

TADA: Making Node-link Diagrams Accessible to Blind and Low-Vision People

Yichun Zhao
yichunzhao@uvic.ca
University of Victoria
Victoria, BC, Canada

Mahadeo A. Sukhai
mahadeo.sukhai@cnib.ca
Canadian National Institute for the Blind
Canada

Miguel A. Nacenta
nacenta@uvic.ca
University of Victoria
Victoria, BC, Canada

Sowmya Somanath
sowmyasomanath@uvic.ca
University of Victoria
Victoria, BC, Canada

ABSTRACT

Diagrams often appear as node-link representations in contexts such as taxonomies, mind maps and networks in textbooks. Despite their pervasiveness, they present accessibility challenges for blind and low-vision people. To address this challenge, we introduce Touch-and-Audio-based Diagram Access (TADA), a tablet-based interactive system that makes diagram exploration accessible through musical tones and speech. We designed TADA informed by an interview study with 15 participants who shared their challenges and strategies with diagrams. We refer to these insights as the *ladder of diagram access*. TADA enables people to access a diagram by: i) engaging in open-ended touch-based explorations, ii) searching for nodes, iii) navigating between nodes and iv) filtering information. We evaluated TADA with 25 participants and found it useful for gaining different perspectives on diagrammatic information.

CCS CONCEPTS

• **Human-centered computing** → **Accessibility systems and tools**; **Interaction techniques**; • **Hardware** → **Touch screens**.

KEYWORDS

Accessibility, Touch, Haptics, Pointing, Gestures, Individuals with Disabilities, Assistive Technologies, Artifact or System

ACM Reference Format:

Yichun Zhao, Miguel A. Nacenta, Mahadeo A. Sukhai, and Sowmya Somanath. 2024. TADA: Making Node-link Diagrams Accessible to Blind and Low-Vision People. 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, 20 pages. <https://doi.org/10.1145/3613904.3642222>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

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

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0330-0/24/05

<https://doi.org/10.1145/3613904.3642222>

1 INTRODUCTION

Larkin and Simon define diagrammatic representations as sets of spatially distributed components with information about the relationships between components [73]. Diagrams often take the form of node-link representations in their different variants (e.g., boxes and arrows, dots and lines [82, 97, 109]), which are a common and powerful way to represent non-linear information such as people networks (e.g., social network diagrams, family trees, organizational structures), relationships between concepts (e.g., mindmaps, taxonomies) and between the constituent elements of systems (e.g., computer network diagrams).

Previous literature has claimed that these representations are challenging to access for blind and low-vision (BLV) people [10, 88, 104]. One author of this paper also has direct lived experience of this as a scientist. Lack of accessibility has a potentially detrimental effect on equity and diversity in occupations that rely on these representations, but also on the daily lives and education of BLV people. However, we are not aware of existing empirical evidence on the importance of this problem. This raises questions such as: how do people currently access information in diagrams? And, to what extent do they face challenges and find workarounds?

We carried out a formative interview study with 15 people to answer these questions and to inform possible solutions. We found that lack of accessibility of diagrammatic representation is a frequent issue affecting BLV people, that their circumstances and motivations to access these are varied and multi-faceted and that existing solutions such as alt-text, asking other people and finding multiple sources for the same information are either insufficient or costly in terms of time, effort and social capital.

As a logical next step, we set out to integrate what we learned and existing experiences from literature (e.g. TeDUB [57, 71, 87, 88], A11yBoard [117] and GeoTablet [96]) to design a new interface that further advances the state of the art. This tool goes beyond existing solutions by: a) preserving spatial relationships from the diagram; b) enabling rich and multi-level access to the information in varied ways; and, c) retaining affordability. TADA runs on inexpensive tablets and integrates touch and speech input and audio output. It provides a range of interaction techniques on node-link representations to access overviews, navigation, information detail, filtering and finding.

We evaluated the atomic elements of the design with 25 people, which confirmed the practicality of our approach and offered ways to fine-tune the design. Our main contributions are:

- The results of an interview study on the importance of accessing diagrammatic information for BLV people.
- The empirically validated design of an interface that allows BLV people to access node-link diagrams.

2 RELATED WORK

This section surveys existing systems and techniques that make visual content accessible to BLV people, with a focus on diagrams. We use the term *diagram* to refer to graphics that represent objects and their relationships, such as node-link diagrams, flowcharts and UML diagrams. Diagrams can be stored as *graphs*. A graph is a mathematical abstraction and an abstract data structure that can be represented in several ways, one popular way being *node-link diagrams*. We use the term *chart* for statistical visualizations such as line and bar charts.

For input/output, we use *tactile* to refer to systems that use the sense of touch for feedback, such as raised-line overlays or pin arrays. We use *haptic* when there is force feedback, including through vibrotactile actuators. *Touch* refers to input from the location of bodily contacts with a surface (usually fingers on screens), even when it does not provide dynamic feedback through those contacts. *Proprioception* refers to people's awareness of the position of their own limbs (e.g., sensing a finger's position relative to another finger).

2.1 Non-spatial Systems

Screen readers (e.g., JAWS, NVDA) [42] read aloud *alternative text* and *image descriptions* to relay details about the visual content to BLV people. There are well-established guidelines for the generation of alt-text [7, 26, 38, 58, 61, 69], and we see a recent push to generate descriptions automatically [28, 34, 51, 56, 85]. Guidelines exist to improve accessibility for diagrams [104] as well as machine learning systems that automatically convert images of diagrams into machine representations [66, 68]. The quality of descriptions can make a difference in their usefulness, however, regardless of quality, alt-text often provides a description that is limited to a single purpose (e.g., a general description that will not suffice for someone looking for detail and a detailed description which is tedious for someone with a passing interest). Kim et al. [70] identify this and other limitations of these approaches and suggest an approach (simulated through *Wizard of Oz* techniques) where instead of a fixed description, users can query figures through natural language questions.

Further evolution of descriptions includes systems that provide hierarchical arrangements of information, which come closer to a continuous spatial approach. The user can move up, down or sideways in the hierarchy, usually with keyboard direction keys. This has been applied to UML and technical diagrams [57, 71, 88], display floor plans [87] and chemical molecule representations [22].

Another example of a system that represents diagrams without preserving spatial relationships is Kevin CASE, which uses a tactile overlay to interact with a table realignment of data flow diagrams for software engineers [15].

Our research differs fundamentally from these approaches because we support spatial perception of the continuous arrangement of data. We believe it leverages spatial memory and conveys important information (see Section 4 for further justification).

2.2 Chart Readers

A group of related systems addresses chart accessibility (e.g., line or bar charts). An initial set of such systems are image description generators specifically for charts [28, 34, 51, 56, 85]. More interactive systems provide translations of data to sound and speech that are navigable and therefore better adaptable to the user's particular needs. These systems differ in how the data is rendered. For example, SeeChart [2] provides a sophisticated way to navigate D3 charts¹ via keyboard, with speech output only and VoxLens has a similar approach but adds speech as input modality [94]. Infsonics [55] provides both audio sonification and speech controlled by keyboard. Other systems such as Chart Reader [103] and Zong et al.'s work [118] support accessible interactions of charts with screen readers. TactualPlot [30] facilitates the exploration of scatterplots through touch and audio. Finally, Slide-Tone and Tilt-Tone use a novel haptic output device to enhance the haptic and sound perception of line charts [41].

Our work differs from chart readers in the type of data represented; we focus on diagrams (formalizable as relationship graphs), whereas most work in charts focuses on the representation of specific types of visualizations such as line and bar charts, which require different interactions and rendering.

2.3 Map Accessibility Systems

Many systems have been proposed to represent maps non-visually. One group provides static tactile feedback (a printed raised-line overlay) with finger tracking, which allows the system to react differently (e.g., with sounds) to different finger positions. NOMAD [86], Mappie [21, 23] and the tactile graphics helper [43] all use variants of this input/output combination. In some cases, instead of fingers, the system senses tangible objects above the surface [37, 75, 77]. The audio output can be speech (Brock et al.'s system names double-tapped geographical objects [20]), or iconic sounds (Albouys-Perrois et al.'s system [3] beeps to indicate correct location of objects).

Replacing static printed overlays with array pins overcomes the fundamental limitation of static displays (which we also consider necessary to avoid in our design). Examples include Audio-Haptic browser [114], ATMap [115], as well as Shimada et al.'s [95], Schmitz et al.'s [93], Ivanchev et al.'s [59], and Holloway et al.'s [54] devices. Earlier systems are summarized by Vidal et al. [107]. A key drawback of these is their specialized hardware, which is still relatively rare, expensive and low-resolution.

Finally, some systems depict maps without relying on haptic means, which makes them more affordable [36]. Many combine computer-generated audio with vibrotactile feedback. Geotablet [96] plays a range of speech, iconic and environmental sounds when touched and TIKISI [9] offers a range of interesting interactions such as zooming, selection and overviews. TouchOver Map [89],

¹D3: <https://d3js.org/>

Timbremap [98] and Open touch/sound maps [62] have similar approaches.

We took inspiration from many of these designs. For example, Yairi et al.'s system [111] guides fingers from one node to another using pitch-modulated sound. However, map representation poses fundamentally different challenges to diagram representation: in maps, there is more emphasis on the shape and location of objects whereas in diagrams the representation of explicit relationships between different objects is key (in maps, these relationships are most often implicit). Consequently, in maps, egocentric navigation is often a key task, which is, arguably, not as relevant for diagrams.

2.4 Images and Raster Renderers

Some researchers have attempted non-visual rendering of generic images (e.g., photographs) to make them accessible through non-visual channels. Haptic renderings (e.g., pin arrays) of images might be too low resolution to effectively allow recognition of elements in raster representations, but with pan, zoom and animation [54] or clever pre-processing (e.g., filtering or simplification [91]) the issues can be ameliorated. Some have also tried audio renderings, for example, by mapping brightness to sound in a sweep of an image [17, 84], by sonifying the brightness of a raster image pixel by pixel depending on touch position [106], or by translating colours and edges of objects into the sound of different instruments [13]. This group of systems informs audio mapping in our work, but our approach differs fundamentally because our representation level is objects rather than pixels.

2.5 Vector- or Object-based Spatial Systems

Maćkowski et al. presented a system based on a raised-line overlay on top of a tablet to represent SVG graphics that users can query at different levels of detail by tapping [79]. *Ad hoc* overlay creation can be tedious, which can benefit from a toolchain as proposed by He et al. [50]. Linespace [99] and PantoGuide [27] are hardware-intensive tactile technologies. Linespace uses 3D printing to create raised line contours on a large surface. PantoGuide enhances static tactile graphics like charts with audio and a hand-mounted system, providing extra tactile information, such as the direction of a search target based on finger location. Also based on a glove, GraVVITAS [45] provides multi-finger vibrotactile feedback on top of a surface based on hand position.

Moving away from systems that require printing or use specialized hardware, Giudice et al. proposed an alternative to access audio descriptions of objects where, instead of a raised line, the high-end tablet vibrates in different ways when the finger traverses different types of graphical elements [44]. Alty and Rigas [4] carried out a relevant study where they investigated encoding the position of objects in the plane through audio timbre and pitch (not integrated into a system).

Two recent relevant systems that use only touch position as input and audio as output are A11yBoard [116, 117] and ImageAssist [83]. These relate to our own approach in that they provide more sophisticated sets of interaction techniques to access the data (segmented and semantically annotated pictures and artwork for ImageAssist and artboards for A11yBoard). The main difference

between these work and ours is that the interaction techniques necessarily differ because of the different types of target graphics.

2.6 Systems for Node-link and Diagram Display

A seminal example of a system for representing node-link diagrams is Audiograf, which provides speech descriptions of objects and other auditory cues under the finger of the user in a touch input device [67]. There are three systems that enable interaction with node-link diagrams for BLV people and go beyond the simple output of Audiograf. TeDUB [57, 71, 87, 88] and GSK [10, 11] share interaction through keyboard or joystick and therefore support navigation of diagrams that are relative (i.e., the current node's spatial position has to be updated mentally by the user as they traverse the diagram). This is fundamentally different from how our system and PLUMB [25, 31–33] operate, where the input device is absolute (stylus for PLUMB and fingers on a tablet for TADA), which makes the current location of exploration continuously accessible to the user through proprioception of their own limbs's position with respect to the device. However, TeDUB allows annotation, and GSK and PLUMB allow creation and editing of links, which we do not consider at this stage. PLUMB supports, however, relatively few interactions to perceive the diagram (it supports only object rendering and link following).

2.7 Interaction Techniques

A few researchers have studied and compared different ways to interact in audio or vibrotactile output touch location-based systems for specific tasks. For example, Kane et al. investigate invocation of commands on mobile phones for blind users [63], Tennison and Gorlewicz [101] compare techniques and strategies to perceive and follow lines, Ramôa et al. [90] and Kane et al. [64] compare ways to direct users to elements on the 2D space, with the latter also covering browsing regions and retrieving localized detail. We integrated this knowledge into the design of our system.

3 FORMATIVE INTERVIEWS: CHALLENGES WITH AND STRATEGIES FOR DIAGRAM ACCESSIBILITY

To understand how individuals currently access diagrams (node-link representations), including their challenges and strategies for access, we conducted remote semi-structured interviews with 15 individuals.

