



## OPEN ACCESS

## EDITED BY

Wang-Kin Chiu,  
The Hong Kong Polytechnic University,  
China

## REVIEWED BY

Sarah Pila,  
Northwestern University,  
United States  
Anna Shvarts,  
Utrecht University,  
Netherlands

## \*CORRESPONDENCE

Stacy A. Doore  
✉ [sadoore@colby.edu](mailto:sadoore@colby.edu)

<sup>1</sup>These authors have contributed equally to this work and share first authorship

## SPECIALTY SECTION

This article was submitted to  
STEM Education,  
a section of the journal  
Frontiers in Education

RECEIVED 16 October 2022

ACCEPTED 08 February 2023

PUBLISHED 11 April 2023

## CITATION

Doore SA, Dimmel J, Kaplan TM,  
Guenther BA and Giudice NA (2023)  
Multimodality as universality: Designing  
inclusive accessibility to graphical information.  
*Front. Educ.* 8:1071759.  
doi: 10.3389/feduc.2023.1071759

## COPYRIGHT

© 2023 Doore, Dimmel, Kaplan, Guenther and Giudice. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Multimodality as universality: Designing inclusive accessibility to graphical information

Stacy A. Doore<sup>1\*†</sup>, Justin Dimmel<sup>2†</sup>, Toni M. Kaplan<sup>3</sup>,  
Benjamin A. Guenther<sup>3,4</sup> and Nicholas A. Giudice<sup>3,5†</sup>

<sup>1</sup>Immersive Navigation Systems and Inclusive Tech Ethics (INSITE) Laboratory, Department of Computer Science, Colby College, Waterville, ME, United States, <sup>2</sup>Research in STEM Education (RiSE) Center, School of Learning and Teaching, University of Maine, Orono, ME, United States, <sup>3</sup>Virtual Environment and Multimodal Interaction (VEMI) Laboratory, University of Maine, Orono, ME, United States,

<sup>4</sup>Department of Psychology, College of Liberal Arts and Sciences, University of Maine, Orono, ME, United States, <sup>5</sup>Spatial Computing Program, School of Computing and Information Science, University of Maine, Orono, ME, United States

Graphical representations are ubiquitous in the learning and teaching of science, technology, engineering, and mathematics (STEM). However, these materials are often not accessible to the over 547,000 students in the United States with blindness and significant visual impairment, creating barriers to pursuing STEM educational and career pathways. Furthermore, even when such materials are made available to visually impaired students, access is likely through literalized modes (e.g., braille, verbal description), which is problematic as these approaches (1) do not directly convey spatial information and (2) are different from the graphic-based materials used by students without visual impairment. The purpose of this study was to design and evaluate a universally accessible system for communicating graphical representations in STEM classes. By combining a multisensory vibro-audio interface and an app running on consumer mobile hardware, the system is meant to work equally well for all students, irrespective of their visual status. We report the design of the experimental system and the results of an experiment where we compared learning performance with the system to traditional (visual or tactile) diagrams for sighted participants ( $n=20$ ) and visually impaired participants ( $n= 9$ ) respectively. While the experimental multimodal diagrammatic system (MDS) did result in significant learning gains for both groups of participants, the results also revealed no statistically significant differences in the capacity for learning from graphical information across both comparison groups. Likewise, there were no statistically significant differences in the capacity for learning from graphical information between the stimuli presented through the experimental system and the traditional (visual or tactile) diagram control conditions, across either participant group. These findings suggest that both groups were able to learn graphical information from the experimental system as well as traditional diagram presentation materials. This learning modality was supported without the need for conversion of the diagrams to make them accessible for participants who required tactile materials. The system also provided additional multisensory information for sighted participants to interpret and answer questions about the diagrams. Findings are interpreted in terms of new universal design principles for producing multisensory graphical representations that would be accessible to all learners.

## KEYWORDS

blind and low vision, multimodal interface, accessible STEM graphics, multisensory interactions, learning system, universal design, assistive technologies

## 1. Introduction

Traditional learning materials used in mainstream science, engineering, technology, and mathematics (STEM) classrooms rely heavily on graphics and images to efficiently convey complex concepts (Kress and Van Leeuwen, 2006). However, these materials are often inaccessible to students with blindness or significant visual impairment (BVI), and this inaccessibility creates significant barriers to STEM educational and career pathways. Current statistics on the number of school-age children that meet the federal definition of visual impairment<sup>1</sup> (including blindness) are often difficult to obtain due to the ways in which incidence data is defined by different states (National Center for Education Statistics (NCES), 2022), and some research suggests that the federal child count underestimates the incidence of visual impairment (Schles, 2021). According to the 2021 American Community Survey (ACS), there are 7.5 million (2.5%) Americans, who are blind or have low vision, including approximately 547,000 children with severe vision difficulty under the age of 18<sup>2</sup> (U.S. Census Bureau, 2020). Of those students, there were approximately 55,249 United States children, youth, and adult students in educational settings who were classified as legally blind<sup>3</sup> (American Printing House for the Blind (APH), 2019). School success opportunities and outcomes can have lifelong impact on BVI individuals. Of the nearly 4 million civilian non-institutionalized working age adults (18–64) with a visual impairment, only 2 million (50%) are employed, another 250,000 (5%) working adults with a visual impairment are classified as unemployed (but still looking for work), with the remaining 1.8 million (45%) of adults with visual impairments classified as not actively engaged in the labor force (McDonnell and Sui, 2019; U.S. Census Bureau, 2020). This compares to 136 million (77%) civilian non-institutionalized working age adults (18–64) without a disability who are employed, 8 million (5%) unemployed (but still looking for work), with the remaining 33 million (18%) of the adult population without a disability classified as not actively engaged in the labor force (U.S. Census Bureau, 2020).

