banda, Paul M. Ajuwon, Jongpil Cheon, Jaehoon Lee, Hye Ran Park, and Sanpalei N. Lyngdoh. 2017. A Survey on the Use of Mobile Applications for People Who Are Visually Impaired. *Journal of Visual Impairment & Blindness* 111, 4 (2017), 307+.
[28] Michael E. Hahn, Corrine M. Mueller, and Jenna L. Gorlewicz. 2019. The Comprehension of STEM Graphics via a Multisensory Tablet Electronic Device by Students with Visual Impairments. *Journal of Visual Impairment & Blindness* 113, 5 (September 2019), 404–418. DOI:https://doi.org/10.1177/0145482x19876463
[29] William L. Heward. 2003. *Exceptional children: An introduction to special education*. Pearson Education, New Jersey.
[30] Crescentia Jung, Shubham Mehta, Atharva Kulkarni, Yuhang Zhao, and Yea-Seul Kim. 2022. Communicating Visualizations without Visuals: Investigation of Visualization Alternative Text for People with Visual Impairments. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (January 2022), 1095–1105. DOI:https://doi.org/10.1109/TVCG.2021.3114846
[31] Susan M. Katz. 2008. Creating Appropriate Graphics for Business Situations. *Business Communication Quarterly* 71, 1 (March 2008), 71–75. DOI:https://doi.org/10.1177/1080569907312877
[32] Nam Wook Kim, Shakila C. Joyner, Amalia Riegelhuth, and Yea-Seul Kim. 2021. Accessible Visualization: Design Space, Opportunities, and Challenges. *Computer Graphics Forum* 40, 3 (June 2021), 173–188. DOI:https://doi.org/10.1111/cgf.14298
[33] David M. Lane and Anikó Sándor. 2009. Designing better graphs by including distributional information and integrating words, numbers, and images. *Psychological Methods* 14, 3 (September 2009), 239–257. DOI:https://doi.org/10.1037/a0016620
[34] Susan J. Lederman and Roberta L. Klatzky. 1997. Relative availability of surface and object properties during early haptic processing. *Journal of Experimental Psychology: Human Perception and Performance* 23, 6 (1997), 1680–1707. DOI:https://doi.org/10.1037/0096-1523.23.6.1680
[35] Susan J. Lederman and Roberta L. Klatzky. 2009. Haptic perception: A tutorial. *Attention, Perception & Psychophysics* 71, 7 (September 2009), 1439–1459. DOI:https://doi.org/10.3758/app.71.7.1439
[36] Fabrizio Leo, Elena Cocchi, and Luca Brayda. 2017. The Effect of Programmable Tactile Displays on Spatial Learning Skills in Children and Adolescents of Different Visual Disability. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25, 7 (July 2017), 861–872. DOI:https://doi.org/10.1109/tnsre.2016.2619742
[37] Juliane Leuders. 2015. Tactile and acoustic teaching material in inclusive mathematics classrooms. *British Journal of Visual Impairment* 34, 1 (December 2015), 42–53. DOI:https://doi.org/10.1177/0264619615610160
[38] Jack M. Loomis, Roberta L. Klatzky, and Nicholas A. Giudice. 2013. Representing 3D Space in Working Memory: Spatial Images from Vision, Hearing, Touch, and Language. In *Multisensory Imagery*, Simon Lacey and Rebecca Lawson (eds.). Springer New York, New York, NY, 131–155. DOI:https://doi.org/10.1007/978-1-4614-5879-1_8
[39] Alan Lundgard and Arvind Satyanarayan. 2022. Accessible Visualization via Natural Language Descriptions: A Four-Level Model of Semantic Content. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (January 2022), 1073–1083. DOI:https://doi.org/10.1109/tvcg.2021.3114770
[40] Marketplace. 2017. Braille versions of textbooks help blind college students succeed. *Marketplace*. Retrieved May 2, 2023 from https://www.marketplace.org/2017/10/12/braille-versions-textbooks-help-blind-college-students-succeed/
[41] Susanna Millar. 2000. Modality and Mind: convergent active processing in interrelated networks as a model of development and perception by touch'. *Touch, representation and blindness* (2000), 99–141.
[42] Nathan W. Moon, Robert L. Todd, David L. Morton, and Emily Ivey. 2012. Accommodating Students with Disabilities in Science, Technology, Engineering, and Mathematics (STEM): Findings from Research and Practice for Middle Grades through University Education. Atlanta, GA: Center for Assistive Technology and Environmental Access, Georgia Institute of Technology.
[43] Priscilla Moraes, Gabriel Sina, Kathleen McCoy, and Sandra Carberry. 2014. Evaluating the accessibility of line graphs through textual summaries for visually impaired users. ACM Press. DOI:https://doi.org/10.1145/2661334.2661368