3.1 Participants

Fifteen adults (8 men and 7 women) ages 18 to 69 years (mean=47.8, median=50, SD=14.17) participated in our 50-90 minute-long interview study. Nine participants (P1-4, P6, P8, P11, P13, P15) are totally blind. Six of them were born totally blind and three lost their vision in adulthood. Six other participants (P5, P7, P9, P10, P12, P14) are legally blind or have low vision, with five of them having congenital or early-onset vision loss, and one experiencing recent vision loss. Appendix A includes other participant details. Our inclusion criteria required: 1) self-identification as legally blind; 2) reliance on the auditory channel or a combination of auditory and other sensory channels for accessing digital information; and

3) utilization of screen readers, braille-based systems, self-devised solutions, or alternative methods for digital information access. All participants encountered diagrams in personal, professional, or educational contexts, either daily or sporadically. Participants were asked to think about three examples of diagrams they had encountered in their lives prior to the study sessions and bring digital examples. Participants received CAD\$30 for their time. The study was pre-approved by the local Research Ethics Board (REB).

3.2 Procedure

The experimenter collected participant information on age, gender, background and vision loss. Then, participants described their encounters with diagrams and, using those as references, answered questions regarding information accessibility. The questions focused on the diagram, the context of the encounter, the importance of the information, motive(s) for accessing the diagram, tools used (if any), the types of tasks when using the diagram, the challenges encountered and their workarounds. The detailed study procedure is in Appendix B.

3.3 Data Collection and Analysis

We analyzed 16 hours of interview data using the thematic analysis method [1, 16] and iteratively coded the interview transcripts. First, one researcher open-coded two transcripts to bootstrap the codebook. With this initial set of codes, we collaboratively discussed and refined the codebook as we coded the remaining transcripts. This resulted in a total of 124 codes grouped into the themes we discuss below.

3.4 Findings

The most basic findings of our study validate the importance of diagrams for our participants. They showed us a multiplicity of diagrams and how they interact and struggle with them, such as textbook diagrams [P2-6, P9, P11, P14], networks and flowcharts [P1, P5, P10, P12, P14] and instructional diagrams [P2, P5, P11]. Participants also highlighted the personal and professional significance of many diagrams, arguing that diagrams enhance existing information [P1, P14], facilitate concept learning [P2-P6, P8, P12] and reflection [P12], are indispensable in certain professions [P1, P2, P7, P11], support recreational and personal activities [P8, P10, P11, P13] and enable collaborations with sighted people [P7, P12, P15].

Their challenges with diagrams were varied. They often missed their existence [P1, P4, P7-8, P12, P15], had unequal access [P1-2, P4-5, P8, P10-15], lacked control over the required level of detail [P1-5, P7-8, P10-12, P14], experienced unfair cognitive burdens such as having to memorize large chunks of information [P1-2, P4-5, P8-9, P11, P13-15] and incurred additional social costs such as those derived from bothering sighted colleagues or acquaintances [P1-3, P8-11, P13-14]. To overcome these challenges, they employ strategies such as asking sighted people for help [P1-13, P15], using OCR/CV tools [P4-5, P7, P9, P15] and finding multiple sources for the same information [P2-4, P6, P9-15].

Participants' comments were rich and insightful. During analysis, we realized that data themes aligned well with a linear structure of the participants' increasing ability to access information that we

call the *ladder of diagram access*, which we use to summarize and organize our description of the insights in the data.

3.4.1 Ladder of Diagram Access. We describe five levels of diagram access, from no access at all (level 1) to comprehensive (level 5). Within each level, we elaborate on aspects related to 1) individual differences in terms of abilities, motivations or tasks they wanted to accomplish, 2) types of diagram properties they attempted to access, and 3) external factors that influenced their process of accessing a diagram.

- L1: *Not Knowing the Existence of a Diagram:* In our study, 7 of 15 participants explicitly told us about experiencing a situation wherein they did not know if a diagram existed. For instance, P12 referred to a document with a diagram and showed us how, when accessed via a screen reader, it skipped the diagram completely: *"It tells us everything on the other [text] pages, and then you hit the page with the diagram, and it stops."* Similarly, P15 commented *"I honestly wouldn't be able to tell you if there was a diagram there [on a webpage] or not."* We classified such instances of interaction as belonging to level 1, which corresponds to the bottom rung on our ladder and refers to the most basic level of information access wherein people have to know the existence of an information representation.
- L2: *Knowing the Existence of a Diagram:* In our study, 2 participants explicitly described situations wherein an individual is made aware of the presence of some kind of embedded object in a document or website, but no further information is available. For example, during the interview, P9 showcased how a screen reader only informed them about a "graphic" and commented *"[it] means nothing to me"* and *"there's basically nothing to give [me] an idea [of the diagram]"*. Similarly, P14 showed us a diagram in a PowerPoint slide which read "embedded object" by the screen reader, signalling that there might be information, but the placeholder failed to inform P14 if the embedded object was a diagram, leaving P14 confused and *"[having] no idea what that [meant]."*

Without such access to the actual diagrammatic information, people cannot decide if the diagram is relevant to their needs and if they should access or ignore the diagram. P10 commented *"[I'm] constantly making the assessment [of] how worthwhile [a diagram is] for me to understand"* and four other participants confirmed this, but they cannot do so with just a placeholder.

We refer to such interactions as level 2 information access, wherein people know there is a placeholder but have too little information to determine whether the information in the visual element is important.

- L3: *Single Static Perspective of a Diagram:* Five participants in our study described examples wherein they were aware that there was a diagram and were able to learn more about it. This type of learning usually involves a static feature like alt-text. For example, P7 commented: *"If the human-generated alternative text is done well, then you can gain a very good picture of the diagram."* P7 and P9 used new computer vision technology to learn more about the diagrams. More commonly, a majority (14 of 15 participants) resorted to asking sighted people but found

that this workaround came with challenges (discussed in level 4).

Despite the benefits of static descriptions, 14 participants also voiced concerns over not having enough control over them, which limited how much they could gain. For example, P12 explained that an inherent limitation of using screen readers is that one does not have control over the types of details one wants to access: “[...] Sometimes, you just want to skim something, and you can’t. You can’t do the ‘cheaty’ short[cut] method of looking at the diagrams [visually] and inferring the information.” Another limitation of text-based descriptions such as alt-text, image descriptions and Optical Character Recognition (OCR) readouts was raised by P5 and P14, who mentioned how the linearity of these aids forces people to piece together different parts of the information on their own, which can be cognitively cumbersome. For example, P14 demonstrated accessing a node-link diagram about Belsky’s (1984) Process Model of the Determinants of Parenting found in a textbook through OCR. P14 described how, while they were learning about the details in the nodes through the textual descriptions, they missed the overall structure and when focusing on the structure, they could not remember the details.

Lastly, P3 and P10 complained about how the information made available is not necessarily the information one needs. P10 said, “In some instances, I could see individuals who [...] want to understand the design or want to understand the layout, and they just don’t want the [provided] statistical information [...] They want to understand the lines, the flows and that sort of information. So it’s a big challenge.” P3 repeatedly accessed the same diagram to enhance understanding but found the process frustrating: “I would try until the bitter end until it looks like I can’t handle this anymore, like I’m frustrated.” Another example of the mismatch between a static description and the user purpose is P3’s comment: “I just scroll past and don’t really think about it again, unless its information that’s important to me, and then I get frustrated that either it’s not all there, or it’s not written in a way that [...] immediately makes sense.”

Six participants in our study described how they sometimes needed to find information through complementary sources. For instance, P2 mentioned finding other people’s written journals when unable to access a diagram of hiking routes. P9 searched for alternative diagrams with text descriptions on search engines once they knew a bit more about the original diagram. However, this workaround is time-consuming and challenging. For example, P3 complained: “It can [be time-consuming and] also frustrating because [the information should be] right there and everybody else has it.”

We classify static descriptions and OCR output as belonging to level 3, wherein an individual knows that a diagram exists and is able to acquire at least one perspective from the diagram through these aids. Using such information, individuals are also able to complete some tasks. However, the lack of control over what information is available makes gaining a deeper understanding challenging and frustrating.

L4: Multiple Perspectives of a Diagram: Eleven participants in our study described scenarios when they attempted to acquire additional perspectives to complement static descriptions (such

as those belonging to level 3). For example, P15 often wants a global perspective first and only after a look at certain details: “You have to tell the basic first and then you tell the details of it.” P7 depicted how he achieves different perspectives by reorganizing the information in their own head: “Sometimes you have to come up with new ways of representing things, right? So that you know this relates to this, but in an alternative way it relates to that as well.”

Participants often reported scenarios wherein they reached out to sighted people (through an app, or in person) to acquire missing perspectives (n=14). For example, P5 uses “Be My Eyes” [39], which allows volunteers to access the mobile phone camera to help them. This way they can ask targeted questions to gain more understanding. Similarly, P3 uses Twitter: “I would probably just tweet the person and [...] ask a specific question about [the diagram].” and P8 mentioned having a direct dialogue with architects to clarify the exact structure of a building.

While such external help was meaningful, 9 participants also commented on the challenges of asking others. For example, the volunteer might not have the necessary expertise, patience or availability to explain the diagram well. P8 and P10 also explicitly highlighted the social costs. P8 told us how “you can sometimes drive people crazy when you really want to understand something, and some people don’t have the ability to explain it to you easily. People get frustrated because they want you to understand their explanation, and if you’re not getting it, then there’s a frustration level.”

We associate these interactions with level 4. Individuals are able to learn more about the diagram than in levels 1-3 by acquiring multiple perspectives of the information and querying. While the mechanisms for querying are not always convenient or sufficient, our participants find that reaching this point in their information access journey is helpful nevertheless.

L5: Effective and Comprehensive Access to Diagrams: As per the subsections above, BLV people are able, through existing technologies or workarounds, to access information in diagrams to accomplish virtually any required task.

However, the ability to access information still does not guarantee equity of information access for BLV people because current aids or strategies, although enabling, are costly in ways that can negate the benefits which motivate representing information as diagrams in the first place. For example, we know that visuals can reduce cognitive load for sighted individuals by, for example, allowing them to rely on spatial memory references instead of having to remember details [12] and the general advantages of diagrammatic representations are well known (e.g., [14, 73, 92]). Many of these advantages likely disappear if access is mediated in slow or cognitively expensive ways. P2 put this eloquently: “What is accessibility? Is accessibility just putting something out there? Or is it being able to give the person the same access, so that they are able to generate, have the same opportunity to get an equal level of understanding as another person?” P14 reflected that “a person with sight really could like, boom, they could kind of answer the simple question really quite quickly, just like looking visually at the [diagram ...] For me, it took more time and effort.”

The final level of the ladder is therefore defined by the efficiency and effectiveness with which people can access the information, not just by whether the information itself is accessible assuming almost unlimited supplies of cognitive abilities, time and effort. The system that we describe in the following sections intends to bring people *closer* to this level of accessibility, but we can only aspire to, not claim success in, making diagrams accessible at this level.

4 DESIGN GOALS AND PRINCIPLES

We set out to design and implement a tool that accomplishes two main overarching goals: **G1**: to help people access diagrammatic information in richer and more interactive ways (consistent with higher rungs of the ladder of diagram access—Section 3.4.1) and **G2**: to reduce common access barriers to this information, such as steep learning curves and expensive or unavailable hardware.

To achieve these goals we selected a set of design principles that guide our design. We chose the design principles to encode what we learned from the formative study (Section 3) and to apply existing knowledge from previous research (e.g., [47, 90, 101]). We applied the overarching goals and the design principles in all stages of the design process, from key foundational choices at the top level (e.g., which modalities to support), to the detail of interaction techniques. The design principles and goals are referenced in the next section to justify design decisions and features (e.g., **G1** or **DP3**).

DP1: *Rely on Spatiality*: Diagrams encode information in the spatial relations between the display objects as well as between objects and the diagram’s frame. For example, objects proximate in visual representations are interpreted as more related to each other than distant objects [73, 78, 110], and diagram designers often use layout to convey intended reading order, importance and other attributes (e.g., in an organization chart the CEO is at the top). We also know that viewers encode elements spatially in memory for agile retrieval while engaging with the visual representation [12], hence the importance of the *mental map* in studies of network visualization [6, 19, 81]. People without vision also rely on spatial object layout, they just do not access it visually, relying instead on other modalities [29] such as proprioception [112], touch [36] or echolocation [5, 102] to build their spatial understanding. In terms of the ladder, spatiality supports the transition from **L3** to **L4** because it enables multiple alternative points of access, akin to visual gaze for visual interfaces. Additionally, preserving spatiality benefits both sighted and BLV people, enabling communication through spatial deixis.

One of our key assumptions is that spatial layout and memory can be effectively leveraged by BLV users. This is consistent with Chundury et al.’s interview study with mobility experts [29], which highlights the importance of spatial awareness to understand data visualizations.

DP2: *Make element categories perceptually different*: Visual diagrams map certain types of elements (e.g., vertices, links and names in a network structure) to specific visual elements (e.g., red square boxes, arrows between boxes and text). An important, but often implicit, characteristic of these mappings is that elements are easily distinguishable from each other [72].

For example, nodes are immediately distinguishable from other elements such as labels or connections. Unlike in visual diagrams, where a long history of conventional mappings exists, other modalities require careful selection of representations in the output space (in our case, sounds) that are fast and clearly distinguishable from each other [29]. This works towards achieving quicker and more efficient understanding [L5].

DP3: *Support Multiple Levels of Information Access*: Our interviews showed that people have diverse varying circumstances and motives to access diagrams. For example, sometimes one wants to gauge the amount of information in a diagram (e.g., to decide whether to study it in more detail later). In some cases, the structure of the diagram is most important, whereas sometimes the specific details are crucial. To support people’s varied motives and circumstances, the design needs to implement multiple forms of access to information [L4], including summaries, overviews, finding and navigation, thereby allowing BLV people to choose the level of detail and granularity. As we have seen in Section 2, most existing systems do not provide comprehensive coverage for the variety of actions necessary in a wide variety of scenarios.

