Spatial Audio-Enhanced Multimodal Graph Rendering for Efficient Data Trend Learning on Touchscreen Devices

Wilfredo J. Robinson M.
Department of Aerospace and Mechanical Engineering, Saint Louis University, United States
wilfredo.robinsonmoore@slu.edu

Medhani Kalal
Department of Aerospace and Mechanical Engineering, Saint Louis University, United States
medhani.kalal@slu.edu

Jennifer L. Tennison
Department of Aerospace and Mechanical Engineering, Saint Louis University, United States
jen.tennison@slu.edu

Nicholas A. Giudice
VEMI Lab, University of Maine, United States
nicholas.giudice@maine.edu

Jenna Gorlewicz
Department of Aerospace and Mechanical Engineering, Saint Louis University, United States
jenna.gorlewicz@slu.edu

ABSTRACT

Touchscreen-based rendering of graphics using vibrations, sonification, and text-to-speech is a promising approach for nonvisual access to graphical information, but extracting trends from complex data representations nonvisually is challenging. This work presents the design of a multimodal feedback scheme with integrated spatial audio for the exploration of histograms and scatter plots on touchscreens. We detail the hardware employed and the algorithms used to control vibrations and sonification adjustments through the change of pitch and directional stereo output. We conducted formative testing with 5 blind or visually impaired participants, and results illustrate that spatial audio has the potential to increase the identification of trends in the data, at the expense of a skewed mental representation of the graph. This design work and pilot study are critical to the iterative, human-centered approach of rendering multimodal graphics on touchscreens and contribute a new scheme for efficiently capturing data trends in complex data representations.

CCS CONCEPTS

• **Human-centered computing** → Accessibility; Ubiquitous and mobile computing; Ubiquitous and mobile devices; Mobile devices; • **Social and professional topics** → User characteristics; People with disabilities; • **Hardware** → Communication hardware, interfaces and storage; Tactile and hand-based interfaces; Touch screens.

ACM Reference Format:
Wilfredo J. Robinson M., Medhani Kalal, Jennifer L. Tennison, Nicholas A. Giudice, and Jenna Gorlewicz. 2024. Spatial Audio-Enhanced Multimodal Graph Rendering for Efficient Data Trend Learning on Touchscreen Devices. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3613904.3641959

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3641959

1 INTRODUCTION AND BACKGROUND

The digitalization of information has significantly increased the availability of educational data visualizations (such as figures, charts, and videos) [1–3]. However, these graphics are largely inaccessible to people with blindness or visual impairments (BVI) due to high visual requirements. While text can be read aloud by Text-to-Speech (TTS) engines and screen readers, graphics must have an associated description, either via alternative text (alt-text) or image recognition tools. Yet, it is challenging to convey the rich information of a data visualization through words alone [4]. To address this issue, multiple approaches have been utilized such as using tactile and 3D printouts, mixed-media and manipulatives [5], graphics with embedded QR codes [6–8], image recognition software [9–11], dedicated touchscreen tablets [12, 13], and refreshable braille displays [14–16]. While the availability and versatility of these options somewhat addresses the challenge of digital information access, they come at a high cost and/or require specific training, limiting mass adoption.

Recent research has focused on utilizing the capabilities of more affordable touchscreen-based tablets and phones to bridge this digital accessibility gap. The ability to leverage the touchscreen's inherent multimodal feedback features (vibrations, sounds, and visual display) to add attributes to graphics has been demonstrated to be beneficial for the interpretability of lines, maps, graphs [17–20], and shapes [20–22]. Guidelines to render these graphics for touchscreens have also been proposed [17, 23]. However, these approaches face a new challenge: placing a high cognitive load on the user [17]. Prior work has found that a multimodal feedback scheme should not use more than three to five feedback cues such as vibrations [24, 25], otherwise the cognitive load placed upon the user could significantly negatively affect their task performance [23]. These constraints challenge the rendering of more complex graphs such as histograms, scatter plots, or increasingly intricate N-dimensional graphs.

A potential solution emerges with the introduction of spatial audio. Spatial audio refers to "the rendering of audio content with spatial attributes in a manner that enables listeners to perceive the sound sources as if they were originating from distinct positions and distances around them, creating a sense of immersion and spatial realism" [26]. Spatial audio has been successfully applied in other fields such as spatial sound scenes [27], 360-degree videos [28], audiovisual learning tools [29], exploration of menus and progress bars on mobile devices [30, 31], and real-world navigation [32, 33]. Spatial audio has also been implemented in the navigation of complex virtual reality (VR) environments, with results demonstrating that spatial learning in the VR environment led to equally accurate spatial learning and navigation for BVI individuals in the real world as from Orientation and Mobility instructors [34, 35]

Spatial audio, like the senses of vision and touch, allows the user to have direct perceptual knowledge of spatial information and spatial relations, which is also important to understanding graphics and data representations. This contrasts with Natural Language (NL) descriptions, which do not convey this information using direct spatial perception but use cognitively mediated linguistic processes, which are slower, require more cognitive effort to process, and are subject to greater error than a perceptual interface [36].

In the context of data visualization and graphics, researchers have used non-speech audio to aid BVI users when exploring two-dimensional tabular data [37], bar charts [38], data in VR environments [39], lines, and line charts [40, 41]. The sonification in [41] mapped discrete y-axis positions to the pitch of the sound produced (higher y value, higher pitch). Similarly, the authors in [42] assigned participants the task of exploring lines and line graphs on a touchscreen, utilizing three types of feedback: vibratory only, auditory only, and bimodal. This study also correlated the y-axis position of the finger to the pitch of the auditory feedback generated, yet discovered no statistically significant difference in performance across the three modalities. None of these studies associated the x-axis position with output to stereo speakers, thereby neglecting the potential spatial information that could have been conveyed through the directionality of the audio output. Finally, a recent study had BVI participants explore Venn diagrams, line charts, and pie charts using two modalities, natural language only vs. vibro-audio exploration, and showed that both modalities were equally viable for learning the graphical information displayed [43]. However, the audio modality in this study did not implement changes in pitch or spatialized speaker output. To our knowledge, spatial audio controlling both pitch and stereo speaker output has yet to have application in the multimodal rendering of touchscreen graphics.

Prior research posits that the ability to extract a high-level summary from a graph, while simultaneously scrutinizing minor details, is essential for users. This approach allows a continuous cycle of questioning and comparison, enabling the user to derive their own insights during the exploration of the graph [44]. However, this issue continues to pose a significant challenge for sighted individuals, BVI individuals, and even the most recent machine learning algorithms [4, 45]. Despite being a challenging aspect of graph comprehension, this task is fundamentally associated with gaining an understanding of the data trends represented in graphs. Discerning these trends is highlighted by a recent review of research literature as a key skill for drawing accurate inferences and insights. This study also showed that comparing trends across multiple graphs can yield deeper insights, further emphasizing the critical role of trend identification in data visualization [46].

Another review, which analyzed 292 papers on accessible graphics from 2010 to 2020, found that only 23 were comparative studies involving BVI participants, and only three focused on graphs and data visualization. Of these, only one study examined whether participants could understand trends in a graph, specifically bar charts [47]. This underscores the urgent need for more research aimed at devising methods that enable BVI individuals to effectively explore graphs, comprehend data trends, and derive meaningful insights.

In this work, we investigate the inclusion of spatial audio in multimodal touchscreen graphics, specifically histograms and scatter plots. These chart types were selected because their characteristics demand a detailed understanding of the overarching data trend. We hypothesize that the implementation of spatial audio in the rendering of these graphs will help the user obtain relevant information related to the distribution of the data on screen and thus support more accurate learning, interpretation, and representation of this information, when compared to the same process using standard audio. To investigate this prediction, we designed a new multimodal feedback scheme for histograms and scatter plots on touchscreen tablets and performed a formative study with five BVI individuals, assessing their accuracy in re-creating graphs from memory and their chart exploration preferences. The application and source code used in this study are publicly available on GitHub [48], enabling further research on multimodal feedback schemes with spatial audio. This design work and pilot study play a crucial role in the iterative, human-centered process of rendering multimodal graphics on touchscreens.