[44] Meredith Ringel Morris, Jazette Johnson, Cynthia L. Bennett, and Edward Cutrell. 2018. Rich Representations of Visual Content for Screen Reader Users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), Association for Computing Machinery, New York, NY, USA, 1–11. DOI:https://doi.org/10.1145/3173574.3173633

[45] S. O'Modhrain, N. A. Giudice, J. A. Gardner, and G. E. Legge. 2015. Designing Media for Visually-Impaired Users of Refreshable Touch Displays: Possibilities and Pitfalls. *IEEE Transactions on Haptics* 8, 3 (September 2015), 248–257. DOI:https://doi.org/10.1109/TOH.2015.2466231

[46] Hari P. Palani, Paul D. S. Fink, and Nicholas A. Giudice. 2020. Design Guidelines for Schematizing and Rendering Haptically Perceivable Graphical Elements on Touchscreen Devices. *International Journal of Human–Computer Interaction* 36, 15 (April 2020), 1393–1414. DOI:https://doi.org/10.1080/10447318.2020.1752464

[47] Hari Prasath Palani and Nicholas A. Giudice. 2017. Principles for Designing Large-Format Refreshable Haptic Graphics Using Touchscreen Devices. *ACM Transactions on Accessible Computing* 9, 3 (February 2017), 1–25. DOI:https://doi.org/10.1145/3035537

[48] Hari Prasath Palani, Jennifer L. Tennison, G. Bernard Giudice, and Nicholas A. Giudice. 2018. Touchscreen-Based Haptic Information Access for Assisting Blind and Visually-Impaired Users: Perceptual Parameters and Design Guidelines. *Advances in Usability, User Experience and Assistive Technology* (June 2018), 837–847. DOI:https://doi.org/10.1007/978-3-319-94947-5_82

[49] Hariprasath Palani, Uro Giudice, and Nicholas A. Giudice. 2016. Evaluation of Non-visual Zooming Operations on Touchscreen Devices. In *Universal Access in Human-Computer Interaction. Interaction Techniques and Environments* (Lecture Notes in Computer Science), Springer International Publishing, Cham, 162–174. DOI:https://doi.org/10.1007/978-3-319-40244-4_16

[50] Venkatesh Potluri, Tadashi E Grindeland, Jon E. Froehlich, and Jennifer Mankoff. 2021. Examining Visual Semantic Understanding in Blind and Low-Vision Technology Users. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ACM. DOI:https://doi.org/10.1145/3411764.3445040

[51] Xin Qian, Eunyee Koh, Fan Du, Sungchul Kim, Joel Chan, Ryan A. Rossi, Sana Malik, and Tak Yeon Lee. 2021. Generating Accurate Caption Units for Figure Captioning. In *Proceedings of the Web Conference 2021*, ACM. DOI:https://doi.org/10.1145/3442381.3449923

[52] Manuele Reani, Alan Davies, Niels Peek, and Caroline Jay. 2018. How do people use information presentation to make decisions in Bayesian reasoning tasks? *International Journal of Human-Computer Studies* 111, (March 2018), 62–77. DOI:https://doi.org/10.1016/j.ijhcs.2017.11.004

[53] Manuele Reani, Niels Peek, and Caroline Jay. 2019. How different visualizations affect human reasoning about uncertainty: An analysis of visual behaviour. *Computers in Human Behavior* 92, (March 2019), 55–64. DOI:https://doi.org/10.1016/j.chb.2018.10.033

[54] L. Penny Rosenblum and Tina S. Herzberg. 2015. Braille and Tactile Graphics: Youths with Visual Impairments Share Their Experiences. *Journal of Visual Impairment & Blindness* 109, 3 (May 2015), 173–185.