DP4: *Complementary and Simultaneous Interactions*: As a consequence of **DP3**, it follows that users should be able to choose the way they engage with diagrams depending on their purpose and circumstances, providing flexibility and choice in how they interact with diagrams [L4] as well as efficiency [L5]. This suggests an approach with multiple interaction techniques, perhaps some of which to achieve the same goal but from a different starting point or with different levels of precision. The interaction techniques could also be complementary or simultaneous; we know that users of tactile maps often use multi-handed operations with one finger or hand being used to keep memory [36, 101] or as a “bookmark”. Therefore we seek to design techniques that seamlessly take place simultaneously or compose and succeed each other.

5 TADA’S DESIGN

Here we describe TADA’s design, starting with a general overview followed by a description of its interaction techniques.

5.1 System Overview

TADA is currently implemented as a multi-platform application with Unity [105] and ChucK [8, 108] (See Appendix C). It is designed for tablet devices equipped with an inbuilt microphone and speaker. TADA’s input is touch (locations of fingers touching the tablet’s surface) and audio (speech commands), and its main output are sounds and speech synthesized in real time. The screen also shows a simple representation of the diagram for people with low or residual vision and for collaboration with sighted individuals (see [40]). The choice of platform and modalities is consistent with **G2**: it requires no specialized hardware and it can run on cheap tablets or even phones (although smaller screens will reduce leverage of **DP1**—see Subsection 7.2). TADA does not rely on vibrotactile input because most tablets do not offer it; however, the current design is compatible with adding this modality.

TADA is designed to access node-link diagrams; therefore the main objects that it represents are nodes, their connecting links, and the information contained in both (labels and attributes).

5.2 Design Process

We followed an iterative user-centered approach to which each author contributed, including one of the authors who is a BLV scientist and diagram user. The process gradually and iteratively transitioned from sketching, low-fidelity prototyping, to high-fidelity prototyping following Buxton’s design process model [24] over the course of 36 design sessions each ranging from 30 minutes to one hour. We devoted two one-hour sessions specifically for testing by our BLV author. In later stages of the design process the focus shifted to the refinement of interaction techniques to work well in conjunction with each other and adjusting threshold values to reasonable defaults (e.g., size of objects and touch drift tolerances). We regularly tested the design with different types of diagrams (e.g., networks and textbook diagrams), to accommodate a variety of node-link representations. We built a randomized diagram generator to help evaluate the robustness, adaptability and scalability of the prototypes.

5.2.1 Sound Design. The main output is audio of two kinds: abstract and speech. Following **DP2**, we map speech and abstract sound (which are easily distinguishable from each other) to different kinds of information: speech renders information contained in nodes and links (node and link names, attributes) and non-speech sounds (tones and musical instrument notes) signal the main objects and their structure. Within the non-speech sound category, we dedicated significant time to finding easily distinguishable sounds. We tested mappings from previous sonification literature [35, 47], environmental sounds, and a variety of simple waveforms (e.g., sine, triangle, saw-tooth) but we settled on utilizing musical instruments to create different timbres, balancing distinguishability, pleasantness and appropriateness.

A node produces a french-horn sound, with higher-pitch representing nodes with more connections. A link produces a plucked guitar sound, and a higher-pitched sound means it is shorter than other links (consistent with how longer strings vibrate at lower frequencies). A sustained french-horn or guitar sound indicates dwelling on a node or link respectively; A discrete sound indicates sweeping over a node or link. A continuous 150Hz pure tone (neither horn nor string) indicates the proximity to a diagram element.

5.3 Interaction Techniques

Our main way to instantiate **DP3** and **DP4** to achieve **G1** is to design a set of interaction techniques and pseudo-modes that support a variety of different ways to access the information contained in the diagram. One challenging aspect of the design process, which is difficult to articulate here in prose, was ensuring that the techniques are consistent and cohesive in both audio design and interaction. These techniques are intentionally crafted, whenever feasible, to work concurrently [**DP4**]. The Video Figure in the supplementary materials conveys some of these operations better than the overview (Figure 1). In the names of the interaction techniques below we use a “+” in the name when two fingers operate simultaneously within the same technique but with different functions. Whenever

a technique was inspired by a previous system, or when there is previous empirical evidence about the interactions that work best, we incorporate that knowledge and mention the sources.

Single-finger Sweep: As a finger sweeps over the screen, the nodes that are touched produce a discrete horn sound and the links that are crossed produce a discrete plucked string sound (Figure 1a). If the movement is faster, the sounds are shorter. If the finger stops on a node or a link, the sound is sustained indefinitely. A small degree of interaction hysteresis is applied in this case: if the movement is very slow, a node grows a bit to accommodate drift, assuming that the user just intends to keep a finger in that node. In practice, the Single-finger Sweep enables quick overviews of the contents of the diagram or a particular area and some degree of detailed exploration. However, it is not appropriate for precise detail retrieval, since that would require systematic tracing of the whole surface, which proved difficult and error-prone in our experience.

Several existing systems have a similar interaction. For example, in AudioGraf [67] and PLUMB [25, 31–33] a similar interaction technique is the only way to interact with the content.

Five-finger Dome: Users activate the Five-finger Dome by placing all the fingers of the hand on the tablet’s surface, which defines a pentagonal area. This plays an audio stream of sounds configured in the following way: one node inside the dome area plays first (french horn), then its connections (string sounds), then another connected node (still within the dome area), with its connection, and so on until all nodes in the dome area are exhausted. Then a short bell sound acts as a separator, and the same sequence plays again (unless the dome has changed position, in which case the sequence changes after the next bell). This technique provides a more refined way to explore connection patterns in localized areas. It is designed to produce sound textures that are perceptibly different depending on the connection patterns. For example, an area of nodes connected in a ring produces a homogeneous rhythmic alternation of node and link sounds, whereas a hub-and-spoke-like pattern produces a sequence that starts with a single node sound followed by many link sounds and then an alternation of node and link sounds. The duration of a full cycle is the same every time, so that denser areas produce faster sequences and vice versa; the overall duration of a cycle is configurable through the system settings.

This technique still supports overview, but in a more abstract way than the Single-finger Sweep, which would be difficult and tedious to use to recognize connectivity patterns. We also anticipated that the dome would require additional expertise to master. The technique is inspired by sonification systems, such as those that provide static renders of charts [4, 47]. However, to our knowledge, our system is the first to leverage higher-level sound textures interactively.

Dwell+Tap: Once a finger dwells on a node or a link², a tap by a second finger in the proximity of the dwelling finger causes the system to read (through speech synthesis) the name of the object and, for each consecutive tap, its attributes. We moved away from Kane et al.’s introduction of “Split Tap” [63, 64, 74], in which the second finger can tap anywhere, to enable multiple of these gestures

²Reaching a node can be achieved through Single-finger Sweep, described above, through Search and Follow, or by navigating the network, described below.

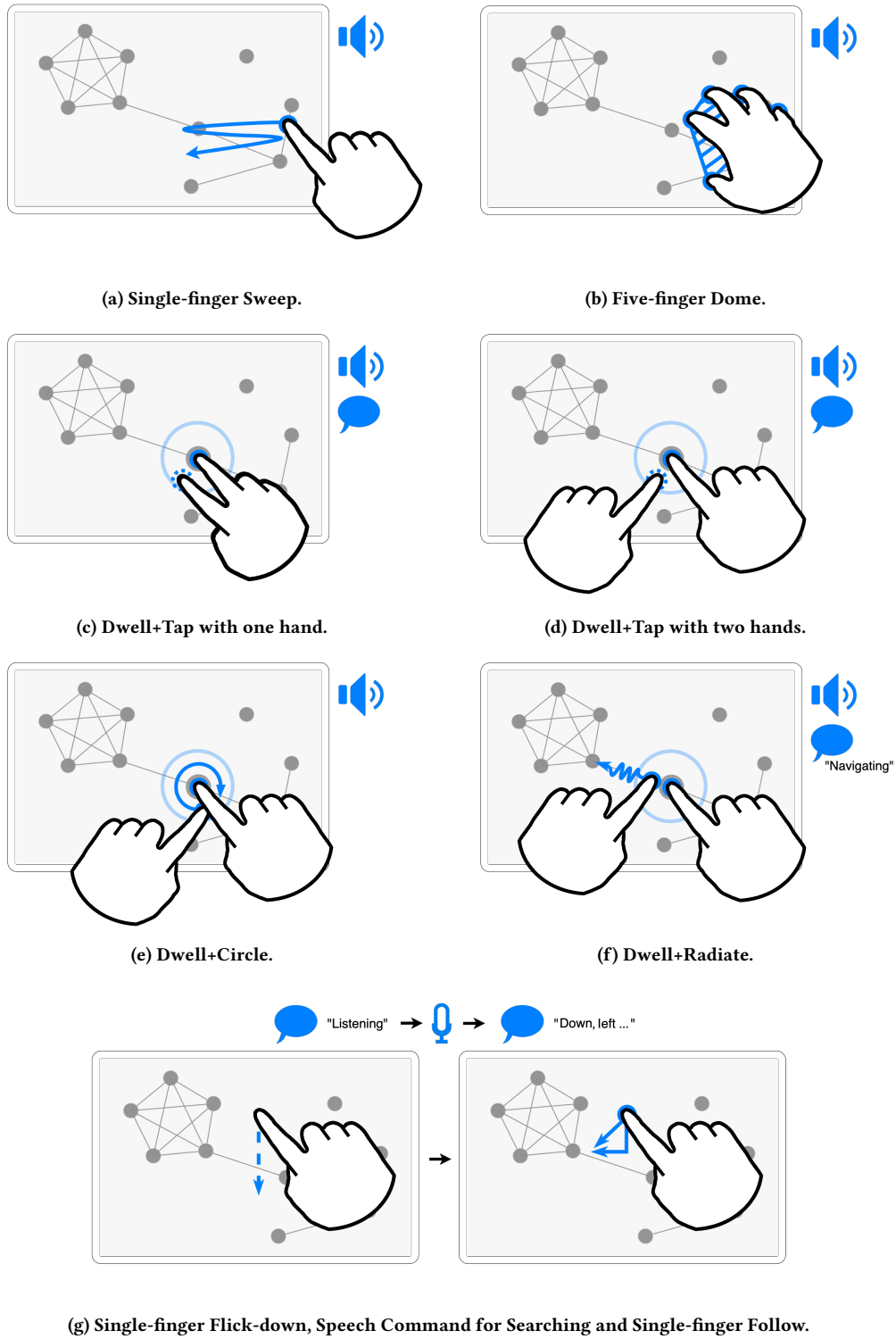


Figure 1: The seven interaction techniques supported by TADA. Continuous arrows indicate touch traces of the finger and discontinuous arrow are flicks. The speakerphone and speech bubbles indicate the use of abstract sounds and speech respectively.

simultaneously on several objects [DP4] and to make the system more robust against accidental taps [80].

Dwell+Circle: Once a finger is on a node, the user might want to know how many times and in which directions it is connected to other nodes. A second finger in proximity, orbiting around a finger dwelling on a node (which could have gotten there in a number of ways), triggers the Dwell+Circle technique. This stops the continuous sound from the dwell on the node and provides an additional tone that becomes louder as the finger approaches the angle of a link. As the link is crossed, it makes the regular string sound, and if the link is dwelt on, this can transition seamlessly to Dwell+Radiate below (in direct support of DP4 and consistent with Hinrichs and Carpendale’s findings about touch gesture transitions in public displays [52]). The fingers do not have to be from different hands but, in certain situations two-touch rotation with a single hand becomes awkward (see [53]).

A specific technique to find links such as Dwell+Circle is not strictly necessary, since a second finger doing Single-finger Sweep could potentially enable finding the links around. However, Dwell+Circle provides the following advantages: 1) it declutters the audio space by interrupting the node dwell sound; 2) it facilitates finding the links through the additional proximity guidance tone, or detecting right away that the node is not connected (absence of the tone); 3) it enables a seamless transition to the Dwell+Radiate technique; and, 4) it enables us to filter out nodes or links that are not relevant to the current node (e.g., close nodes, or links just crossing). We are not aware of a similar technique in any existing system.

Dwell+Radiate: After finding a node and a link of interest, the user will often want to follow the link to the corresponding neighbour node. This is a common form of navigation through the network. While one of the fingers still dwells on the node and the other finger is on a link (from the preceding Dwell+Circle gesture), the user can start moving this second finger away from the origin node (instead of the circular motion in Dwell+Circle), which triggers Dwell+Radiate. The tone sound now becomes louder as the finger approaches the neighboring node and plays a fanfare twist of the french-horn sound of the node to indicate success, at which point the original dwelt node finger can be released to start a new Dwell+Circle gesture. There are two main supports for the user tracing the line to the new node. First, the proprioceptive information of the two fingers can help the user move one finger away from the other (similar to a pinch gesture). Second, the string sound stops playing if the finger deviates from the string, at which point the user can seek it again by moving perpendicularly up and down. The design of this technique is inspired by existing work on tracing lines [46, 101], although it introduces new elements (e.g., filtering out other links, seamless transition to further navigation) and it is seen here for the first time together with other forms of interaction.

5.4 Pseudo-modes

Besides the interaction techniques described above, TADA offers several pseudo-modes, often triggered through speech commands, that support different kinds of access. A Single-finger Flick-down invokes the tablet’s speech recognition system, which responds

with “listening”. Recognized commands are Searching and Filtering, which we discuss below.