2 MULTIMODAL FEEDBACK: STANDARD VS SPATIAL AUDIO SCHEMES

Multimodal rendering of graphs on touchscreens typically uses four main feedback mechanisms: sonification, vibration, text description, and visual cues. Sonification specifically refers to the use of non-speech audio to convey information or perceptualize data [49] and is output by the device's speakers. Vibration utilizes the device's built-in motor (usually an eccentric rotating mass motor or linear actuator) to alert the user of an event and to represent lines or objects that might require tracing. TTS is provided by the device's operating system and is used to read aloud all text, add context, and provide more detailed information about important sections of the graphic [17, 23]. Finally, the graphic itself is visually displayed (optionally with magnification or high contrast) on the touchscreen. The integration of these four feedback mechanisms into the rendering of a graph is called a multimodal feedback scheme.

A multimodal feedback scheme on a touchscreen allows the user to explore the graph with one or multiple fingers, and the graph provides different feedback cues depending on the finger's location and the type of information being touched at that location. Within a feedback scheme, sonification and vibration are used to provide spatial information to the user, whereas text descriptions voiced with TTS engines are used to relay semantic information of the graph, such as the meaning of a bar in a bar chart [23].

A limitation presents itself when the user requires more precise information about their finger's location on the touchscreen. In a

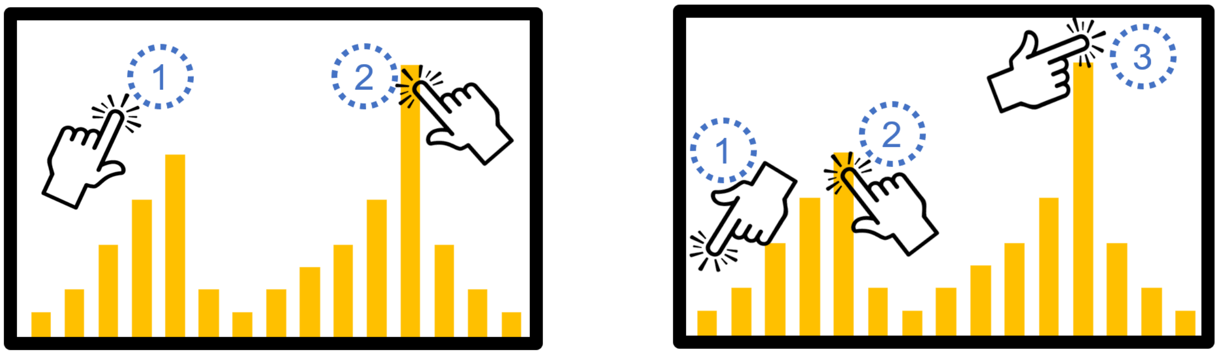

Figure 1: Standard Feedback Scheme (left): 1: touch on empty space produces standard sound; 2: touch on bar (or data point) stops sound and produces vibration. Spatial Audio Feedback Scheme (right): 1: touch on bottom left empty space produces directional sound (left speaker only) with low pitch; 2: touch on bar (or data point) stops sound and produces vibration; 3: touch on top right empty space produces directional sound (right speaker only) with high pitch.

feedback scheme, a sound is used to denote meaningful empty space, such as the interior of a shape or a notable section of the graph. This sound is applied universally across the screen whenever the user's finger encounters such spaces. To ascertain their exact position on the screen – top left, middle, or bottom right – users would need to devise their own mental mapping and localization strategy. Alternatively, the developer of the feedback scheme would need to incorporate additional vibrations or audio cues for each area of the screen. Previous research has acknowledged the ineffectiveness of adding too many signals of the same modality [23].

We hypothesize that spatial audio can provide more detailed localization information without increasing the number of cues, thereby avoiding further cognitive load issues. This is achieved through the use of directional audio (control of stereo speaker channels) and pitch control. For instance, when the user's finger moves to the right edge of the screen, only the right speaker will produce sound and vice versa when moving to the left. The pitch of the sound will vary depending on the vertical position of the finger, from high pitch at the top of the screen and low pitch at the bottom. In essence, we aim to convey spatial information more efficiently by manipulating pitch and stereo output channels while reducing the overall state space (and cognitive load) of multimodal cues.

For the scenarios presented in this paper, both standard and spatial audio schemes were developed and compared for histograms and scatter plots. The hardware modifications, the size of the graphical elements on screen, the vibration profiles, and the number of vibro-audio cues were selected following the guidelines in [17, 23, 24]. In both schemes, vibrations were used to represent the data (bins in histograms and data points in scatter plots). The main difference was in the approach to sonification: the standard approach used the same sound to denote empty space, regardless of the finger's location on the screen, while the newly developed approach modified the sound's direction and pitch depending on the finger's location on the screen (see Figure 1 and Video Figure 1).

3 HARDWARE AND SOFTWARE IMPLEMENTATION

Multimodal histograms and scatter plots were developed for use on an Android platform. This section details the small modifications made to the tablet hardware that align with previous research to support user exploration [23], and the implementation of the new multimodal feedback scheme with spatial audio in Android.

The Android application for this research study was deployed on a Samsung Galaxy Tab S5e, with a screen size of 10.5" and a resolution of 2560 x 1600 pixels. Following prior research [23], minor hardware modifications were made to the tablet: (1) tape was used as a bounding contour, allowing the user to identify the active area of the screen, (2) anti-slip bumper dots on the lower corners on the back of the tablet, and (3) two pop sockets on the upper corners of the back of the tablet to provide a more comfortable exploration angle for the users.

The software application was developed entirely in the open-source Android Studio Electric Eel 2022.1.1 Patch 1. All of the imported libraries and objects mentioned in this section are open-source and either built-in to Android or completely programmed in Java.

The wave sounds representing empty space were obtained from Freesound.org [50] and trimmed with Audacity [51]. A wave sound was selected to represent empty space because of its continuous, unintrusive sound profile, which previous research has illustrated is important for a high-quality user experience, particularly if a majority of the screen is empty space [52, 53].

All spatial audio sonification was controlled by Android's built-in SoundPool library, which allows for near-analog control of the changes in pitch and audio channel output (left or right). Regarding pitch control, the wave sound used in this study was not a monotone pitch, which means it must be characterized as a frequency spectrum. The most common frequency in this spectrum is 293 Hz (D4 tone) and thus will be used as the reference point to explain pitch modulation implementation. The default tone (293 Hz, D4) plays when the user is at the top of the screen. The pitch continuously adjusts every instant the user moves vertically 1 pixel on the screen (Y axis), reaching 146.5 Hz (D3 tone) in the middle and 23

Hz (F#0 tone) at the bottom. Reducing the tone any lower than this would remove it from the human hearing range (approximately 20 Hz – 20 kHz), and also make it nonreproducible for the device speakers. The spatial audio scheme produces 44 semitone changes (22 tones, 3.7 octaves) when the user moves between the top and bottom of the screen.

Regarding audio channel output control (left or right), both speakers output audio at 100% of the volume when the user's finger is in the middle of the screen, and this distribution adjusts every pixel the user moves horizontally (X axis). Once the user reaches the right side of the screen, the right speaker outputs audio at 100% volume and the left speaker outputs audio at 0%. The inverse behavior happens if the user reaches the left of the screen. To this end, each participant was provided with a pair of circumaural stereo headphones (KVIDIO) to receive audio output during the study. It is important to note that the versatility of this design with the built-in Android SoundPool library provides the developer a substantial degree of granularity to modify both the pitch and audio channel output behaviors to adjust to their target user demographic.