[55] Wolff-Michael Roth, G. Michael Bowen, and Michelle K. McGinn. 1999. Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching* 36, 9 (November 1999), 977–1019. DOI:https://doi.org/10.1002/(sici)1098-2736(199911)36:9<977::aid-tea3>3.0.co;2-v

[56] Audrey C. Rule, Greg P. Stefanich, Robert M. Boody, and Belinda Peiffer. 2011. Impact of Adaptive Materials on Teachers and their Students with Visual Impairments in Secondary Science and Mathematics Classes. *International Journal of Science Education* 33, 6 (April 2011), 865–887.

[57] M. K. Smith, W. B. Wood, W. K. Adams, C. Wieman, J. K. Knight, N. Guild, and T. T. Su. 2009. Why Peer Discussion Improves Student Performance on In-Class Concept Questions. *Science* 323, 5910 (January 2009), 122–124. DOI:https://doi.org/10.1126/science.1165919

[58] Ian Spence. 1990. Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance* 16, 4 (1990), 683–692. DOI:https://doi.org/10.1037/0096-1523.16.4.683

[59] Ian Spence. 2005. No Humble Pie: The Origins and Usage of a Statistical Chart. *Journal of Educational and Behavioral Statistics* 30, 4 (December 2005), 353–368. DOI:https://doi.org/10.3102/10769986030004353

[60] Ian Spence and Stephan Lewandowsky. 1991. Displaying proportions and percentages. *Applied Cognitive Psychology* 5, 1 (January 1991), 61–77. DOI:https://doi.org/10.1002/acp.2350050106

[61] Sonja Steinbach. 2022. Looking beyond vision: Supports for students who are blind or visually impaired in mathematics. *NWEA* (2022).

[62] Saiganesh Swaminathan, Thijs Roumen, Robert Kovacs, David Stangl, Stefanie Mueller, and Patrick Baudisch. 2016. Linespace: A Sensemaking Platform for the Blind. ACM. DOI:https://doi.org/10.1145/2858036.2858245

[63] Touch Graphics Inc. 2022. Touch Graphics. *Touch Graphics*. Retrieved from https://www.touchgraphics.com

[64] Matthew Turk. 2014. Multimodal interaction: A review. *Pattern Recognition Letters* 36, (January 2014), 189–195. DOI:https://doi.org/10.1016/j.patrec.2013.07.003

[65] US Census Bureau. 2020. 2019 American community survey 1-year estimates: Employment status by disability status and type, Table B18120. Retrieved May 2, 2023 from https://data.census.gov

[66] ViewPlus. 2023. ViewPlus – Delivering Sense Ability. Retrieved from https://viewplus.com/

[67] Deborah L. Voltz, Nettye Brazil, and Alison Ford. 2001. What Matters Most in Inclusive Education: A Practical Guide for Moving Forward. *Intervention in School and Clinic* 37, 1 (September 2001), 23–30. DOI:https://doi.org/10.1177/105345120103700105

[68] Peter Wayner. 2011. Illustrating Your Life In Graphs And Charts. *New York Times*, B9.

[69] WebAIM.org. 2019. Survey on screen reader usage #9. Retrieved May 2, 2023 from https://webaim.org/projects/screenreadersurvey9/

[70] Wai Yu and Stephen Brewster. 2003. Evaluation of multimodal graphs for blind people. *UAIS* 2, 2 (June 2003), 105–124. DOI:https://doi.org/10.1007/s10209-002-0042-6

[71] Leonardo Zani Zamprogno, Bruno Merlin, João Ferreira, Heleno Fülber, and Allan Veras. 2021. Multimodal Tactile Graphics Using T-TATIL, A Mobile Application for Tactile Exploration by Visually Impaired People. In *Universal Access in Human-Computer Interaction. Design Methods and User Experience* (Lecture Notes in Computer Science), Springer International Publishing, Cham, 482–498. DOI:https://doi.org/10.1007/978-3-030-78092-0_33

[72] Kim T. Zebehazy and Adam P. Wilton. 2021. Graphic Reading Performance of Students with Visual Impairments and Its Implication for Instruction and Assessment. *Journal of Visual Impairment & Blindness* 115, 3 (May 2021), 215–227. DOI:https://doi.org/10.1177/0145482x211016918