Searching with Single-finger Follow: After a Single-finger Flick-down, saying “Search for [information]” invokes the search mode. The user can search for information contained in any node, link, or for the links or node names themselves. For example, in a social network diagram, the user can flick down and utter “searching for Asha”. The system repeats the speech command from the user and processes the search. The system responds verbally with the results, for example, it could say “Nothing found”, or “Found [number of results] result(s)” depending on the search. To locate the results, the user places a single finger on the screen. The system then provides directions (left, right, bottom, up) verbally that direct the user’s finger to the location of the nearest result. For example, if the node for “Asha” is in the lower left direction relative to the finger, the system would say “down, left” on a loop. The verbal directions adapt to the movement of the finger. If the finger moves diagonally, the directional instruction combines horizontal and vertical prompts. Alternatively, when the finger moves only horizontally or vertically, the system will provide only directions in that axis until the horizontal or vertical location of the target is reached, and then the other directions are provided. The pacing of the speech prompts also encodes the remaining distance to the target. A slower pacing means the finger is closer. When the finger reaches a target node the system plays the same fanfare-like horn sound at the end of Dwell+Radiate.

We considered several ways to guide users to a target. We incorporate Ramôa et al.’s approach, which found that voice-based guidance is the most efficient method for helping people pinpoint a location, and it does not require prior training [90]. Similar techniques are utilized in ImageAssist [83] and AccessOverlays [64], but without the axial and proximity enhancements.

Filtering: The same Single-finger Flick-down triggers a filtering operation if the user says: “Filter by [attribute, value]”. The system recognizes the end of the speech command, repeats the command back and enters into filtering mode. The filtering mode reduces the complexity of a diagram by deemphasizing nodes and links that do not comply with the filter condition (a value of an attribute). For example, one can filter in a genealogical diagram by attribute gender: female. The attributes and values that users can filter for are the same ones accessible through the Dwell+Tap technique. When the filter is on, only the nodes that meet the filter criterion and their connections are fully accessible, but to preserve some degree of awareness of the filtered-out elements, these nodes and their connections still produce sounds, but at a fainter volume and with a static white noise played on top of them. This signifies that these elements have been filtered out and are not currently the focus of the person’s attention.

Additional Information: People can listen to the alt-text of the diagram by flicking a single finger to the right. Flicking with two fingers to the left plays the audio legend: a description of the audio mappings (e.g., plays the horn sound and says “node”).

Overall the techniques and pseudomodes intend to bring people from L3 up (we do not address L1-2). For example, Single-finger

Sweep, Five-finger Dome and Dwell+Circle all offer multiple [DP3], complementary [DP2, DP4] and spatially-indexed access [DP1] to locations decided/remembered by the user, enabling people to reach L4. Moreover, Dwell+Tap, Dwell+Circle and Dwell+Radiate are designed to alternate each other efficiently, bringing people closer to L5.

6 EVALUATIVE STUDY

The main goal was to test whether participants could use the fundamental features of TADA effectively and to detect usability issues. This formal study complemented the continuous informal testing of the system during design and implementation that we carried out on ourselves and others, including a blind co-author.

6.1 Participants

Twenty-five participants (14 female, 10 male, 1 transgender) who self-identified as legally blind and were 25 to 101 in age (mean=51.92, median=52, SD=16.35) volunteered for the study. Because visual perception can vary significantly, we analyze by group: group 1 (20 participants) could not visually access diagrams on the tablet at all; group 2 (5 participants) could perceive visual information from diagrams to a varying degree. The specifics of participants' demographic information are in Appendix D. Participants received CAD\$30 compensation. This study was pre-approved by the local REB.

6.2 Location and Apparatus

Experiments took place in 3 different locations, all dedicated private spaces with a table. TADA ran on a 10.6" Lenovo Tab M10 Plus tablet with Android 12 that we augmented with a thin 3D-printed bezel to prevent participants from inadvertently sliding their fingers off its edge. We used a 3D-printed tactile overlay of a simple diagram to speed up training. We captured the hands of the participant from above with a tripod-mounted video camera.

6.3 Procedure and Measurements

Participants provided consent and verbally answered demographic questions. After an introduction to the system's purpose, participants completed six blocks of activities. Each block consisted of a training phase with the tactile overlay and without it and then interaction with the tablet to answer questions or complete task requests. When the participant asked an additional question or required further assistance, the experimenter recorded the task as "completed with assistance". We designed the experiment's diagrams to be representative of reasonably complex diagrams that one might encounter in real life but also to have unequivocally correct or incorrect answers (see also scalability discussion—Subsection 7.2). The diagrams used for training and to answer the questions were different. After each block, the experimenter requested general comments about the interaction techniques. After all tasks, participants verbally answered a NASA TLX questionnaire. We changed the original TLX questionnaire scale from 21 to 10 points to make it more convenient for verbal interaction.

We collected: a) demographic data; b) correctness of answers or success in activity; c) comments on interaction techniques; d)

video of the tablet and hands; e) experimenter's observations and; f) NASA TLX responses. The experiment took 45-90 minutes.

6.4 Tasks and Stimuli

We designed six experimental task blocks, each testing an interaction technique or basic functionality of the system. For example, the first block tests participants' ability to achieve an overview with the Single-finger Sweep technique of Section 5.3.

The design of the tasks (Table 1, Appendix E) was informed by our formative study (tasks mentioned by participants) and prior work [76, 113]. These cover many tasks prevalent in node-link diagram research such as overview, clusters, adjacency, and connectivity aspects. Overview refers to the full diagram, whereas summary can be just of certain areas. We selected the experimental tasks to balance: a) being representative of a wide variety of common tasks; b) having quantifiable outcomes; and c) assessing one or more of the interaction techniques offered by TADA. We created the stimulus diagrams (Figure 2) to meet the goals and constraints of our experiment. We prioritized simple layouts to make the experiment manageable and easy to explain, to avoid confusion over required outcomes and to reduce time to achieve familiarity. Stimuli 1 and 2 contained multiple networks with different configurations within the same diagram so that we could evaluate their ability to contrast and compare quickly without excessive memory strain.

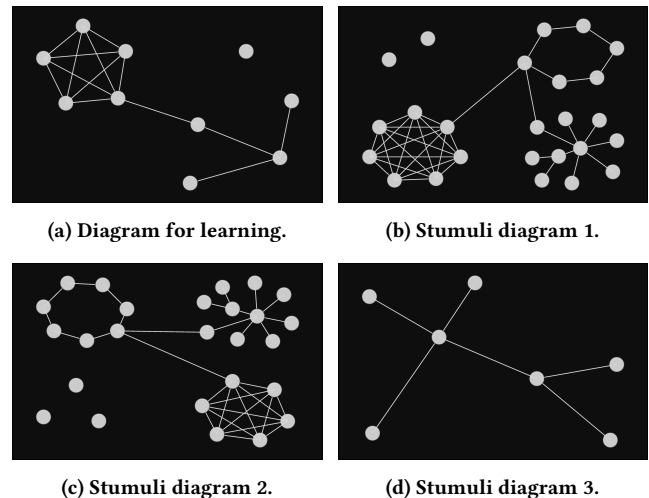


Figure 2: Diagrams presented in TADA as stimuli.

6.5 Data Analysis

We employed a mixed-methods approach to analyze the collected data. All the interview responses were analyzed qualitatively using a thematic analysis approach [1, 16], wherein we first used deductive coding and then organized the codes into themes using affinity diagramming. For questions Q1.2-Q6 we also counted the number of correct and incorrect responses. We referred to the gestures video to better understand participants' feedback regarding challenges and common causes of error but did not perform a systematic gesture analysis which we leave to future work.

Table 1: Summary of the six experimental blocks.

Block	Technique	#Q	Example Question/Action
Overview	Single-finger Sweep	5	Q1.1: Describe the overall layout of the diagram. Q1.2: Which quadrant of the diagram has the most connections between people? Q1.3: Which quadrant has the least connections? Q1.4: Which quadrant of the diagram has the most people? Q1.5: Which quadrant has the least people?
Summary	Five-finger Dome	5	Q2.1: Which quadrant of the diagram has the most connections between people? Q2.2: Which quadrant has the least connections? Q2.3: Which quadrant of the diagram has the most people? Q2.4: Which quadrant has the least people? Q2.5: Which quadrant has people with unequal numbers of friends?
Detail	Dwell+Tap	1	Q3: Given a person, what is the name and professional background associated with it?
Search	Flick Down, Speech Command, Single-finger Follow	1	Q4: Locate Bob.
Connections	Dwell+Circle	1	Q5: How many friends does Bob have?
Navigate	Dwell+Radiate	1	Q6: Navigate from Bob to any of his friends.

6.6 Results

This section summarizes the quantitative and qualitative results of our study.

6.6.1 Task and Question Outcomes. Figures 3a and b present a summary of participants’ responses to Q1.2-Q6. A majority of the participants from both groups answered all the questions correctly (22 out of 25) and 7 out of 25 participants required assistance with more than one question, such as additional clarifications regarding the underlying technical concepts. Lastly, there were three instances of three participants who did not reach the correct answer.

For Q1.1, all participants from both groups could correctly describe the overall layout of the diagram. Three participants [P1, P8, P14] described the layout by explaining how the nodes were spatially organized. For example, P8 from group 2 said: *“That’s a group of people [on the lower left ... There are] single individuals [on the upper left].”* Our expectation when designing TADA was that participants would also be able to identify details such as connectivity between regions and nodes, and we found this to be the case; 22 out of 25 participants from both groups described the diagram at a richer level. For example, P6 from group 1 described it as *“[On the lower right,] there is a person in the middle, I think, that’s connected to all other people around ... as I am going in circle here.”*

To address Q1.4-Q6, we occasionally provided additional clarifications or assistance to help the participant better understand the questions or to complete the task. For instance, Q1.4 and Q2.3 require identifying the quadrant with the most nodes. Three participants in Q1.4 and 2 participants in Q2.3 believed that the quadrant having the most connections is the one with the most nodes. To clarify the question, we described the difference between the density of nodes and the density of links, and this helped the participants correctly identify the answers. Another instance of such clarification involved prompting participants to think more carefully about the question. For example, to answer Q2.5 (people with unequal numbers of friends), three participants provided an answer but we recognized that were unsure about the underlying technical concept, and therefore, we prompted them to think more carefully

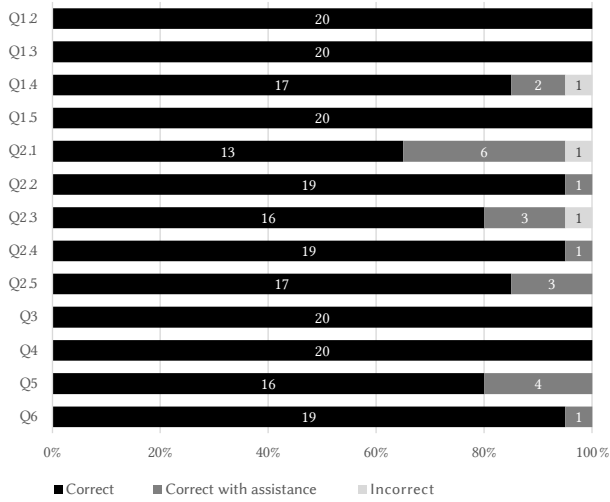
about what the question was asking and then they answered correctly. Six participants required assistance with their gesture(s), such as covering a quadrant more completely with the Five-finger Dome interaction.

Three participants from group 1 did not correctly answer Q2.1 (most connections in a quadrant), Q1.4 and Q2.3 (most nodes in a quadrant). They confused the region where only one node was highly connected with the region where all nodes were connected to each other. We do not think this was due to misinterpretation of the audio because they acquired correct answers for all remaining questions.

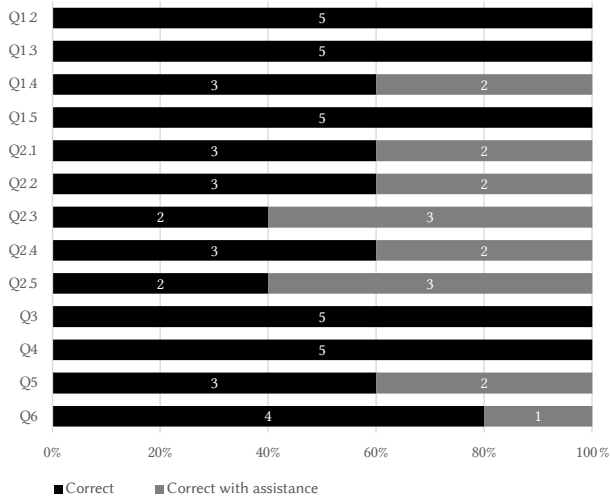
6.6.2 Perceived Workloads. Figure 4 summarizes participant responses to the NASA TLX questionnaire [48, 49]. Overall, we visually infer from Figure 4 that the workload ratings are generally towards the left side of the tables (i.e., the workloads were generally considered low). Mental demand (group 1’s mean rating: 3.35; group 2: 3.8), physical demand (1: 2.15; 2: 4.6) and effort (1: 2.7; 2: 3.2), received slightly higher mean ratings from both groups than other categories, for which the mean ratings were below 2. Participants who gave higher scores in these categories reasoned that the cognitive effort needed to understand and interpret the audio information was relatively high (n=4) and so was the physical effort for performing some of the gestures (n=2). P8 explicitly attributed the workload to the effort required to answer the questions due to their age (101 years). In the cases of physical demand, this was due to a lack of experience with performing gestures [P8] and low dexterity of fingers [P12].

6.6.3 Participant Feedback. Overall, participants described TADA as *“fun”* [P2, P11, P16], *“intuitive”* [P3, P4, P5, P22], *“useful”* or *“helpful”* [P1, P3, P4, P5, P7, P11] and *“innovative”* [P1, P18, P22, P25].