Multi-finger detection and tracking were implemented by importing and modifying the SimpleFingerGestures library [54], which allows for tracking of up to 5 fingers on screen. This library also allows for the creation of custom multi-finger gestures to be implemented if needed, but the implementation of these was omitted to avoid a considerable expansion of the current study's scope.

All vibration cues were controlled by using a slightly modified version of the Vibration Libraries from the Android Plugins for the accessibility-focused Quorum Programming Language [55]. The vibration profile used was triggered by the "vibrateForever()" method, which initiates a max amplitude, constant vibration at the natural frequency of the built-in motor (200 Hz) when the user touches a bin in a histogram or a data point in a scatter plot. The vibration would stop when the user left the data-relevant portion of the graph.

Additionally, a proportional scaling algorithm was implemented such that this multimodal scheme would work on screens of varying sizes. This feature allows the graphics and the data-relevant coordinates responsible for triggering vibration feedback to adjust to the screen dimensions of any Android phone or tablet.

Finally, the finger's location was logged every 0.3 seconds using the built-in File, FileWriter, and SimpleDateFormat objects. This logging period was selected because it allowed the tablet to write to disk multiple times per second without affecting the application's framerate, as any slowdown in framerate could negatively affect the user experience [56]. A pseudo-code for the algorithm used in the app is shown in Algorithm 1.

Algorithm 1 General Multimodal Feedback Scheme with and without Spatial Audio

```
# Detect the presence, location, and color of the pixel under the finger
finger_location, pixel_color =
detect_finger_location_and_pixel_color()
# Check if the finger is on "empty space"
if pixel_color == "empty space":
    # Stop all vibrations and play the "wave" sounds
    stop_vibrations()
    play_wave_sounds()
    # Check if spatial audio is activated for this exploration task
    if spatial_audio_activated():
        # Modify the proportion of the audio sent to each audio output channel (left or right)
        # based on finger's X axis position.
        modify_audio_direction(finger_location.x)
        # Modify the pitch of the audio based on the finger's Y axis position.
        modify_audio_pitch(finger_location.y)
# Check if the finger is on a bin (histograms) OR a data point (scatter plots)
elif pixel_color in ["bin", "data point"]:
    # Stop all audio output and start vibrating.
    stop_audio_output()
    start_vibrating()
# In all feedback schemes, log the X and Y locations of all fingers on the screen every 0.3 seconds.
log_finger_locations_every_0_3_seconds()
```

4 RESEARCH STUDY DESIGN

This was an iterative and formative evaluation of our design schemes involving users with lived experience. As such, the purpose of this work was to serve as a pilot study into multimodal feedback schemes with integrated spatial audio for histograms and scatter plots to help guide further design and development of optimized schemes for inclusive data presentation. It aimed to answer the following questions: (1) how does the presence of spatial audio during graph exploration affect the extraction of data-related features (inflection points in the data trend, number of modes in histograms, curve characteristics in scatter plots) as evaluated through participants' re-creation of the graph?, (2) how does the presence of spatial audio during graph exploration affect the physical features (orientation and scale) represented in the participants' subsequent graph re-creation?, and (3) would users prefer to use spatial audio in the exploration of multimodal histograms and scatter plots on a touchscreen? The selection of a re-creation task as the primary assessment method was deliberate, as it required the participants to not only grasp the essential data-related insights from the graph but also reconstruct them from memory in a physical format. This approach necessitates a significantly more cognitively complex process than simply reading the graphs and provides us with deeper insights into the nuanced impacts of spatial audio on the underlying representation built up from learning. This re-creation task is a novel approach to gaining insights into a participant's understanding of the data, as previous studies using similar accessible or multimodal approaches have typically only measured learning performance through comprehension and data extraction questions [17, 21, 23, 37]. This study was approved by the presiding university's Institutional Review Board (IRB).

4.1 Participants

Five individuals with BVI were recruited to participate as volunteers from a local orientation and mobility summer camp for young adults with BVI (see Table 1). Young adults were chosen as the primary

Table 1: Participant Demographics for our Research Study

#	Relevant Courses	Braille Reading	Access Aids	Age	Grade Level	Self-reported Vision Status
1	Pre-algebra, Algebra 1&2, geometry	Yes	Braille Display	18	Graduated	Cortical Blindness (PNET)
2	Pre-algebra, Algebra 1&2, geometry	Yes	JAWS, Braille Display, BrailleNote Touch	16	11	Persistent Hyperplastic Primary Vitreous
3	Algebra 2, geometry, trigonometry	Yes	Braille Display, Voice over, JAWS, NVDA, Narrator, TalkBack, ChromeVox	18	12	Retinopathy of Prematurity (ROP)
4	Pre-algebra, Algebra 1&2, geometry, statistics	No	None	22	Graduated	Legally Blind
5	Pre-algebra, Algebra 1&2, geometry	Yes	JAWS, phone, iPad, focus 40	17	12	Retinopathy of Prematurity (ROP)

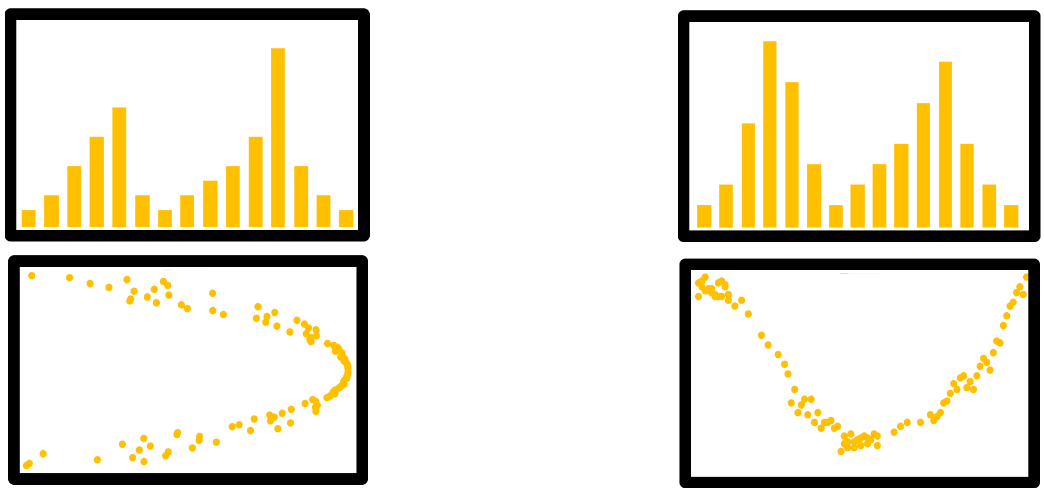

Figure 2: Histograms and Scatter Plots presented to the participants during the research study. Top left: a histogram with 15 bins and a bimodal distribution, explored with standard audio. Top right: a histogram with 15 bins and a bimodal distribution, explored with spatial audio. Bottom left: a scatter plot with 100 points and 1 inflection point, explored with standard audio. Bottom right: a scatter plot with 100 points and 1 inflection point, explored with spatial audio.

demographic due to the alignment of their education level with the content presented in this study. It is important to note that while not all participants had taken statistical courses, their foundational knowledge of algebra and geometry facilitated the explanations of both histograms and scatter plots during the training phase.

4.2 Graphs

The histograms consisted of a bimodal distribution with 15 bins, and the scatter plots consisted of 100 points following a trend with a single inflection point (see Figure 2). The characteristics for these graphs were selected to strike a balance between two factors: (1) spatial complexity of the graph, and (2) limitations on the size of objects on the screen. In the context of this study, the term "spatial complexity" is an abstraction of how much information the user must obtain from the graph to understand the trend of the data. For example: a histogram with 3 modes is more spatially complex than a histogram with a normal distribution (a single mode), and a line graph with multiple inflection points is more spatially complex than a straight line.