[73] Braille and Tactile Graphics. Retrieved May 2, 2023 from https://assist.vt.edu/content/assist_vt_edu/en/services/braille.html

APPENDIX

A APPENDIX

Table 1A: BLV participant demographics

Participant ID #	Sex	Age	Blindness and Residual Vision	Avenue of Graph Exposure
1	F	36	Total blindness	Primary, secondary, and post-secondary school; social work
2	F	52	Leber congenital amaurosis, no usable vision	Primary, secondary, and post-secondary school
3	M	62	Light/dark perception, low vision	School and current occupation; investment
4	F	40	Retinitis pigmentosa, no usable vision	Prior to vision loss: school; after vision loss: current occupation
5	M	67	Total blindness	Past occupation in engineering
6	M	48	Light/dark perception, no usable vision	Past and current occupation
7	F	69	Total blindness	Past occupation; tactile graphics educator
8	M	23	Total blindness	Primary and secondary school
9	M	36	Achromatopsia, low vision	School and current occupation
10	F	24	Retinoblastoma, total blindness	Primary and secondary school
11	M	23	Albinism, no usable vision	Primary and secondary school
12	F	43	Retinal detachment, no usable vision	Current occupation
13	F	33	Retinopathy, low vision	Primary, secondary, and post-secondary school
14	M	30	Total blindness	Secondary and post-secondary school
15	M	19	Total blindness	Primary and secondary school
16	M	31	Retinitis pigmentosa, low vision	Primary and secondary school
17	M	28	Light/dark perception	School and past occupation
18	M	31	Ocular hypertension, total blindness	Secondary and post-secondary school
19	M	47	Congenital nystagmus, low vision	School and past occupation

B APPENDIX

Table 1B: Venn diagrams used in the user study with NL and VA descriptions

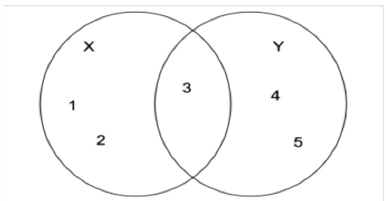
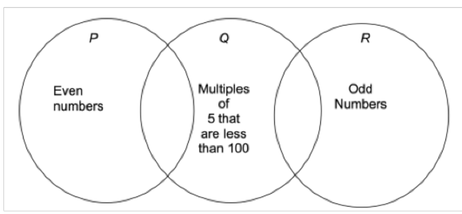
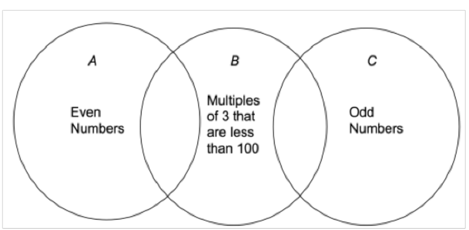
Venn Diagrams in User Study	Natural Language Description	Vibro-Audio Overview
Practice Venn Diagram 	This is a Venn diagram that uses two circles and one intersection to represent two sets of numbers. The circle on the left is labeled X and contains the numbers 1, 2, and 3. The circle on the right is labeled Y and contains the numbers 3, 4, and 5.	This is a Venn diagram that uses two circles and one intersection to represent two sets of numbers.
Experimental Venn Diagram 1 	This is a Venn diagram that uses three circles and two intersections to represent relationships among three sets of numbers. The circle on the left is labeled P and represents even numbers. The circle in the middle is labeled Q and represents multiples of 5 that are less than 100. The circle on the right is labeled R and represents odd numbers.	This is a Venn diagram that uses three circles and two intersections to represent relationships among three sets of numbers.
Experimental Venn Diagram 2 	This is a Venn diagram that uses three circles and two intersections to represent relationships among three sets of numbers. The circle on the left is labeled A and represents even numbers. The circle in the middle is labeled B and represents multiples of 3 that are less than 100. The circle on the right is labeled C and represents odd numbers.	This is a Venn diagram that uses three circles and two intersections to represent relationships among three sets of numbers.