Audio Design. Eight participants explicitly noted that associating distinct sounds to nodes and links was a good strategy. P5 complimented: *“I like the distinction of the sounds. You know, you got a very distinct pluck with the string, and then the horn honks. [And] the navigational sound [of] orientation [is] a good contrast with the*



(a) Group 1 (no visual diagram perception) results.



(b) Group 2 (partial visual diagram perception) results.

Figure 3: Correctness of answers to quantitative questions listed in Table 1.

other two.”. P10 found the use of speech and non-speech sounds for different purposes effective: “When you are just gathering a sense of the general layout, I think the sounds are perfect because if [the system] starts to tell you [such information in speech] all at once, you cannot distinguish [the information]”.

Participants found that the audio they received when performing specific interaction techniques was helpful. Participants found the audio feedback when performing the Single-finger Sweep was informative to gain an overview of a diagram (n=16). The paced link sounds heard when using the Five-finger Dome gesture helped many understand the ways in which nodes were connected within a sub-region of the diagram (n=11). Four participants explicitly noted that speech was appropriate to convey detailed information using

	1	2	3	4	5	6	7	8	9	10
Mental	1	4	8	3	2	2	0	0	0	0
Physical	8	6	3	1	2	0	0	0	0	0
Temporal	14	5	1	0	0	0	0	0	0	0
Performance	8	9	2	1	0	0	0	0	0	0
Effort	4	9	4	0	1	1	1	0	0	0
Frustration	14	2	3	0	0	1	0	0	0	0

(a) Group 1 NASA-TLX results.

	1	2	3	4	5	6	7	8	9	10
Mental	0	2	0	0	3	0	0	0	0	0
Physical	0	1	2	0	0	0	1	1	0	0
Temporal	3	1	1	0	0	0	0	0	0	0
Performance	1	2	1	1	0	0	0	0	0	0
Effort	0	2	2	0	0	1	0	0	0	0
Frustration	3	0	2	0	0	0	0	0	0	0

(b) Group 2 NASA-TLX results.

Figure 4: NASA-TLX workload rating histograms. The number in each cell indicates the number of participants selecting that workload rating.

the Dwell+tap interaction. Lastly, 3 participants explicitly noted that the directional feedback when performing the Single-finger Follow gesture was logical, helpful and efficient.

Participants also pointed to some limitations of our audio design. Three participants [P2, P10, P11] mentioned that the audio summary from the Five-finger Dome interaction (Block 2) was overwhelming and difficult to process. In Q2.1, five participants confused the connected links sound of one node as referring to information about the whole quadrant. Four of these participants gave the correct answer after we reminded them of how the audio was designed.

Interaction Techniques. Participants provided mixed feedback on the interaction techniques. For instance, P9 and P19 mentioned that the Five-finger Dome (Block 2) gesture was easy to perform. For the Dwell+Circle (Block 5) gesture P14, P23, and P24 found it easy to perform and P3, P5, and P18 said that it provided clear indications of the directions of connected links. Seven participants noted Dwell+Tap (Block 3) was simple, and P6 especially liked the “stability” and “consistency” of repeated tapping during dwelling. As for Dwell+Radiate (Block 6), P3, P16, and P23 found it to be straightforward and easy to use.

Some participants needed assistance with performing the gestures, especially the Five-finger Dome and the Dwell+Circle. Three participants [P8, P13, P16] commented that the Five-finger Dome gesture felt awkward to perform. We think this was partly due to this gesture requiring participants to maintain consistent contact by applying a certain level of pressure. Six participants also found Dwell+Circle awkward to perform. One reason for this was that the dwelling finger was observed to be sometimes drifting as participants performed the circling gesture, resulting in accidental interactions with other nodes or links. Additionally, because the node sound is muted during this interaction to prioritize the proximity guidance audio, participants also did not realize that their

finger had drifted unless they lifted it up and placed it down to confirm that the dwelling finger was still on the node.

Five participants [P2, P4, P11, P12, P25] attributed their difficulty with performing gestures to individual experiences. For example, they highlighted that varied motor skills [P2, P11, P12] can challenge touch-based interactions. P2, P4, and P24 noted that more familiarity with the device’s dimensions and gesture detection could enhance interaction. Lastly, three participants also had long nails that made it difficult to perform the gestures precisely or comfortably.

7 DISCUSSION

This Section addresses the evaluation study, the system design and its applicability.

7.1 Discussion and Limitations: Evaluative Study

Most participants successfully completed the experimental task blocks, with some occasional assistance. Our results highlight that a majority of the participants could explore the spatial layout (**DP1**—participants correctly answered spatial questions) and distinguish between the different graph elements (**DP2**—participants correctly associated sounds and speech to different elements to answer questions about nodes, links and information contained therein). Additionally, each interaction technique and pseudo-mode enabled participants to gain different perspectives about the diagram (**DP3**—participants correctly identified different levels of information from overview to detail, from global to local and from spatial to semantic).

Designing accessible gestures is a complex problem. Our design aligns with existing guidelines towards this goal such as Kane et al.’s [65], who recommend multi-touch gestures for BLV people and suggest allowing approximate targeting methods [65]. Providing audio cues for targets can also help [46]. We learned about two additional low-level issues: feedback about the required amount of pressure people have to apply is often required (e.g., this was a challenge for the Five-Finger Dome), and dwell-based gestures suffer from finger drifting (e.g. with Dwell+Circle). Further exploring how such challenges can be addressed will help in designing accessible gestures.

However, there are also areas that require further investigation. While TADA’s design addresses **DP4** by enabling simultaneous gestures, our tasks did not evaluate this scenario. We also did not perform systematic analyses of the video recordings to investigate details of gesture performance. In future, we will carry out a detailed gesture analysis of the video and a study that tests the benefits of more complex interaction chains. Moreover, our stimuli were relatively simple and constrained by the study’s goals. An “in-the-wild” exploration of diagram access can advance our understanding of situational aspects and the scalability of our approach.

7.2 Discussion and Limitations: TADA’s Design

Our formative study gathered evidence supporting the importance of making diagrams accessible to BLV people and exposed the challenges and frustrations of accessing this kind of information. We also learned that it is important to make information available at

least at **L3**—static descriptions. This finding aligns with current academic and commercial solutions that focus on improving accessibility to visual elements by encouraging alt-text and improving image descriptions [7, 38]. However, textual descriptions are often not enough because people require different information depending on their circumstances. Our formative interviews encouraged us to facilitate access to multiple perspectives of a diagram [**G1**]. This can be achieved in different ways. For example, one could ask a human (a common workaround) or a large language model to retrieve a particular aspect of a diagram (e.g., the number of nodes, whether it is densely connected). Some available systems such as Bing already enable this kind of interaction³. Nevertheless, diagrams are often created because they offer a better representation of the subject matter than text; for example, for representing structures that are inherently not linear. Moreover, we often access diagrams not to obtain a piece of information, but to remember topics or to build our understanding and mental models [73]. TADA was designed under the assumption that spatial interactive engagement with a diagram is as beneficial for BLV people as it is for sighted people. Thus, we encoded this as **DP1**. The comments from the formative interviews and the results from the evaluation study show early support for our proposed approach.

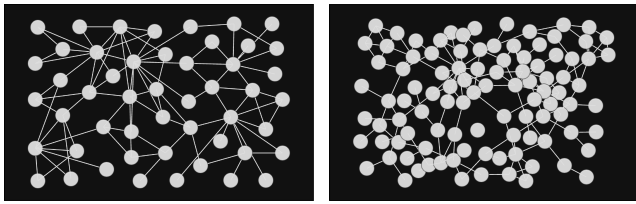
The current implementation of TADA chiefly uses touch input (without tactile feedback) and audio output to provide interactivity. However, other modality choices are possible, such as vibrotactile output (e.g., [89, 96]) and pin arrays [107], which are progressively becoming more affordable and higher resolution. We think that some of these modalities offer advantages for some tasks. For example, following a link as done in our study might be much faster with a tactile interface, especially if challenges around resolution of the tactile display (e.g., pin arrays) can be improved. Exploring such solutions and comparing the benefits and limitations of different modalities may be useful for the design of future interfaces.

Some of the design choices that we made, such as not implementing different zoom levels or having a pannable canvas (e.g., [83, 93, 115]) constrain the complexity of diagrams displayable by TADA. We chose not to support panning or zooming because we think that they would prevent users from constructing a useful spatial mental map of the content [**DP1**] (this also applies to maps [36], which our formative study participants’ experiences also confirmed). Nevertheless, it is realistic to expect more complex diagrams (e.g., Figure 5ab) which would be difficult to explore from small screen spaces. Figure 5a shows the most complex diagram that we have found, informally, to still be interactable with in TADA with a 10.6" tablet. In the future, research is required to better understand how spatial understanding can be retained while enabling the exploration of larger pieces of information, especially on smaller displays like tablets and mobile phones.

7.3 Applicability of TADA

The current version of TADA loads diagrams described in GraphML [18]. Although most existing diagrams are not in this format, we expect the translation from other formats, including raster images, to be addressed by advances in computer vision (see, e.g., [66, 68, 85]).

³E.g., Bing: <http://www.bing.com>



(a) A diagram with maximum complexity. (b) A diagram with too much complexity.

Figure 5: Complex diagram examples for TADA.

TADA is currently designed to represent node-link diagrams. We believe that most diagrams can be represented in the graph structures offered by node-link diagrams. For example, a genealogical tree is a graph with people as nodes and a few types of edges between them. Some other chart types, such as Gantt charts, can also be expressed this way but their representation might be more awkward.

However, there are three diagram features that TADA does not currently implement and would extend the applicability of the system. First, TADA only supports non-directional links. Second, in its current version, TADA only uses straight links, which can make it inelegant to display some types of layouts. Third, a mechanism to convey the area or region where nodes exist can further facilitate representation through containment.

8 CONCLUSION

We presented insights into approaches for improving diagram accessibility for BLV people. We conducted a formative study and learned about the challenges with diagram access. We summarized these challenges and some workarounds participants employed as steps in a “ladder of diagram access” that ranges from not knowing that a diagram even exists to taking advantage of a diagram as a representation superior to text. Informed by our study, we designed TADA, a tool that provides multi-faceted access to node-link diagrams through the careful design and selection of a range of complementary interaction techniques, modes and sound designs that work on an inexpensive tablet. Based on an evaluation study we contribute an understanding of how audio and touch-based interactions can be leveraged to make diagrams more accessible.

ACKNOWLEDGMENTS

We thank Alex Thomo for his invaluable guidance at the initial stages of this work. This research is supported by the University of Victoria, NSERC DG 2020-04401, and the NSERC CREATE VADA graduate program.

REFERENCES

- [1] Anne Adams, Peter Lunt, and Paul Cairns. 2008. *A qualitative approach to HCI research* (1 ed.). Cambridge University Press, 138–157. <https://doi.org/10.1017/CBO9780511814570.008>
- [2] Md Zubair Ibne Alam, Shehnaz Islam, and Enamul Hoque. 2023. SeeChart: Enabling Accessible Visualizations Through Interactive Natural Language Interface For People with Visual Impairments. (Feb 2023). <https://doi.org/10.1145/3581641.3584099> arXiv:2302.07742 [cs].
- [3] Jérémy Albouys-Perrois, Jérémy Laviole, Carine Briant, and Anke M. Brock. 2018. Towards a Multisensory Augmented Reality Map for Blind and Low Vision People: a Participatory Design Approach. In *Proceedings of the 2018 CHI*

- Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–14. <https://doi.org/10.1145/3173574.3174203>
- [4] James L. Alty and Dimitrios Rigas. 2005. Exploring the use of structured musical stimuli to communicate simple diagrams: the role of context. *International Journal of Human-Computer Studies* 62, 1 (Jan 2005), 21–40. <https://doi.org/10.1016/j.ijhcs.2004.08.003>
- [5] Ronny Andrade, Jenny Waycott, Steven Baker, and Frank Vetere. 2021. Echolocation as a Means for People with Visual Impairment (PVI) to Acquire Spatial Knowledge of Virtual Space. *ACM Transactions on Accessible Computing* 14, 1 (Mar 2021), 1–25. <https://doi.org/10.1145/3448273>
- [6] Daniel Archambault and Helen C. Purchase. 2013. Mental Map Preservation Helps User Orientation in Dynamic Graphs. In *Graph Drawing*, Walter Didimo and Maurizio Patrignani (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 475–486.
- [7] American Anthropological Association. 2022. Guidelines for Creating Image Descriptions. Retrieved 2023 from <https://americananthro.org/accessibility/image-descriptions/>
- [8] Jack Atherton and Ge Wang. 2018. Chunity: Integrated Audiovisual Programming in Unity. In *Proceedings of the 2018 conference on new interfaces for musical expression*. 6.
- [9] Sina Bahram. 2013. Multimodal eyes-free exploration of maps: TIKISI for maps. *ACM SIGACCESS Accessibility and Computing* 106 (Jun 2013), 3–11. <https://doi.org/10.1145/2505401.2505402>
- [10] Suzanne P. Balik, Sean P. Mealin, Matthias F. Stallmann, and Robert D. Rodman. 2013. GSK: universally accessible graph sketching. In *Proceeding of the 44th ACM technical symposium on Computer science education*. ACM, Denver Colorado USA, 221–226. <https://doi.org/10.1145/2445196.2445266>
- [11] Suzanne P. Balik, Sean P. Mealin, Matthias F. Stallmann, Robert D. Rodman, Michelle L. Glatz, and Veronica J. Sigler. 2014. Including blind people in computing through access to graphs. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '14)*. Association for Computing Machinery, New York, NY, USA, 91–98. <https://doi.org/10.1145/2661334.2661364>
- [12] Dana H. Ballard, Mary M. Hayhoe, Polly K. Pook, and Rajesh P. N. Rao. 1997. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20, 4 (1997), 723–742. <https://doi.org/10.1017/S0140525X97001611>
- [13] Michael Banf, Ruben Mikalay, Baris Watzke, and Volker Blanz. 2016. PictureSensation – a mobile application to help the blind explore the visual world through touch and sound. *Journal of Rehabilitation and Assistive Technologies Engineering* 3 (Oct 2016), 2055668316674582. <https://doi.org/10.1177/2055668316674582> Citation Key: banf-PictureSensationMobileApplication-2016.
- [14] Adam Binks, Alice Toniolo, and Miguel A. Nacenta. 2022. Representational transformations: Using maps to write essays. *International Journal of Human-Computer Studies* 165 (2022), 102851. <https://doi.org/10.1016/j.ijhcs.2022.102851>
- [15] Paul Blenkhorn and D Gareth Evans. 1998. Using speech and touch to enable blind people to access schematic diagrams. *Journal of Network and Computer Applications* 21, 1 (1998), 17–29. <https://doi.org/10.1006/jnca.1998.0060>
- [16] Richard E Boyatzis. 1998. *Transforming qualitative information: Thematic analysis and code development*. sage.
- [17] Vincent Bragard, Thomas Pellegrini, and Julien Pinquier. 2015. *Pyc2Sound: a Python tool to convert images into sound*. 4 pages. <https://doi.org/10.1145/2814895.2814912> Citation Key: bragard-Pyc2SoundPythonTool-2015.
- [18] Ulrik Brandes, Markus Eiglsperger, Jürgen Lerner, and Christian Pich. 2010. *Graph Markup Language (GraphML)*. CRC Press, Taylor & Francis Group, Boca Raton.
- [19] Anke Brock. 2014. Interactive Maps for Visually Impaired People: Design, Usability and Spatial Cognition. (2014).
- [20] Anke Brock and Christophe Jouffrais. 2015. Interactive audio-tactile maps for visually impaired people. *ACM SIGACCESS Accessibility and Computing* 113 (Nov 2015), 3–12. <https://doi.org/10.1145/2850440.2850441>
- [21] Anke Brock, Philippe Truillet, Bernard Oriola, and Christophe Jouffrais. 2014. Making Gestural Interaction Accessible to Visually Impaired People. In *Haptics: Neuroscience, Devices, Modeling, and Applications (Lecture Notes in Computer Science)*, Malika Auvray and Christian Duriez (Eds.). Springer, Berlin, Heidelberg, 41–48. https://doi.org/10.1007/978-3-662-44196-1_6
- [22] Andy Brown, Steve Pettifer, and Robert Stevens. 2003. Evaluation of a non-visual molecule browser. In *Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility (Assets '04)*. Association for Computing Machinery, New York, NY, USA, 40–47. <https://doi.org/10.1145/1028630.1028639>
- [23] Emeline Brule, Gilles Bailly, Anke Brock, Frederic Valentin, Grégoire Denis, and Christophe Jouffrais. 2016. MapSense: Multi-Sensory Interactive Maps for Children Living with Visual Impairments. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 445–457. <https://doi.org/10.1145/2858036.2858375>
- [24] Bill Buxton. 2010. *Sketching user experiences: getting the design right and the right design*. Morgan kaufmann.
- [25] Matt Calder, Robert F. Cohen, Jessica Lanzoni, and Yun Xu. 2006. PLUMB: an interface for users who are blind to display, create, and modify graphs. In

- Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06)*. Association for Computing Machinery, New York, NY, USA, 263–264. <https://doi.org/10.1145/1168987.1169046>
- [26] DIAGRAM Center. 2019. Image Description Guidelines. Retrieved 2023 from <http://diagramcenter.org/table-of-contents-2.html/>
- [27] Elyse D. Z. Chase, Alexa Fay Siu, Abena Boadi-Agyemang, Gene S-H Kim, Eric J Gonzalez, and Sean Follmer. 2020. PantoGuide: A Haptic and Audio Guidance System To Support Tactile Graphics Exploration. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event Greece, 1–4. <https://doi.org/10.1145/3373625.3418023>
- [28] 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 (2019), 249–260. <https://doi.org/10.1111/cgf.13686> Citation Key: choi-VisualizingNonVisualEnabling-2019a.
- [29] Pramod Chundury, Biswaksen Patnaik, Yasmin Reyazuddin, Christine Tang, Jonathan Lazar, and Niklas Elmqvist. 2022. Towards Understanding Sensory Substitution for Accessible Visualization: An Interview Study. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (Jan 2022), 1084–1094. <https://doi.org/10.1109/TVCG.2021.3114829> Citation Key: chundury-UnderstandingSensorySubstitution-2022.
- [30] Pramod Chundury, Yasmin Reyazuddin, J. Bern Jordan, Jonathan Lazar, and Niklas Elmqvist. 2023. TactualPlot: Spatializing Data as Sound using Sensory Substitution for Touchscreen Accessibility. *IEEE Transactions on Visualization and Computer Graphics* (2023), 1–11. <https://doi.org/10.1109/TVCG.2023.3326937>
- [31] Robert F. Cohen, Valerie Haven, Jessica A. Lanzoni, Arthur Meacham, Joelle Skaff, and Michael Wissell. 2006. Using an audio interface to assist users Who are visually impaired with steering tasks. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06)*. Association for Computing Machinery, New York, NY, USA, 119–124. <https://doi.org/10.1145/1168987.1169008> Citation Key: cohen-UsingAudioInterface-2006b.
- [32] Robert F. Cohen, Arthur Meacham, and Joelle Skaff. 2006. Teaching Graphs to Visually Impaired Students Using an Active Auditory Interface. *SIGCSE Bull.* 38, 1 (mar 2006), 279–282. <https://doi.org/10.1145/1124706.1121428>
- [33] Robert F. Cohen, Rui Yu, Arthur Meacham, and Joelle Skaff. 2005. PLUMB: displaying graphs to the blind using an active auditory interface. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility (Assets '05)*. Association for Computing Machinery, New York, NY, USA, 182–183. <https://doi.org/10.1145/1090785.1090820> Citation Key: cohen-PLUMBDisplayingGraphs-2005.
- [34] Charlie Cross, Deniz Cetinkaya, and Huseyin Dogan. 2020. Transforming Diagrams' Semantics to Text for Visually Impaired. In *Design, User Experience, and Usability. Interaction Design (Lecture Notes in Computer Science)*, Aaron Marcus and Elizabeth Rosenzweig (Eds.). Springer International Publishing, Cham, 339–350. https://doi.org/10.1007/978-3-030-49713-2_24 Citation Key: cross-TransformingDiagramsSemantics-2020.
- [35] Alberto De Campo. 2007. Toward a data sonification design space map. In *Proceedings of the International Conference on Auditory Display*. 342–347.
- [36] Julie Ducasse, Anke M. Brock, and Christophe Jouffrais. 2018. *Accessible Interactive Maps for Visually Impaired Users*. Springer International Publishing, Cham, 537–584. https://doi.org/10.1007/978-3-319-54446-5_17
- [37] Julie Ducasse, Marc J-M Macé, Marcos Serrano, and Christophe Jouffrais. 2016. Tangible Reels: Construction and Exploration of Tangible Maps by Visually Impaired Users. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2186–2197. <https://doi.org/10.1145/2858036.2858058> Citation Key: ducasse-TangibleReelsConstruction-2016.
- [38] Emory J. Edwards, Michael Gilbert, Emily Blank, and Stacy M. Branham. 2023. How the Alt Text Gets Made: What Roles and Processes of Alt Text Creation Can Teach Us About Inclusive Imagery. *ACM Transactions on Accessible Computing* (Mar 2023), 3587469. <https://doi.org/10.1145/3587469>
- [39] Be My Eyes. 2023. Be My Eyes - See the world together. Retrieved 2023 from <https://www.bemyeyes.com/>
- [40] Danyang Fan, Kate Glazko, and Sean Follmer. 2022. *Accessibility of Linked-Node Diagrams on Collaborative Whiteboards for Screen Reader Users: Challenges and Opportunities*. Springer International Publishing, Cham, 97–108. https://doi.org/10.1007/978-3-031-09297-8_6
- [41] Danyang Fan, Alexa Fay Siu, Wing-Sum Adrienne Law, Raymond Ruihong Zhen, Sile O'Modhrain, and Sean Follmer. 2022. Slide-Tone and Tilt-Tone: 1-DOF Haptic Techniques for Conveying Shape Characteristics of Graphs to Blind Users. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–19. <https://doi.org/10.1145/3491102.3517790> Citation Key: fan-SlideToneTiltTone1DOF-2022b.
- [42] The American Foundation for the Blind. 2023. Screen Readers. Retrieved 2023 from <https://www.afb.org/blindness-and-low-vision/using-technology/assistive-technology-products/screen-readers>
- [43] 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*. ACM Press, Lisbon, Portugal, 97–106. <https://doi.org/10.1145/2700648.2809868>
- [44] 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*. ACM, Boulder Colorado USA, 103–110. <https://doi.org/10.1145/2384916.2384935>
- [45] Gagatay Goncu and Kim Marriott. 2011. *GravVITAS: Generic Multi-touch Presentation of Accessible Graphics*. Lecture Notes in Computer Science, Vol. 6946. Springer Berlin Heidelberg, Berlin, Heidelberg, 30–48. https://doi.org/10.1007/978-3-642-23774-4_5
- [46] Jenna L. Gorlewicz, Jennifer L. Tension, 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 (Sep 2020), 1–30. <https://doi.org/10.1145/3403933>
- [47] Brian Hansen, Leya Breanna Baltaxe-Admony, Sri Kurniawan, and Angus G. Forbes. 2019. Exploring Sonic Parameter Mapping for Network Data Structures. In *Proceedings of the 25th International Conference on Auditory Display (ICAD 2019)*. Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne, 67–74. <https://doi.org/10.21785/icad2019.055> Citation Key: hansen-ExploringSonicParameter-2019.
- [48] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (Oct 2006), 904–908. <https://doi.org/10.1177/154193120605000909>
- [49] Sandra G. Hart and Lowell E. Staveland. 1988. *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research*. Human Mental Workload, Vol. 52. North-Holland, 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [50] Liang He, Zijian Wan, Leah Findlater, and Jon E. Froehlich. 2017. TacTILE: A Preliminary Toolchain for Creating Accessible Graphics with 3D-Printed Overlays and Auditory Annotations. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Baltimore Maryland USA, 397–398. <https://doi.org/10.1145/3132525.3134818> Citation Key: he-TacTILEPreliminaryToolchain-2017.
- [51] Marti A Hearst. 2023. Show It or Tell It? Text, Visualization, and Their Combination. (2023).
- [52] Uta Hinrichs and Sheelagh Carpendale. 2011. Gestures in the Wild: Studying Multi-Touch Gesture Sequences on Interactive Tabletop Exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Vancouver, BC, Canada) (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 3023–3032. <https://doi.org/10.1145/1978942.1979391>
- [53] Eve Hoggan, John Williamson, Antti Oulasvirta, Miguel Nacenta, Per Ola Kristensson, and Anu Lehtio. 2013. Multi-Touch Rotation Gestures: Performance and Ergonomics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Paris, France) (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 3047–3050. <https://doi.org/10.1145/2470654.2481423>
- [54] Leona Holloway, Swamy Ananthanarayan, Matthew Butler, Madhuka Thisuri De Silva, Kirsten Ellis, Gagatay Goncu, Kate Stephens, and Kim Marriott. 2022. Animations at Your Fingertips: Using a Refreshable Tactile Display to Convey Motion Graphics for People who are Blind or have Low Vision. In *The 24th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Athens Greece, 1–16. <https://doi.org/10.1145/3517428.3544797>
- [55] Leona M Holloway, Gagatay Goncu, Alon Ilisar, Matthew Butler, and Kim Marriott. 2022. Infsonics: Accessible Infographics for People who are Blind using Sonification and Voice. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–13. <https://doi.org/10.1145/3491102.3517465> Citation Key: holloway-InfsonicsAccessibleInfographics-2022.
- [56] Anett Hoppe, David Morris, and Ralph Ewerth. 2021. Evaluation of Automated Image Descriptions for Visually Impaired Students. In *Artificial Intelligence in Education (Lecture Notes in Computer Science)*, Ido Roll, Danielle McNamara, Sergey Sosnovsky, Rose Luckin, and Vania Dimitrova (Eds.). Springer International Publishing, Cham, 196–201. https://doi.org/10.1007/978-3-030-78270-2_35 Citation Key: hoppe-EvaluationAutomatedImage-2021.
- [57] Mirko Horstmann*, Martin Lorenz, A Watkowski, G Ioannidis, Otthein Herzog, Alasdair King, D Gareth Evans, Cornelius Hagen, Christoph Schlieder, A-M Burn, et al. 2004. Automated interpretation and accessible presentation of technical diagrams for blind people. *New Review of Hypermedia and Multimedia* 10, 2 (2004), 141–163. <https://doi.org/10.1080/13614560512331326017>
- [58] W3C Web Accessibility Initiative (WAI). 2017. Tips and Tricks. Retrieved 2023 from <https://www.w3.org/WAI/tutorials/images/tips/>
- [59] Mihail Ivanchev, Francis Zinke, and Ulrike Lucke. 2014. *Pre-journey Visualization of Travel Routes for the Blind on Refreshable Interactive Tactile Displays*. Lecture Notes in Computer Science, Vol. 8548. Springer International Publishing, Cham, 81–88. https://doi.org/10.1007/978-3-319-08599-9_13
- [60] j1mmyto9. 2022. Speed to text in Unity iOS use Native Speech Recognition. Retrieved 2023 from <https://github.com/j1mmyto9/speech-and-text-unity-ios-android>