In previous studies exploring the multimodal rendering of graphs on touchscreens, the graphs used were bar charts with 4 or less bars [21, 23] and lines with 3-4 inflection points [19]. For this study, we chose histograms with a higher spatial complexity. These graphs needed to meet the guidelines and limitations related to bin/bar size established in [17, 23]. Thus, a bimodal histogram (multiple inflection points) with 15 bins (more than three times those of previous studies) covers the complexity requirements and object-size limitations on our available screen real-estate.

Regarding scatter plots, previous research states that bars, lines, and angles should be at least 2 mm wide for them to be detectable by the user and 4 mm for them to be traceable [17, 23]. As such, the data points on the tablet were 5 mm in diameter, to account for previous guidelines and include a 1 mm buffer because a single data point is much smaller than the elements referenced in the guidelines. Prior research has rendered scatter plots with 20 data points [57] and 12 data points [58] on tactile graphics, but not on touchscreens. Thus, a 100-point scatter plot with a single inflection point was chosen as it provided a balance between spatial complexity and rendering constraints to support our initial investigation.

While we acknowledge that histograms and scatter plots can be significantly more complicated, our approach is to incrementally increase complexity while optimizing the user interface and search strategies through a human-centered design process. This work represents the next iteration of complexity for multimodal graph rendering.

4.3 Procedure

After being provided informed consent, participants were asked to complete a demographics questionnaire (see Table 1) after which the study commenced. Participants were first presented with a training phase where the experimenters explained the definition and purpose of histograms and scatter plots. This allowed participants to freely explore both graphs on the touchscreen and as an embossed printout to establish familiarity with the context and to ask any questions before proceeding.

After exploring the graphic on the touchscreen and embossed printout, participants were asked to re-create the graphs from memory by manipulating Wikki Stix on a plastic canvas the same size as the touchscreen. Progression past the training session was criterion-based, such that participants had to correctly answer a total of five comprehension questions related to the trend of the data in the graphs. Example questions related to histograms were: (1) "How many bins could you find?" and (2) "Which bin was the tallest/shortest?". Example questions related to scatter plots were: (1) "What direction do you feel the points are generally going?" and (2) "Where in the graph do you think the points are the most clustered together?". Though some participants needed more time than others, all participants successfully completed the training session.

In the experimental session, participants were given two tasks: (1) freely explore a histogram or scatter plot and try to understand the trend of the data within a 2-minute time limit, and (2) use WikkiStix to re-create the trend of the graph on the plastic canvas within a 2-minute time limit. The 2-minute time limit ensured participants would be able to finish all trials within 60-minutes, which was found to be sufficient from pilot testing. On average, the 60-minutes were distributed as follows: 28-30 minutes for training, 16 minutes for graph explorations and re-creations, 5-7 minutes for post-study preference-related questions, and 8-10 minutes for introductions, additional participant questions, and breaks (if requested). Each participant was tasked to complete two trials for each graph type wherein one trial used a standard multimodal feedback scheme (sonification, vibration) and the other used a modified multimodal feedback scheme (spatial audio, vibration). This modified feedback scheme retained all the attributes of the standard feedback scheme, with the singular distinction being the spatialization of the sonification. The order in which the graphs and multimodal feedback schemes were presented was alternated for each participant to mitigate potential learning effects. When changing sonification modalities, the trend of the data would change but not the graph complexity (see Figure 2).

For each participant, the data collection setup consisted of: (1) tablet software to log the temporal location of all the participants' fingers on screen during each of the exploration tasks, (2) a fixed tripod to record participants' re-creations, and (3) a second fixed tripod to video record the movement of the participants' hands during the entirety of the study.

After study completion, participants were asked for their preferences related to the implementation of spatial audio in their exploration assignments and which type of graph they deemed easiest to explore. Additionally, two experimenters independently rated the performance of the participants' re-creations by comparing them to the original graph on the touchscreen.

5 RESULTS AND DISCUSSION

This section details our findings regarding: (1) users' performance when re-creating the graphs from memory, (2) users' preferences related to the use of spatial audio for graph exploration, and (3) the main exploration strategies employed by users when exploring histograms and scatter plots. Prior to presenting the results, we establish interrater reliability for the experimenters that evaluated user performance. Our key takeaways can be summarized as follows: (1) Spatial audio has the potential to increase the identification of trend-related insights of information but may also skew the mental representation of the overall graph. Future iterations should include a calibration phase to allow users to learn the full range of audio pitch and direction before exploring, which may mitigate this bias. (2) Participants prefer exploring histograms and scatter plots with spatial audio, potentially finding it more useful when accessing less structured graphs like scatter plots. (3) A larger sample size is needed to further investigate more user exploration strategies and compare them with the pilot findings observed in this study. Future studies should also investigate the use of anchor points and multi-finger exploration strategies. The following subsections will provide more detail on these key takeaways.

5.1 Interrater Reliability Analysis

Before analyzing results, we established interrater reliability for two experimenters who independently rated the participants' re-creation performance based on four criteria: (1) identification of inflection points in the trend; (2) identification of all histogram modes and scatter plot curve features; (3) orientation of the graph re-creation; and (4) scale of the graph re-creation. Criteria (1) and (2) are data-related insights, while criteria (3) and (4) are related to the physical representation of the participants' mental image of the graph. These four criteria were evaluated with 1-4 points, where 1 point indicated the worst performance and 4 points indicated near-perfect performance (see Table 2).

Each graph type had the four criteria from the rubrics rated independently by each experimenter. As this data was ordinal, not nominal, a Weighted Kappa Analysis with linear weights was used

Table 2: Graph Re-creation Rubric

Rubric Component	4 points	3 points	2 points	1 point
Inflection Points	Participant identified the correct number of inflection points	Participant identified +- 1 inflection points	Participant identified +- 2 inflection points	Participant identified +- 3 (or higher) inflection points
Trend	Histograms: - Participant identified all modes - Participant showed their location and relative height difference with a max +- 15% localization error Scatterplots: - Participant identified both sides of the curve and connected them through the inflection point	Histograms: - Participant incorrectly identified max +- 1 modes - Participant showed their location and relative height difference with a max +- 30% localization error Scatter plots: - Participant identified at least one side of the curve and tried to connect it to the inflection point	Histograms: - Participant incorrectly identified max +- 2 modes OR - Participant showed their location and relative height difference with a max +- 45% localization error Scatterplots: - Participant identified one side of the curve but did not try to connect it to the inflection point	Histograms: - Participant incorrectly identified max +- 3 modes OR - Participant showed their location and relative height difference with a localization error greater than 45% Scatterplots: - Participant either did not identify any sides of the curve or incorrectly identified the inflection point
Orientation	10 degrees > re-creation tilt >= 0 degrees	30 degrees > re-creation tilt >= 10 degrees	60 degrees > re-creation tilt >= 30 degrees	re-creation tilt >= 60 degrees
Scale	100% >= re-creation size > 90%	90% >= re-creation size > 75%	75% >= re-creation size > 50%	re-creation size <= 50%

instead of the standard unweighted Cohen's Kappa to establish interrater reliability. The resulting Kappa scores ranged from 0.6 to 1.00, with 14/16 scores higher than 0.7, which illustrate reliability among the evaluators while also highlighting the sensitivity of the results given the variables. The singular Kappa score of 0.6 was a result of three, single-point differences in rating between experimenters. Usually, this sensitivity is countered by a large sample size, which is a limitation of this study. The significance (p) of these tests were distributed as follows: < 0.05 on 9/16 tests, $0.06 - 0.08$ on 6/16 tests, and a single 0.1. The significance (p) of these interrater reliability tests were obtained only to validate these interrater ratings and do not have any bearing on the participants' re-creation performance results.

5.2 Participant Performance Re-creating Graph Trends

Participants were tasked with re-creating the trend of each histogram and scatter plot they explored (see Figure 3). Due to our low sample size and the intent of our study being formative for iterative design, we limited our analysis to basic descriptive statistics to support our qualitative analysis. Percentages are used solely for the purpose of discussing performances across two groups and no inferential statistical analyses were performed.