Table 2B: Pie Charts used in the user study with NL and VA descriptions

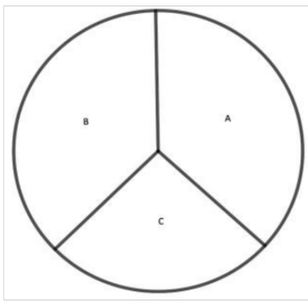
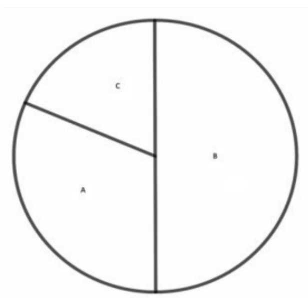
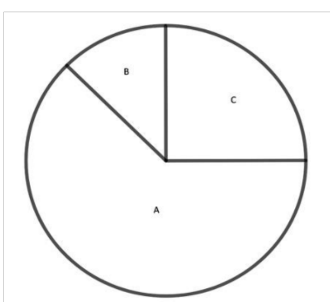
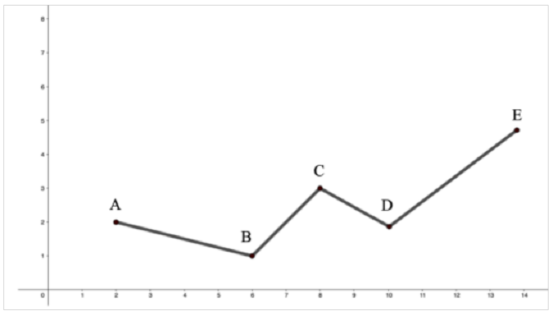
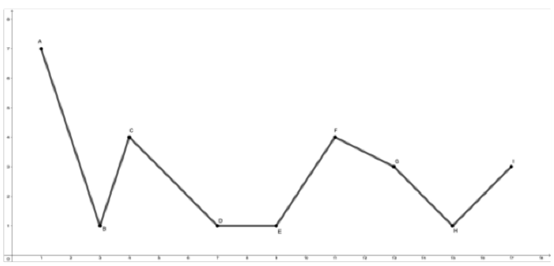
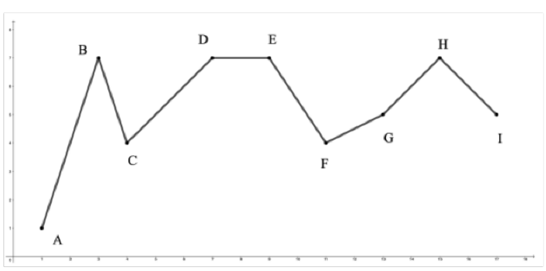
Pie Charts in User Study	Natural Language Description	Vibro-Audio Overview
Practice Pie Chart 	This figure is a pie graph with three regions labeled A, B, and C. Regions A, B, and C are all the same size.	This figure is a pie graph with three regions labeled A, B, and C.
Experimental Pie Chart 1 	This figure is a pie graph with three regions labeled A, B, and C. Region A is twice the size of region C, and region B is the same size as regions A and C together.	This figure is a pie graph with three regions labeled A, B, and C.
Experimental Pie Chart 2 	A pie graph with three regions labeled A, B, and C. Region A is two and a half times the size of region C. Region C is twice the size of region B.	This is a pie graph with three regions labeled A, B, and C.

Table 3B: Line graphs used in the user study with NL and VA descriptions

Line Graphs in User Study	Natural Language Description	Vibro-Audio Overview
Practice Line Graph 	This is a line graph with 5 data points and 4 segments. The data points are labeled from left to right A through E. The vertical axis is ruled in one-unit increments from 0 to 8. The horizontal axis is ruled in one-unit increments from 0 to 14. In the graph, from left to right, the line begins at data point A at coordinate (2,2), then decreases to data point B at coordinate (6,1), sharply increases to data point C at coordinate (8,3), then decreases to data point D at coordinate (10,2) and sharply increases to data point E at coordinate (14,5).	This is a line graph with 5 data points and 4 segments.
Experimental Line Graph 1 	This is a line graph with 9 data points and 8 segments. The data points are labeled from left to right A through I. The vertical axis is ruled in one-unit increments from 0 to 8. The horizontal axis is ruled in one-unit increments from 0 to 18. In the graph, from left to right, the line begins at data point A at coordinate (1,7), steeply decreases to data point B at coordinate (3,1), sharply increases to data point C at coordinate (4,4), then decreases to data point D at coordinate (7,1) and is steady to data point E at coordinate (9,1). The line then steeply increases to data point F at coordinate (11, 4), then gently decreases to data point G at coordinate (13, 3) and more steeply decreases to data point H at coordinate (15, 1). The line finally increases to data point I at coordinate (17,3).	This is a line graph with 9 data points and 8 segments.
Experimental Line Graph 2 	This is a line graph with 9 data points and 8 segments. The data points are labeled from left to right A through I. The vertical axis is ruled in one-unit increments from 0 to 8. The horizontal axis is ruled in one-unit increments from 0 to 18. In the graph, from left to right, the line begins at data point A at coordinate (1,1), steeply increases to data point B at coordinate (3,7), sharply decreases to data point C at coordinate (4,4), then increases to data point D at coordinate (7,7) and is steady to data point E at coordinate (9,7). The line then steeply decreases to data point F at coordinate (11, 4), then gently increases to data point G at coordinate (13, 5) and more steeply increases to data point H at coordinate (15, 7). The line finally decreases to data point I at coordinate (17,5).	This is a line graph with 9 data points and 8 segments.