- [61] 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 (Jan 2022), 1095–1105. <https://doi.org/10.1109/TVCG.2021.3114846> Citation Key: jung-communicatingVisualizationsVisuals-2022.
- [62] Nikolaos Kaklanis, Konstantinos Votis, and Dimitrios Tzovaras. 2013. A mobile interactive maps application for a visually impaired audience. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility (W4A '13)*. Association for Computing Machinery, New York, NY, USA, 1–2. <https://doi.org/10.1145/2461121.2461152>
- [63] Shaun K. Kane, Jeffrey P. Bigham, and Jacob O. Wobbrock. 2008. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility - Assets '08*. ACM Press, Halifax, Nova Scotia, Canada, 73. <https://doi.org/10.1145/1414471.1414487>
- [64] Shaun K. Kane, Meredith Ringel Morris, Annuska Z. Perkins, Daniel Wigdor, Richard E. Ladner, and Jacob O. Wobbrock. 2011. Access overlays: improving non-visual access to large touch screens for blind users. In *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11)*. Association for Computing Machinery, New York, NY, USA, 273–282. <https://doi.org/10.1145/2047196.2047232>
- [65] Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable gestures for blind people: understanding preference and performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 413–422. <https://doi.org/10.1145/1978942.1979001>
- [66] Aniruddha Kembhavi, Mike Salvato, Eric Kolve, Minjoon Seo, Hannaneh Hajishirzi, and Ali Farhadi. 2016. A Diagram is Worth a Dozen Images. In *Computer Vision – ECCV 2016 (Lecture Notes in Computer Science)*, Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling (Eds.). Springer International Publishing, Cham, 235–251. https://doi.org/10.1007/978-3-319-46493-0_15 Citation Key: kembhavi-DiagramWorthDozen-2016.
- [67] Andrea R. Kennel. 1996. Audiograf: a diagram-reader for the blind. In *Proceedings of the second annual ACM conference on Assistive technologies (Assets '96)*. Association for Computing Machinery, New York, NY, USA, 51–56. <https://doi.org/10.1145/228347.228357>
- [68] Daesik Kim, YoungJoon Yoo, Jeeseo Kim, Sangkuk Lee, and Nojun Kwak. 2018. Dynamic Graph Generation Network: Generating Relational Knowledge from Diagrams. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Salt Lake City, UT, 4167–4175. <https://doi.org/10.1109/CVPR.2018.00438> Citation Key: kim-DynamicGraphGeneration-2018.
- [69] Gyeongri Kim, Jiho Kim, and Yea-Seul Kim. 2023. “Explain What a Treemap is”: Exploratory Investigation of Strategies for Explaining Unfamiliar Chart to Blind and Low Vision Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–13. <https://doi.org/10.1145/3544548.3581139>
- [70] Jiho Kim, Arjun Srinivasan, Nam Wook Kim, and Yea-Seul Kim. 2023. Exploring Chart Question Answering for Blind and Low Vision Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3544548.3581532>
- [71] Alasdair King, Paul Blenkhorn, David Crombie, Sijo Dijkstra, Gareth Evans, and John Wood. 2004. Presenting UML Software Engineering Diagrams to Blind People. In *Computers Helping People with Special Needs (Lecture Notes in Computer Science)*, Klaus Miesenberger, Joachim Klaus, Wolfgang L. Zagler, and Dominique Burger (Eds.). Springer, Berlin, Heidelberg, 522–529. https://doi.org/10.1007/978-3-540-27817-7_76
- [72] Stephen M. Kosslyn. 1989. Understanding charts and graphs. *Applied Cognitive Psychology* 3, 3 (1989), 185–225. <https://doi.org/10.1002/acp.2350030302>
- [73] Jill H. Larkin and Herbert A. Simon. 1987. Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science* 11, 1 (1987), 65–100. <https://doi.org/10.1111/j.1551-6708.1987.tb00863.x>
- [74] Jonathan Lazar, Daniel F Goldstein, and Anne Taylor. 2015. *Ensuring digital accessibility through process and policy*. Morgan kaufmann.
- [75] David Ledo, Miguel A. Nacenta, Nicolai Marquardt, Sebastian Boring, and Saul Greenberg. 2012. The HapticTouch toolkit: enabling exploration of haptic interactions. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (TEI '12)*. Association for Computing Machinery, New York, NY, USA, 115–122. <https://doi.org/10.1145/2148131.2148157>
- [76] Bongshin Lee, Catherine Plaisant, Cynthia Sims Parr, Jean-Daniel Fekete, and Nathalie Henry. 2006. Task taxonomy for graph visualization. In *Proceedings of the 2006 AVI workshop on Beyond time and errors: novel evaluation methods for information visualization (BELIV '06)*. Association for Computing Machinery, New York, NY, USA, 1–5. <https://doi.org/10.1145/1168149.1168168>
- [77] Nicolai Marquardt, Miguel A. Nacenta, James E. Young, Sheelagh Carpendale, Saul Greenberg, and Ehud Sharlin. 2009. The Haptic Tabletop Puck: tactile feedback for interactive tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '09)*. Association for Computing Machinery, New York, NY, USA, 85–92. <https://doi.org/10.1145/1731903.1731922>
- [78] Richard E. Mayer. 2002. *Multimedia learning*. Psychology of Learning and Motivation, Vol. 41. Academic Press, 85–139. [https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6)
- [79] Michał Maćkowski, Piotr Brzoza, Mateusz Kawulok, Rafał Meisel, and Dominik Spinczyk. 2023. Multimodal Presentation of Interactive Audio-Tactile Graphics Supporting the Perception of Visual Information by Blind People. *ACM Transactions on Multimedia Computing, Communications, and Applications* (Mar 2023), 3586076. <https://doi.org/10.1145/3586076> Citation Key: mackowski-MultimodalPresentationInteractive-2023.
- [80] David Mcgookin, Stephen Brewster, and WeiWei Jiang. 2008. Investigating Touchscreen Accessibility for People with Visual Impairments, Vol. 358. <https://doi.org/10.1145/1463160.1463193>
- [81] Daniel R. Montello. 2015. *Spatial Cognition*. Elsevier, 111–115. <https://doi.org/10.1016/B978-0-08-097086-8.72057-5>
- [82] Tamara Munzner. 2014. *Visualization Analysis and Design* (0 ed.). A K Peters/CRC Press. <https://doi.org/10.1201/b17511>
- [83] Vishnu Nair, Hanxiu “Hazel” Zhu, and Brian A. Smith. 2023. ImageAssist: Tools for Enhancing Touchscreen-Based Image Exploration Systems for Blind and Low Vision Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17. <https://doi.org/10.1145/3544548.3581302> arXiv:2302.09124 [cs].
- [84] NASA and the Harvard-Smithsonian Center for Astrophysics. 2020. Data sonification: Sounds from around the milky way. <https://chandra.si.edu/photo/2020/sonify/>
- [85] OpenAI. 2023. GPT-4 Technical Report. arXiv:2303.08774 (Mar 2023). <https://doi.org/10.48550/arXiv.2303.08774> arXiv:2303.08774 [cs].
- [86] Don Parkes. 1988. NOMAD”: An audio-tactile tool for the acquisition, use and management of spatially distributed information by partially sighted and blind persons. In *Proceedings of Second International Conference on Maps and Graphics for Visually Disabled People*. 24–29.
- [87] Helen Petrie, Neil King, Anne-Marie Burn, and Peter Pavan. 2006. Providing interactive access to architectural floorplans for blind people. *British Journal of Visual Impairment* 24, 1 (Jan 2006), 4–11. <https://doi.org/10.1177/0264619606060010>
- [88] Helen Petrie, Christoph Schlieder, Paul Blenkhorn, Gareth Evans, Alasdair King, Anne-Marie O’Neill, George T. Ioannidis, Blaithin Gallagher, David Crombie, Rolf Mager, and Maurizio Alafaci. 2002. TeDUB: A System for Presenting and Exploring Technical Drawings for Blind People. In *Computers Helping People with Special Needs (Lecture Notes in Computer Science)*, Klaus Miesenberger, Joachim Klaus, and Wolfgang Zagler (Eds.). Springer, Berlin, Heidelberg, 537–539. https://doi.org/10.1007/3-540-45491-8_102
- [89] Benjamin Poppinga, Charlotte Magnusson, Martin Pielot, and Kirsten Rasmussen-Gröhn. 2011. TouchOver map: audio-tactile exploration of interactive maps. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. ACM, Stockholm Sweden, 545–550. <https://doi.org/10.1145/2037373.2037458>
- [90] Gaspar Ramoá, Vincent Schmidt, and Peter König. 2022. Developing Dynamic Audio Navigation UIs to Pinpoint Elements in Tactile Graphics. *Multimodal Technologies and Interaction* 6, 12 (Dec 2022), 113. <https://doi.org/10.3390/mti6120113> Citation Key: ramoa-DevelopingDynamicAudio-2022.
- [91] Ravi Rastogi and Dianne T. V. Pawluk. 2013. Dynamic Tactile Diagram Simplification on Refreshable Displays. *Assistive Technology* 25, 1 (Mar 2013), 31–38. <https://doi.org/10.1080/10400435.2012.685567>
- [92] Mike Scaife and Yvonne Rogers. 1996. External cognition: how do graphical representations work? *International journal of human-computer studies* 45, 2 (1996), 185–213.
- [93] Bernhard Schmitz and Thomas Ertl. 2012. Interactively Displaying Maps on a Tactile Graphics Display. (2012).
- [94] Ather Sharif, Olivia H. Wang, Alida T. Muongchan, Katharina Reinecke, and Jacob O. Wobbrock. 2022. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–19. <https://doi.org/10.1145/3491102.3517431>
- [95] Shigenobu Shimada, Haruka Murase, Suguru Yamamoto, Yusuke Uchida, Makoto Shimojo, and Yutaka Shimizu. 2010. Development of Directly Manipulable Tactile Graphic System with Audio Support Function. In *Computers Helping People with Special Needs (Lecture Notes in Computer Science)*, Klaus Miesenberger, Joachim Klaus, Wolfgang Zagler, and Arthur Karshmer (Eds.). Springer, Berlin, Heidelberg, 451–458. https://doi.org/10.1007/978-3-642-14100-3_68
- [96] Mathieu Simonnet, Cécile Bothorel, Luiz Felipe Maximiano, and André Thépaut. 2012. GeoTablet, une application cartographique pour les personnes déficientes visuelles.
- [97] Robert Spence. 2014. *Information Visualization: An Introduction*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-07341-5>
- [98] Jing Su, Alyssa Rosenzweig, Ashvin Goel, Eyal De Lara, and Khai N. Truong. 2010. Timbremap: enabling the visually-impaired to use maps on touch-enabled devices. In *Proceedings of the 12th international conference on Human computer*

interaction with mobile devices and services. ACM, Lisbon Portugal, 17–26. <https://doi.org/10.1145/1851600.1851606>