The re-creation rubric had two variables intimately associated with the trend of the data in the graph on screen: Inflection and Trend. Participants' performance related to these variables can be summarized as follows: (1) Inflection: participants performed 3.25% better (*mean* = 0.13/4.00 points, *SD* = 0.63) when identifying the inflection points of the trends in histograms with spatial audio, but there was no numerical difference in performance for this feature when re-creating scatter plots. (2) Trend: participants performed 9.50% better (*mean* = 0.38/4.00 points, *SD* = 0.75) when identifying the correct number of modes in histograms with spatial audio and performed 6.25% better (*mean* = 0.25/4.00 points, *SD* = 1.06) when identifying both sides of the curve and connecting them through the inflection point on scatter plots with spatial audio. This indicates that spatial audio has great potential to increase the identification of data-trend-related insights when exploring histograms and scatter plots.

However, results vary when analyzing the variables more closely related to the physical re-creation of the graph from memory: Orientation and Scale. Participants' performance related to these variables can be summarized as follows: (1) Orientation: participants performed 3.25% worse (*mean* = -0.13/4.00 points, *SD* = 0.25) when orienting their re-creation of histograms with spatial audio but performed 6.25% better (0.25/4.00 points, *SD* = 0.35) when orienting their re-creations of scatter plots with spatial audio. (2) Scale: participants performed 15.75% worse (*mean* = -0.63/4.00 points, *SD* = 1.11) when scaling their re-creations of histograms with spatial audio and 12.5% worse (*mean* = -0.50/4.00 points, *SD* = 0.71) when scaling their re-creations of scatter plots with spatial audio. This indicates that even though spatial audio increased the participants'

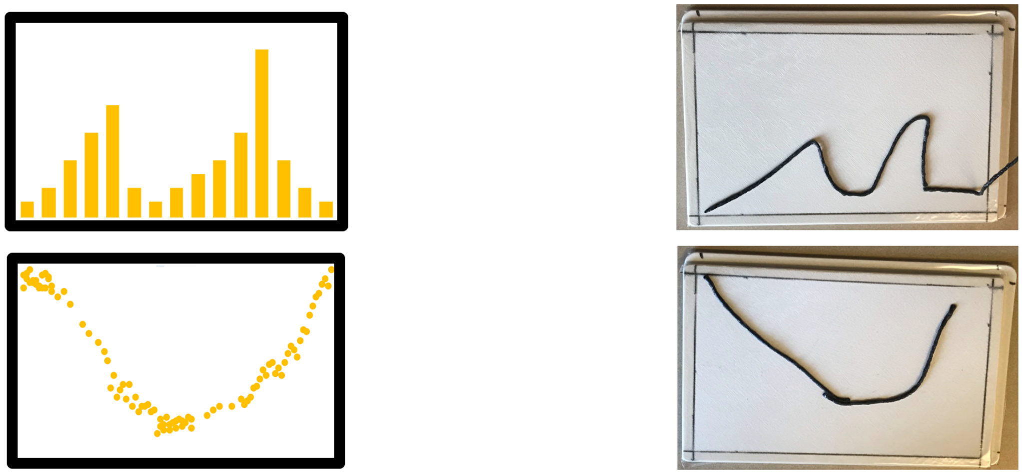

Figure 3: Samples of participants' re-creations for both histograms and scatter plots. Top left: a histogram explored with standard audio. Top right: Participant 1's re-creation of the trend in the histogram. Bottom left: a scatter plot explored with spatial audio. Bottom right: Participant 4's re-creation of the trend in the scatter plot.

data-trend-related takeaways, it may have also altered the participant's mental image of the graph displayed on screen, especially with respect to scale.

We believe that the constant changing of pitch and direction in the received audio could cause participants to stretch their mental image of the graph. For example, if a participant is exploring a graph and the pitch has become very low and most of the audio is coming from their right earcup, it may be difficult for them to determine how far down or right on the tablet screen they are without an initial calibration phase to learn the full range of audio pitch and direction. Without having a baseline modulus to provide this frame of reference, participants may perceptually stretch or bend the graph in certain ways, which can bias or otherwise affect their mental image and physical re-creation of it. Future research must explore whether this alteration is a universal consequence of all spatial audio implementation and whether it bears positive or negative implications on the exploration and comprehension of information within accessible graphs.

5.3 Participant Preferences Regarding Spatial Audio for Graph Exploration

Participants were asked if they preferred exploring the graphs with or without spatial audio. Four out of five participants preferred exploring histograms with spatial audio, and all participants preferred exploring the scatter plots with spatial audio. Participants emphasized that the use of pitch and directional audio gave them a better understanding of their finger's location on the tablet. Participant 1 mentioned that spatial audio "is more pleasant to listen to. It does help knowing if you are closer to the top or closer to the bottom. Especially if you are looking for something specific." They also gave an example related to their education: "if a teacher gave you information, like a number, then [based on its value] I can find it near the bottom, or near the top. I think functionally, the different pitches are helpful!" Participants 3, 4, and 5 shared a similar sentiment to the utility of spatial audio. On the other hand, Participant 2 mentioned that they preferred spatial audio for scatter plots only, because "it's easier for me to determine the pattern of the histogram without spatial audio". We speculate that the participant felt the structured nature of the histogram already provided sufficient information about the graph features and data trend. It is possible that participants would consider spatial audio more useful in less structured graphs, such as scatter plots, where data points can be located anywhere on the screen, rather than being aligned around the X-axis.

Second, participants were asked which of the graphs were easier to explore, and all participants stated that histograms were much simpler to explore than scatter plots, regardless of the sonification modality. Participant 2 summarized it as "Scatter plots [were more difficult] because I just could not really tell where the clusters and dots were". Participant 1 shared a similar sentiment: "The frustration is looking for or identifying the dot. It takes a while to pinpoint where it is". These takeaways provide critical insight on the unique attributes that spatial audio provides but also on the challenges that exist in multimodal renderings of charts, particularly unstructured formats.

5.4 Main Exploration Strategies for Histograms and Scatter Plots

To uncover exploration strategies of participants, finger coordinates were logged during graph exploration on the tablet. After analyzing the exploration patterns of all participants, derived from the logged temporal finger coordinates and video footage, there are some distinct strategies that were employed for both histograms and scatter plots. This section focuses on explaining and showcasing these strategies with the obtained data.

All participants initially approached exploring graphs by discretely tapping on the screen as opposed to slow, continuous motions that are more conducive for exploration. During training, it

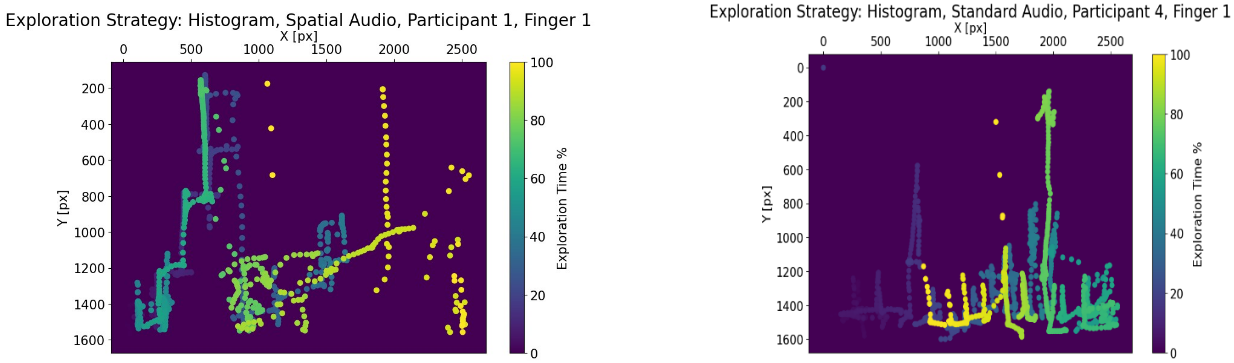

Figure 4: Main strategies when exploring histograms. Points are rendered in chronological order, with initial exploration points rendered with dark blue and changing to light yellow the longer a participant explores. Participant 1 (left) traced the bins from left to right, and then restarted the same left-to-right process. On the other hand, Participant 4 (right) traced the graph from left to right, and back tracked the tracing from right to left.