C APPENDIX

Table 1C: Graph questions answered by participants

Graph Type	Question Type	Question	Correct Answer
Practice Venn Diagram	Criterion -	The number of circles represented in this diagram is. . .	2
	Criterion -	The number 5 belongs to. . .	Y only
Exp. Venn Diagram 1	Criterion -	The number of circles represented in this diagram is. . .	3
	Criterion -	The number of intersections represented in this diagram is. . .	2
	Test -	The intersection of P and Q would contain which of the following groups of numbers?	10, 20, 30
	Test -	Based on this Venn Diagram, the number 85 would belong to. . .	The intersection of Q & R
	Test -	Based on this Venn Diagram, the number 102 would belong to. . .	P only
	Test -	Which of the following numbers is in both P and R?	None
Exp. Venn Diagram 2	Criterion -	The number of circles represented in this diagram is. . .	3
	Criterion -	The number of intersections represented in this diagram is. . .	2
	Test -	The intersection of A and B would contain which of the following groups of numbers?	24, 30, 36
	Test -	Based on this Venn Diagram, the number 81 would belong to. . .	The intersection of B & C
	Test -	Based on this Venn Diagram, the number 213 would belong to. . .	C only
	Test -	Which of the following numbers is in both A and C?	None
Practice Pie Chart	Criterion -	How many regions are in the pie graph?	3
	Criterion -	Region A is. . .	The same size as region C
Exp. Pie Chart 1	Criterion -	How many regions are in the pie graph?	3
	Criterion -	The regions are labeled using. . .	Letters
	Test -	Rank the regions from largest to smallest.	B, A, C
	Test -	What is the area of the largest region as a percent of the whole circle?	50%
	Test -	What is approximately region A as a percent of the whole circle?	33%
	Test -	The smallest region is approximately what percent of the largest region?	33%
Exp. Pie Chart 2	Criterion -	How many regions are in the pie graph?	3
	Criterion -	The regions are labeled using. . .	Letters
	Test -	Order the regions by size from largest to smallest.	A, C, B
	Test -	What is the approximate combined area of the smallest and largest regions as a percent of the total?	75%
	Test -	What is the approximate combined area of regions B and C as a percent of the total?	35%
	Test -	Region C is approximately what percent of Region A?	40%
Practice Line Graph	Criterion -	How many line segments are plotted in the graph?	4
	Criterion -	What is the lowest point of the graph?	B
Exp. Line Graph 1	Criterion -	How many data points are plotted in the graph?	9
	Criterion -	How many line segments are plotted in the graph?	8
	Test -	Which two data points represent the steepest decrease in value?	A to B
	Test -	The distance between data points C and D decreases. . .	More steeply than F to G
	Test -	What is the highest point of the graph?	A
	Test -	The length of the horizontal segment is. . .	2 units
Exp. Line Graph 2	Criterion -	How many data points are plotted in the graph?	9
	Criterion -	How many line segments are plotted in the graph?	8
	Test -	Which two data points represent the steepest increase in value?	A to B
	Test -	The distance between data points C and D increases. . .	More steeply than F to G
	Test -	What is the lowest point of the graph?	A
	Test -	The length of the horizontal segment is. . .	2 units