- [99] Saiganesh Swaminathan, Thijs Roumen, Robert Kovacs, David Stangl, Stefanie Mueller, and Patrick Baudisch. 2016. Linespace: A Sensemaking Platform for the Blind. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 2175–2185. <https://doi.org/10.1145/2858036.2858245>
- [100] Unity Technologies. 2023. Maximize Multiplatform Game Development | Unity. Retrieved 2023 from <https://unity.com/solutions/multiplatform>
- [101] Jennifer L. Tennison and Jenna L. Gorlewicz. 2019. Non-visual Perception of Lines on a Multimodal Touchscreen Tablet. *ACM Transactions on Applied Perception* 16, 1 (Jan 2019), 1–19. <https://doi.org/10.1145/3301415>
- [102] Lore Thaler and Melvyn A Goodale. 2016. Echolocation in humans: an overview. *Wiley Interdisciplinary Reviews: Cognitive Science* 7, 6 (2016), 382–393.
- [103] John R Thompson, Jesse J Martinez, Alper Sarikaya, Edward Cutrell, and Bongshin Lee. 2023. Chart Reader: Accessible Visualization Experiences Designed with Screen Reader Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–18. <https://doi.org/10.1145/3544548.3581186>
- [104] Márcio Josué Ramos Torres and Regina Barwaldt. 2019. Approaches for diagrams accessibility for blind people: a systematic review. In *2019 IEEE Frontiers in Education Conference (FIE)*. 1–7. <https://doi.org/10.1109/FIE43999.2019.9028522> Citation Key: torres-ApproachesDiagramsAccessibility-2019.
- [105] Unity Technologies. 2023. Unity Real-Time Development Platform. Retrieved 2023 from <https://unity.com/>
- [106] Shin'ichiro Uno, Yasuo Suzuki, Takashi Watanabe, Miku Matsumoto, and Yan Wang. 2018. Soundbased Image and Position Recognition System SIPReS. In *Proceedings of the 24th International Conference on Auditory Display - ICAD 2018*. The International Community for Auditory Display, Houghton, Michigan, 67–72. <https://doi.org/10.21785/icad2018.005> Citation Key: uno-SoundbasedImagePosition-2018.
- [107] Fernando Vidal-Verdu and Moustapha Hafez. 2007. Graphical Tactile Displays for Visually-Impaired People. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 15, 1 (Mar 2007), 119–130. <https://doi.org/10.1109/TNSRE.2007.891375>
- [108] Ge Wang and Perry R Cook. 2003. ChucK: A Concurrent, On-the-fly, Audio Programming Language. In *Proceedings of the 2003 International Computer Music Conference*. 8.
- [109] Colin Ware. 2013. *Information Visualization: Perception for Design*. Elsevier.
- [110] Robin Williams. 2015. *The non-designer's design book: Design and typographic principles for the visual novice*. Pearson Education.
- [111] Ikuko Eguchi Yairi, Masamitsu Takano, Motoki Shino, and Minoru Kamata. 2008. Expression of paths and buildings for universal designed interactive map with due consideration for visually impaired people. In *2008 IEEE International Conference on Systems, Man and Cybernetics*. 524–529. <https://doi.org/10.1109/ICSMC.2008.4811330>
- [112] Naohide Yamamoto and Amy L. Shelton. 2005. Visual and proprioceptive representations in spatial memory. *Memory & Cognition* 33, 1 (Jan 2005), 140–150. <https://doi.org/10.3758/BF03195304>
- [113] Yalong Yang, Kim Marriott, Matthew Butler, Gagatay Goncu, and Leona Holloway. 2020. Tactile Presentation of Network Data: Text, Matrix or Diagram?. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3313831.3376367> Citation Key: yang-TactilePresentationNetwork-2020.
- [114] Limin Zeng and Gerhard Weber. 2010. Audio-Haptic Browser for a Geographical Information System. In *Computers Helping People with Special Needs (Lecture Notes in Computer Science)*, Klaus Miesenberger, Joachim Klaus, Wolfgang Zagler, and Arthur Karshmer (Eds.). Springer, Berlin, Heidelberg, 466–473. https://doi.org/10.1007/978-3-642-14100-3_70
- [115] Limin Zeng and Gerhard Weber. 2012. ATMap: Annotated Tactile Maps for the Visually Impaired. In *Cognitive Behavioural Systems (Lecture Notes in Computer Science)*, Anna Esposito, Antonietta M. Esposito, Alessandro Vinciarelli, Rüdiger Hoffmann, and Vincent C. Müller (Eds.). Springer, Berlin, Heidelberg, 290–298. https://doi.org/10.1007/978-3-642-34584-5_25
- [116] Zhuohao Zhang and Jacob O. Wobbrock. 2022. A11yBoard: Using Multimodal Input and Output to Make Digital Artboards Accessible to Blind Users. In *Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 1–4. <https://doi.org/10.1145/3526114.3558695>
- [117] Zhuohao (Jerry) Zhang and Jacob O. Wobbrock. 2023. A11yBoard: Making Digital Artboards Accessible to Blind and Low-Vision Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–17. <https://doi.org/10.1145/3544548.3580655>
- [118] Jonathan Zong, Crystal Lee, Alan Lundgard, JiWoong Jang, Daniel Hajas, and Arvind Satyanarayan. 2022. Rich Screen Reader Experiences for Accessible Data Visualization. *Computer Graphics Forum* 41, 3 (2022), 15–27. <https://doi.org/10.1111/cgf.14519>

A FORMATIVE STUDY: PARTICIPANT DETAILS

Table 2 presents a summary of participant details in the formative study.

B FORMATIVE STUDY: INTERVIEW QUESTIONS

Here is the list of interview questions asked for each scenario of a diagram encounter:

- I. Describe briefly the diagrammatic information from your encounter.
- II. Describe briefly the context of the encounter in terms of when, where, how, who, and why.
- III. To which degree were you successful in accessing the diagrammatic information? (from 0, not successful at all, to 5, perfectly successful)
- IV. How often do you encounter such a scenario?
- V. What were the motives for or objectives of accessing the information?
- VI. What was the importance of the information?
- VII. Was it an individual, collaborative, or hybrid setting?
- VIII. What was the surrounding physical environment? For example, was it public or private? Was it personal or professional?
- IX. What media was the information presented in? Was it printed, electronic, verbal, or other?
- X. What did you do when you accessed the information? What tools did you use to access the information, if any?
- XI. What types of tasks do you do with the diagrammatic information? (Some examples include getting a general idea, searching, exploring details, browsing, navigating, and annotating.)
- XII. What went well and smoothly when trying to access the diagram?
- XIII. What were the main challenges that you encountered when trying to access the diagram?
- XIV. How did you try to address these challenges, if at all?
- XV. Do you have any ideas that would help you access diagrammatic information in better ways?

C TADA'S IMPLEMENTATION DETAILS

TADA is built using Unity [105], ChucK [108], and Chunity [8]. Unity, a game engine, uses C# scripts to interpret diagram data, construct the user interface and interactive diagram components, identify various gestures and their interactions with diagrams, and initiate the corresponding audio feedback. Unity's multi-platform support allows the system to operate on tablets with different operating systems [100]. ChucK, a programming language that generates concurrent, strongly-timed procedural audio, renders TADA's audio. Chunity, which merges ChucK and Unity, serves as the bridge between the two. Native speech recognition is enabled by open-source codes [60], allowing speech engines to run on either iOS / iPadOS or Android.

Table 2: Details about participants from the formative study.

PID	Age	Gender	Degree of Vision Loss	Onset	Background
P1	35-40	Female	Total	Birth	Employee resource group lead
P2	38	Male	Total	Young adult	Government employee; Salesperson; Worked in the sports and recreation industry
P3	35-40	Female	Total	Birth	Retail
P4	18	Male	Total	Birth	Student
P5	52	Male	Blurry, glasses do not help	Recent years (3 years ago)	Hardware support, IT
P6	32	Female	Total	Birth	Government employee
P7	60	Male	3% vision	3 years old	Teacher
P8	60	Male	Total	7 years old	Businessman
P9	40	Female	Congenital Rubella Syndrome and Glaucoma. Fluctuating vision	Lifelong progression	Human resources, volunteer resources
P10	63	Male	1% vision	Birth	Banking and assistive technology
P11	69	Male	Total	Early seventies	Self-employed
P12	56	Female	Legally blind, low vision	Losing vision for the past decade	Professor
P13	62	Female	Total	2 years old	Information officer, now retired
P14	42	Female	Legally blind, significant vision loss	Birth, lifelong progression	Research and teaching
P15	50	Male	Total	Vision loss from birth, total vision loss from 2 years ago	Wood products engineer

D EVALUATIVE STUDY: PARTICIPANT DETAILS

Table 3 shows a summary of participant details from the evaluative study.

E EVALUATIVE STUDY: EXPERIMENTAL BLOCKS

This appendix presents the details of the experimental blocks from the evaluative study.

The 6 experimental blocks each focused on an interaction technique. The interaction techniques evaluated are (1) Single-finger Sweep; (2) Five-finger Dome; (3) Dwell+Tap; (4) Single-finger Flick-down, Speech Command for Searching and Single-finger Follow; (5) Dwell+Circle; and (6) Dwell+Radiate. We tasked the participants to only use the dedicated interaction technique to answer its associated question(s) and the questions asked about the test diagrams are listed below corresponding to the blocks.

E.1 Gain an Overview with Single-finger Sweep

We used test diagram 1 (Figure 2b) which was a more complex diagram than the tactile diagram, and it has distinct characteristics in the diagrammatic structure in its four quadrants meant for comparison. We tasked participants to gain an overview with Single-finger Sweep and answer the following questions about the test diagram:

- Q1.1: Describe the overall layout of the diagram.
- Q1.2: Which quadrant of the diagram has the most connections between people? (Answer: Bottom left.)
- Q1.3: Which quadrant has the least connections? (Answer: Top left.)

Q1.4: Which quadrant of the diagram has the most people? (Answer: Bottom right.)

Q1.5: Which quadrant has the least people? (Answer: Top left.)

For the first question (Q1.1), we asked participants to sweep and explore the entire diagram and provide answers regarding the general layout. We then asked more specific questions (Q1.1 to 1.4) for participants to explicitly think in terms of quadrants and in terms of the connectivity and density of nodes. We asked participants to not count exactly but to gather a general overview.

E.2 Gain a Summary of a Sub-region with Five-finger Dome

To learn about the effectiveness of the Five-finger Dome within sub-regions, we tasked participants to gather the summaries of sub-regions using Five-finger Dome and answer the following questions:

- Q2.1: Which quadrant of the diagram has the most connections between people? (Answer: Bottom right.)
- Q2.2: Which quadrant has the least connections? (Answer: Bottom left.)
- Q2.3: Which quadrant of the diagram has the most people? (Answer: Top right.)
- Q2.4: Which quadrant has the least people? (Answer: Bottom left.)
- Q2.5: Which quadrant has people with unequal numbers of friends? (Answer: Top right.)

We asked the same specific questions about connectivity and density of nodes but switched the diagram to test diagram 2 (Figure 2c) to make sure participants cannot reuse the same answers

Table 3: Details about participants from the evaluation study.

PID	Age	Gender	Degree of Vision Loss	Onset	Able to Perceive Diagrams Visually?	Background
P1	52	Female	Legally blind, less than 3% peripheral vision left	10 years old	No	Massage therapist
P2	48	Male	Light perception in left eye, blurred vision in right eye	2022	No	Construction business
P3	25	Male	Total	Birth	No	Technology-related
P4	56	Female	Total	21 years old	No	Senior management
P5	46	Trans-gender female	Total	Birth	No	Volunteer for non-profit organizations and local communities
P6	53	Female	Total	Birth	No	Language teacher
P7	70	Female	Total	5 years old	No	Program coordinator; physiotherapist
P8	101	Female	Legally blind, usable vision remaining in left eye.	Gradually losing	Yes	Therapeutic dietitian
P9	73	Male	Total	11 years old	No	IT technologist
P10	52	Female	Legally blind, RP, no central vision, some peripheral vision left	18 years old	No	Manager
P11	65	Male	Total	Birth, light and colour perception until age seven	No	Worker at a disability organization
P12	58	Female	Legally blind, partial optic atrophy with peripheral vision	Prematurely	Yes (able to visually perceive some)	Unemployed
P13	35	Female	Legally blind	Birth	Yes (able to visually perceive some)	Program coordinator
P14	41	Male	Light perception in left eye, total blind in right eye	Birth	No	Assistive tech support
P15	56	Female	Legally blind, 3% vision in right eye.	Birth, gradual vision loss	Yes (able to visually perceive some)	Previously professional musician; After vision loss, teaching music and assistive technology
P16	46	Female	Legally blind, totally blind in right eye, some narrow vision in left eye	Birth	No	Customer service and sales
P17	32	Female	Total	Birth	No	Government work
P18	50	Female	Legally blind	2013	No	Medical assistant
P19	58	Male	RP, legally blind	2008	Yes	Accounting
P20	37	Female	Legally blind	9 years old	No	Non-profit sector
P21	68	Male	Legally blind, some peripheral vision left	2014	No	Previously transit driver
P22	61	Male	Legally blind, some peripheral vision left	30 years old	No	Software engineering
P23	38	Male	Total	Young adult	No	Government employee; Salesperson; Worked in the sports and recreation industry
P24	28	Male	Legally blind	2016	No	Assistive tech support
P25	49	Female	Legally blind, can only see light and shadow	Birth, gradual vision loss	No	Executive director at non-profit organization

from test diagram 1. It included modifications to test diagram 1 such as changing the positions and numbers of the nodes in each quadrant. Additionally, because we also want to evaluate explicitly the effectiveness of the pacing of audio, Q2.5 was asked to see if participants could pick up the variations in the pacing of the link audio sounds within an area. There can be multiple answers to Q2.5, and we asked participants to identify the most obvious one where there is one person very popular in the top right quadrant while the others are not so that this node's link sounds are paced much faster. Again, we asked participants to not count exactly but to gather a general overview, and we did not evaluate what details this interaction could potentially achieve.

E.3 Gain Textual Details with Dwell+Tap

We asked a question about the details of a diagram element from performing the Dwell+Tap interaction:

Q3: Given a person, what is the name and professional background associated with it?

We used test diagram 3 (Figure 2d) for this task and the remaining three tasks (Task 4 to Task 6). We do not have to design different diagrams for the remaining interactions because the answers would not affect answering later questions. In Q3, we provided participants with a node to examine that had a name other than "Bob" which was going to be used in later questions.

E.4 Search with Single-finger Flick-down, Speech Command and Single-finger Follow

We tasked the participants to locate a piece of information and answer the following question:

Q4: Where is Bob? Locate Bob.

To eliminate other factors affecting the accuracy of speech recognition such as participants' accents and background noise, we conducted a Wizard of Oz experiment where we asked participants to only look for where Bob is in the diagram, and the system is only configured to search for Bob.

E.5 Explore Connected Links with Dwell+Circle

We asked participants to perform Dwell+Circle to identify the number of connected links of a node:

Q5: How many friends does Bob have? (Answer: 4.)

Dwell+Circle aims to explore a node's connected links in detail, gathering the number of connections, their directions and their relative distances. Q5 only asked about the number of connections because, in the interest of time, we decided to focus on whether one could complete a full circle and count the number of connected links without exploring other details. Q6 below partially explored the directional information of the connected links.

E.6 Navigate with Dwell+Radiate

Following Task 5, we asked participants to navigate to any nodes of the neighbouring node using Dwell+Radiate:

Q6: Navigate from Bob to any of his friends.

Dwell+Radiate works in combination with Dwell+Circle. Using the Dwell+Circle interaction to explore the connected links, one

can locate a link of interest and then decide to navigate toward its direction using the Dwell+Radiate interaction. Therefore, Q6 evaluated the Dwell+Circle interaction partially and the Radiate interaction.