was suggested to users to use swiping, slow motions to explore the graph on screen. It was noted that some participants took longer to adapt to this type of interaction with the screen, but others, such as Participant 4, attempted to employ multi-finger approaches from the beginning. Participants also employed interesting strategies of quickly moving one finger to the edge of the screen and then back to a previous location to reorient themselves in identifying elements of the graph. For histograms, a common strategy employed when exploring across bins was to return to the previous bin to re-check and compare heights as they explored the new bin. These strategies articulate organic, successful approaches that were developed on-the-fly by participants and provide interesting initial insights into strategies that may be helpful in exploring multimodal graphics.

Three areas that were challenging for participants that were revealed through exploration strategies include: 1) following straight line paths (noted in histograms), 2) finding individual points, and 3) distinguishing single points from clusters of dense points (points 2 and 3 both noted in scatter plots). Interestingly, it was also observed that finding clusters of dense points was challenging in the training session as well, even in embossed versions, indicating that this is a challenge that stretches across mediums and warrants further research. We believe the difficulty in point 2 can be attributed to two main factors: the tablet’s computation time and the vibration motor’s response time. As stated in [17], a Galaxy touchscreen can have a computation delay as high as 10 ms and a mechanical actuation delay as high as 10 ms which, when combined, can become a perceivable lag for users. Considering the data points had a diameter of 5 mm and the speed at which participant “swiped”, the participants could quickly move their finger over a point and then leave it before the vibration motor had a chance to reach a discernible vibration amplitude.

One potential solution is to assign individual threads the task of solely controlling the sonification and vibration feedback cues. Another would be to include feedback when a user is close to a data point, but not necessarily over it. For instance, a slower (but noticeable) vibration could start when the user is within a certain proximity to a data point, serving as an alert that a data point is within reach. The introduction of this expanded operational radius could afford the tablet motor sufficient time to achieve noticeable vibration levels.

For histograms, participants would most commonly start exploring from the bottom left of the graph, then slowly move their finger to the right until they felt a bin. Once a bin was detected (sonification stops and vibration starts), participants tried to trace that bin as high as they could, until they felt their finger leave the bin. Once back in empty space (vibration stops, sonification resumes), participants would start moving either to the bottom of the tablet and then right or start moving to the right immediately. We believe this exploration strategy allowed the participants to identify both the number of bins and their individual height. If the participant were to reach the right-most end of the tablet before the 2-minute time limit was reached, the participant would either go back to the left-most side of the tablet and re-trace the histogram (see Figure 4, left) or start tracing back from the right of the tablet to the left (see Figure 4, right).

For scatter plots, participants typically started their exploration from either the top left or top right corner of the screen. They employed a sweeping motion, moving either left to right while descending (see Figure 5, left), or vertically from top to bottom while shifting sideways (see Figure 5, right). Upon detecting a data point (sonification stops and vibration starts), participants decelerated their sweeping motions to ascertain the presence of nearby data points. Consequently, participants predominantly focused their exploration around the scatter plot’s inflection points due to the higher point density in these areas, which facilitated easier detection. As the exploration time neared its end, participants reverted to quick sweeping motions in an attempt to identify data points across the remaining screen area.

One participant, Participant 4, employed a unique strategy of using two fingers throughout exploration - one as an anchor point, and one as an exploration point. This strategy was organically employed by the participant during training and was used across audio modalities and graphs, resulting in high mean re-creation scores of 14.7/16 for histograms and 14/16 for scatter plots. For

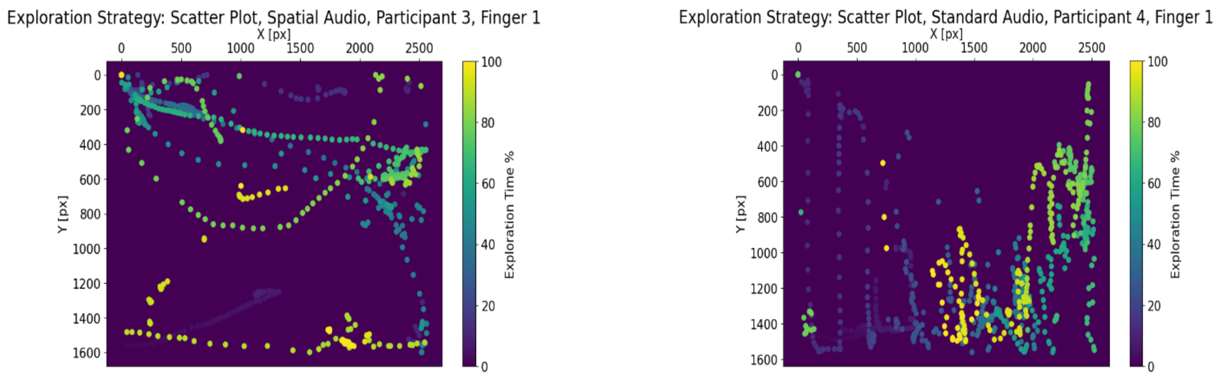

Figure 5: Main strategies when exploring scatter plots. Points are rendered in chronological order, with initial exploration points rendered with dark blue and changing to light yellow the longer a participant explores. Participant 3 (left) traced the graph from top to bottom, sweeping left to right, while Participant 4 (right) traced the graph from left to right, sweeping up and down.

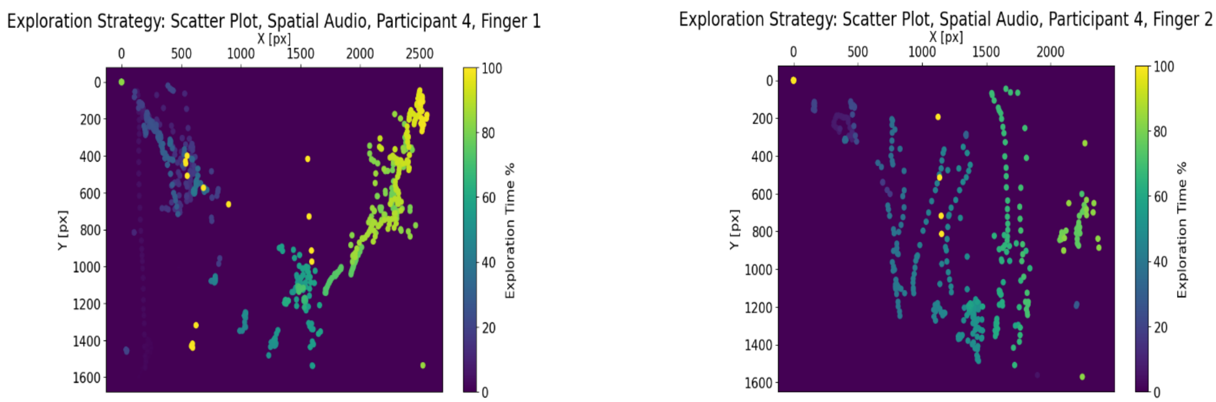

Figure 6: Participant 4's anchor-finger strategy when exploring the scatter plot with spatial audio and two fingers. Points are rendered in chronological order, with initial exploration points rendered with dark blue and changing to light yellow the longer a participant explores. Finger 1 (left) does not show clear sweeping motions, rather it shows different sections of the screen where Finger 1 was used as an anchor point. Finger 2 (right) followed sweeping motions starting in the top left, growing wider as the finger swept from left to right, until reaching the rightmost end of the screen.

histograms, Participant 4 had a clearly defined "anchor" finger and a free exploration finger. For scatter plots, the roles for the fingers could change depending on the exploration area on the screen. The unstructured nature of the data in scatter plots could be the reason for this observation, as the participant needed more reference (anchor) points in order to make sense of the mental image being created (see Figure 6). Additionally, Participant 4's increase in performance when using a two-finger strategy corroborates previous findings and observations. These include instances where participants proposed that multi-finger strategies could potentially augment their performance during a line-tracing study [19], instances where participants exhibited improved proficiency in discerning distinct vibration profiles on a touchscreen when employing multiple fingers [25], and instances where experimenters discovered that instructing participants on how to navigate graphs on touchscreens became more straightforward when an additional finger was used as a reference point [23]. This suggests that this anchor-finger strategy is worth investigating in further research, as it may prove to be a valuable tool that could assist other individuals with visual impairments when exploring multimodal histograms and scatter plots.

6 STUDY LIMITATIONS

This research illustrates the richness of iterative, design-based user studies, even with small sample sizes, uncovering qualitative insights that will shape the next iteration of multimodal graph rendering using spatial audio. Limitations of this study focus specifically on four main points:

1. Low Sample Size: Because the sample size of our study was small, we focused on descriptive statistics and qualitative

findings but we note that generalizability of such findings is reserved for larger scale studies in future work.

2. Complexity of Charts: While histograms and scatter plots offer more complex chart formats than simple bar charts or line graphs, we acknowledge that even the charts used in this work are relatively simplistic in design. We note, however, that this meets the research where it is, and this work uncovers challenges that need to be addressed before being able to successfully render more complex charts, which will also be a focus of future work.
3. Time Limits: In order to keep the total study time within a one-hour interval, we had to limit the exploration time to 2 minutes per graph. Some participants felt that this time was too little, and thus, the re-creation results may be lower because they were incomplete. Future work will dive deep into single chart types to allow for longer exploration and re-creation times.
4. Technical Limitations: This work illustrated technical limitations of multimodal rendering schemes, such as participants being unable to distinguish single points in high density point clouds and being unable to trace along key elements such as bars, in straight line paths. Future work will investigate how to enable such capabilities while working within the hardware and software constraints of the platform.

7 CONCLUSION AND FUTURE WORK

This work presented the design of a multimodal feedback scheme with integrated spatial audio for the exploration of both histograms and scatter plots on touchscreens and evaluated it through a research study with five BVI individuals. Design schemes with and without spatial audio were shared and made publicly available for future use. Results showed that spatial audio has potential to increase the identification of the trend-related insights when exploring data-intensive plots like histograms and scatter plots, at the cost of potentially skewing the mental representation of the graph. While more research is needed to ascertain the affordances and limitations of spatial audio in multimodal graph rendering, there was a high preference for using spatial audio among study participants.

Future work will focus on adding unique sonification to the top of histogram bins, using individual processor threads and early notification areas around scatter plot points to address motor delays in vibration response times, and testing different design iterations of this multimodal feedback scheme to determine the most efficient approach for providing information within a limited exploration time period. Future studies should also investigate the users' strategies when exploring histograms and scatter plots and compare them to the strategies presented in this work. This comparison would give developers more insight into the exploration strategies employed by BVI users that maximize accurate and efficient information extraction and learning when exploring multimodally rendered graphs on a touchscreen, which are critical to understand for supporting the training and adoption of new technological approaches.

Spatial audio conveys direct perceptual information about the arrangement of elements on screen - something that is not possible through natural language and alt-text descriptions alone. Thus, this work serves as a first stepping stone towards obtaining novel multimodal feedback schemes with integrated spatial audio to support BVI individuals in independently exploring and extracting trends from data-intensive graphs such as histograms and scatter plots.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under award numbers #2048428 and #2106394. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. In addition, the author(s) would like to express their sincere gratitude to the individuals who participated in this study.

REFERENCES

[1] M. F. Rice and K. R. Ortiz, ""Evaluating Digital Instructional Materials for K-12 Online and Blended Learning"," *TechTrends,* vol. 65, pp. 977-992, 2021.
[2] K. Tosh, S. Doan, A. Woo and D. Henry, Digital Instructional Materials: What Are Teachers Using and What Barriers Exist?, Santa, Monica: RAND Corporation, 2020.
[3] L. M. Federer, "Long-term availability of data associated with articles in PLOS ONE," *PLOS ONE,* vol. 17, pp. 1-13, August 2022.
[4] J. Zong, C. Lee, A. Lundgard, J. Jang, D. Hajas and A. Satyanarayan, "Rich Screen Reader Experiences for Accessible Data Visualization," *Computer Graphics Forum (Proc. EuroVis),* 2022.
[5] Omnicor Inc., *"Wikki Stix for the Sight-Impaired",* wikkistix.com, 2023.
[6] C. M. Baker, L. R. Milne, J. Scofield, C. L. Bennett and R. E. Ladner, "Tactile Graphics with a Voice: Using QR Codes to Access Text in Tactile Graphics," in *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility*, New York, NY, USA, 2014.
[7] C. M. Baker, L. R. Milne, R. Drapeau, J. Scofield, C. L. Bennett and R. E. Ladner, "Tactile Graphics with a Voice," vol. 8, January 2016.
[8] T. Götzelmann and K. Winkler, "SmartTactMaps: A Smartphone-Based Approach to Support Blind Persons in Exploring Tactile Maps," in *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, New York, NY, USA, 2015.
[9] J. Brzostek-Pawłowska, "Multimedia Mathematical Communication in a Diverse Group of Students," *Journal of Telecommunications and Information Technology,* p. 92–103, June 2019.
[10] M. Elgendy, T. Guzsvinecz and C. Sik-Lanyi, "Identification of Markers in Challenging Conditions for People with Visual Impairment Using Convolutional Neural Network," *Applied Sciences,* vol. 9, 2019.
[11] D. Mikułowski and J. Brzostek-Pawłowska, "Multi-Sensual Augmented Reality in Interactive Accessible Math Tutoring System for Flipped Classroom," in *Intelligent Tutoring Systems: 16th International Conference, ITS 2020, Athens, Greece, June 8–12, 2020, Proceedings*, Berlin, 2020.
[12] M. Minhat, N. L. Abdullah, R. Idrus and P. Keikhosrokiani, "TacTalk: Talking tactile map for the visually impaired," in *2017 8th International Conference on Information Technology (ICIT)*, 2017.
[13] Q.-W. Lin, S.-L. Hwang and J.-L. Wang, "Establishing a Cognitive Map of Public Place for Blind and Visual Impaired by Using IVEO Hands-On Learning System," in *Human-Computer Interaction. Users and Contexts of Use*, Berlin, 2013.
[14] T. Hu, S. Xuan, Y. Gao, Q. Shu, Z. Xu, S. Sun, J. Li and X. Gong, "Smart Refreshable Braille Display Device Based on Magneto-Resistive Composite with Triple Shape Memory," *Advanced Materials Technologies,* vol. 7, p. 2100777, 2022.
[15] Orbit Research, Graphiti®– a Breakthrough in Non-Visual Access to All Forms of Graphical Information, Orbitresearch.com, 2023.
[16] Freedom Scientific, *Focus 80 Blue 5th Gen,* Freedomscientific.com, 2023.
[17] H. P. Palani, P. D. S. Fink and N. A. Giudice, "Design Guidelines for Schematizing and Rendering Haptically Perceivable Graphical Elements on Touchscreen Devices," *International Journal of Human–Computer Interaction,* vol. 36, pp. 1393-1414, 2020.
[18] H. P. Palani, J. L. Tennison, G. B. Giudice and N. A. Giudice, "Touchscreen-Based Haptic Information Access for Assisting Blind and Visually-Impaired Users: Perceptual Parameters and Design Guidelines," in *Advances in Usability, User Experience and Assistive Technology*, Cham, 2019.
[19] J. L. Tennison and J. L. Gorlewicz, "Non-Visual Perception of Lines on a Multimodal Touchscreen Tablet," *ACM Trans. Appl. Percept.,* vol. 16, February 2019.
[20] N. A. Giudice, H. P. Palani, E. Brenner and K. M. Kramer, "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*, New York, NY, USA, 2012.

[21] M. E. Hahn, C. M. Mueller and J. L. Gorlewicz, "The Comprehension of STEM Graphics via a Multisensory Tablet Electronic Device by Students with Visual Impairments," *Journal of Visual Impairment & Blindness,* vol. 113, pp. 404-418, 2019.

[22] ViTAL, ViTAL: Equalizing Education through Accessible Experiences.

[23] J. L. Gorlewicz, J. L. Tennison, P. M. Uesbeck, M. E. Richard, H. P. Palani, A. Stefik, D. W. Smith and N. A. Giudice, "Design Guidelines and Recommendations for Multimodal, Touchscreen-Based Graphics," *ACM Trans. Access. Comput.,* vol. 13, August 2020.

[24] J. L. Tennison, P. M. Uesbeck, N. A. Giudice, A. Stefik, D. W. Smith and J. L. Gorlewicz, "Establishing Vibration-Based Tactile Line Profiles for Use in Multimodal Graphics," *ACM Trans. Appl. Percept.,* vol. 17, May 2020.

[25] J. Pasquesi and J. L. Gorlewicz, "Investigating Multi-Touch Vibrations on Mobile Touchscreens," in *2021 IEEE World Haptics Conference (WHC)*, 2021.

[26] F. Rumsey, Spatial Audio, New York: Routledge, 2001.

[27] L. McCormack and A. Politis, "SPARTA & COMPASS: Real-Time Implementations of Linear and Parametric Spatial Audio Reproduction and Processing Methods," in *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*, 2019.

[28] A. Hirway, Y. Qiao and N. Murray, "A QoE and Visual Attention Evaluation on the Influence of Spatial Audio in 360 Videos," in *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 2020.

[29] K. Yang, B. Russell and J. Salamon, "Telling Left From Right: Learning Spatial Correspondence of Sight and Sound," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[30] A. Walker and S. Brewster, "Spatial Audio in Small Screen Device Displays.," *Personal and Ubiquitous Computing,* vol. 4, pp. 144-154, June 2000.

[31] A. Pirhonen, S. Brewster and C. Holguin, "Gestural and Audio Metaphors as a Means of Control for Mobile Devices," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2002.

[32] J. M. Loomis, R. G. Golledge and R. L. Klatzky, "Navigation System for the Blind: Auditory Display Modes and Guidance," *Presence: Teleoper. Virtual Environ.,* vol. 7, p. 193–203, April 1998.

[33] Y. Vazquez-Alvarez, I. Oakley and S. Brewster, "Auditory display design for exploration in mobile audio-augmented," *Personal and Ubiquitous Computing,* vol. 16, pp. 1-13, December 2011.

[34] E. Connors, E. Chrastil, J. Sánchez and L. Merabet, "Action Video Game Play and Transfer of Navigation and Spatial Cognition Skills in Adolescents who are Blind," *Frontiers in human neuroscience,* vol. 8, p. 133, March 2014.

[35] L. Merabet, E. Connors, M. Halko and J. Sánchez, "Teaching the Blind to Find Their Way by Playing Video Games," *PloS one,* vol. 7, p. e44958, September 2012.

[36] N. Giudice, "Navigating without Vision: Principles of Blind Spatial Cognition," in *Handbook of Behavioral and Cognitive Geography*, vol. 1, 2018, pp. 260-288.

[37] R. Ramloll, W. Yu, B. Riedel and S. Brewster, "Using Non-speech Sounds to Improve Access to 2D Tabular Numerical Information for Visually Impaired Users," January 2001.

[38] S. Wall and S. Brewster, "Feeling What You Hear: Tactile Feedback for Navigation of Audio Graphs," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2006.

[39] D. Hayatpur, H. Xia and D. Wigdor, "DataHop: Spatial Data Exploration in Virtual Reality," in *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, 2020.

[40] S. H. Choi and B. N. Walker, "Digitizer Auditory Graph: Making Graphs Accessible to the Visually Impaired," in *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, 2010.

[41] L. M. Brown and S. Brewster, "Drawing by Ear: Interpreting Sonified Line Graphs," January 2003.

[42] R. L. Klatzky, N. A. Giudice and C. R. Bennett, "Touch-Screen Technology for the Dynamic Display of 2D Spatial Information Without Vision: Promise and Progress," *Multisensory Research,* 2014.

[43] J. R. Brown, S. A. Doore, J. A. Dimmel, N. Giudice and N. A. Giudice, "Comparing Natural Language and Vibro-Audio Modalities for Inclusive STEM Learning with Blind and Low Vision Users," 2023.

[44] S. N. Friel, F. R. Curcio and G. W. Bright, "Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications," *Journal for Research in Mathematics Education,* no. 2, pp. 124-158, 2001.

[45] P. Mishra, S. Kumar and M. K. Chaube, "Graph Interpretation, Summarization and Visualization Techniques: A Review and Open Research Issues," *Multimedia Tools and Applications,* vol. 82, no. 6, pp. 8729-8771, 2023.

[46] N. Glazer, "Challenges with graph interpretation: A review of the literature," *Studies in Science Education,* vol. 47, pp. 183-210, September 2011.

[47] M. Butler, L. M. Holloway, S. Reinders, C. Goncu and K. Marriot, "Technology Developments in Touch-Based Accessible Graphics: A Systematic Review of Research 2010-2020," *CHI Conference on Human Factors in Computing Systems (CHI '21),* no. Article 278, pp. 1-15, 2021.

[48] W. J. Robinson Moore, "CHI 2024 Spatial Audio Development Application," 2024. [Online]. Available: https://github.com/wrobinsonmoore/CHI_2024_Spatial_Audio_App. [Accessed 13 February 2024].

[49] G. Kramer, Auditory display : sonification, audification, and auditory interfaces, Santa Fe Institute, N.M.: Reading, Mass. : Addison-Wesley, 1994.

[50] davidgtr1, *"Ocean waves, Relaxing Waves, Tranquil Geo",* Freesound.org, 2023.

[51] Audacity Team, *"Audacity: Free, open source, cross-platform audio software",* Audacity Team, 2023.

[52] S. Paszkiel, P. Dobrakowski and A. Łysiak, ""The impact of different sounds on stress level in the context of EEG, cardiac measures and subjective stress level: a pilot study"," *Brain sciences,* vol. 10, p. 728, 2020.

[53] E. O. Dijk and A. Weffers, "Breathe with the Ocean: a System for Relaxation using Audio, Haptic and Visual Stimuli," 2011.

[54] A. Gupta, *"SimpleFingerGestures Android Library",* GitHub, 2018.

[55] A. Stefik, *"Quorum Language",* GitHub, 2023.

[56] Y. Su, X. Gao and X. Liu, ""How does frame-loss affect users' perception of smoothness?"," *CCF Transactions on Pervasive Computing and Interaction,* vol. 3, pp. 199-221, 2021.

[57] T. Watanabe and H. Mizukami, "Effectiveness of Tactile Scatter Plots: Comparison of Non-visual Data Representations," *International Conference on Computers Helping People with Special Needs,* vol. 10896, pp. 628-635, 2018.

[58] C. Engel, E. F. Muller and G. Weber, "SVGPlott: an accessible tool to generate highly adaptable, accessible audio-tactile charts for and from blind and visually impaired people," *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments,* pp. 186-195, 2019.