This large disparity between the employment rates for BVI adults and the general population without a disability (50% vs. 77%) helps to motivate our work to improve information access for advancing into STEM related careers and the opportunities that are available with

advanced STEM education. As workplaces become more automated, future labor market skills needed to maintain United States progress and innovation will require more diversity of perspectives for complex problem solving, therefore “all learners must have an equitable opportunity to acquire foundational STEM knowledge” (Honey et al., 2020). In order to support these equitable opportunities in STEM, there is a profound need for accessible STEM training tools and learning materials to provide learning access across future labor contexts and for people of all ages. NSF STEM participation data does not breakdown participation by disability type (Blaser and Ladner, 2020). What is well documented is that the dearth of accessible materials for BVI learners at all levels and how this presents acute challenges for inclusive STEM courses (Moon et al., 2012). To take one example, the study of geometry, as manifested in secondary schools, is inextricably bound with what has been described as the *diagrammatic register* – a communication modality in which mathematical concepts are conveyed through logical statements (written in words) that are linked to diagrams (Dimmel and Herbst, 2015). The primary challenge for BVI learners is that geometry diagrams visually convey properties that do not explicitly describe spatial information in the accompanying text, such as whether a point is on a line, or whether two lines intersect. Thus, verbal descriptions containing additional information are necessary to make the diagrams accessible to BVI learners, however, these longer descriptions use additional words that can increase the cognitive load of making sense of the representation, with the long descriptions often still failing to convey key spatial content (Doore et al., 2021). As a result, BVI learners spend significantly more time and are far less accurate than their sighted peers in interpreting diagrammatic representations due to the lack of consistent standards for graphical content metadata, including description annotations (Sharif et al., 2021; Zhang et al., 2022). While extended length description recommendations for graphical representations have evolved and improved over time (Hasty et al., 2011; W3C, 2019), few guidelines for natural language descriptions of diagrams, charts, graphs, and maps are grounded in any theoretical framework with some notable exceptions using category theory (Vickers et al., 2012) spatial cognition theory (Trickett and Trafton, 2006), semiotic theory (Chandler, 2007), and linguistic theory (Lundgard and Satyanarayan, 2021). We view our work as complementary to this body of theoretically grounded research, embedding structured natural language descriptions into accessible multisensory data representations that use haptics, spatial audio, and high contrast visuals to help with the interpretation of graphic information.

Beyond the challenges of creating accessible information ecosystems in classrooms for all learners, the STEM visualization access challenge has received growing attention at a broader societal level as the use of graphical representations has been shown to play an important role in conveying abstract concepts and facilitating the deeper meaning of scientific texts (Khine, 2013). The information access gap inevitably contributes to the lower academic performance observed in math and science among BVI students in comparison both to other subjects and also to their non-visually impaired peers within STEM disciplines (Cryer et al., 2013). The limited availability of blind-accessible materials can also force teachers to adopt content that employs phrasing, structure, or terminology that does not correspond with the teacher’s preferred method of instruction or intended curriculum. This lack of access to educational materials can

1 Federal regulations define visual impairment (including blindness) as “an impairment in vision that, even with correction, adversely affects a child’s educational performance” [34CFR Sec. 300.8(c)(13)]. Some states have elaborated on this definition by specifying minimum levels of visual acuity or a restriction in the visual field. Thus, a child may qualify as having a visual impairment in one state but may not qualify in another <https://ies.ed.gov/ncser/pubs/20083007/index.asp>.

2 The children referred to range in age from 0 to 17 years and only included those children that had serious difficulty seeing even when wearing glasses as well as those that are blind.

3 The students referred to range in age from 0 to 21 years as well as certain qualifying adult students and only included those students with vision loss that functioned at/met the legal definition of blindness. Legal blindness is a level of vision loss that has been defined by law to determine eligibility for benefits. It refers to explicitly to those who have a central visual acuity of 20/200 or less in the better eye with the best possible correction, or a visual field of 20 degrees or less.

make classroom learning and information interpretation difficult, resulting in BVI students falling significantly behind standard grade level content (Lundgard and Satyanarayan, 2021).

There is thus an urgent need for a universal accessibility solution providing inclusive information access to STEM content supporting the same level of learning, understanding, and representation—i.e., functionally equivalent performance—for all learners. By universal, we mean the solution should use only those accessibility supports that could reasonably be expected to be familiar and available to all learners—i.e., the solution would not require specialized hardware or knowledge of specialized systems of communication, such as braille (National Federation of the Blind (NFB), 2009). By *functionally equivalent*, we mean the representations built up from different modalities will be associated with similar behavioral performance on STEM tasks (e.g., accuracy and success rate; Giudice et al., 2011). Evidence for such functional equivalence has been observed across many tasks and is explained by the development of a sensory-independent, 'spatial' representation in the brain, called the spatial image, which supports similar (i.e., statistically equivalent) behavior, independent of the learning modality (for reviews, see Loomis et al., 2013; Giudice, 2018). Functional equivalence has been demonstrated with learning from many combinations of inputs (visual, haptic, spatialized audio, spatial language), showing highly similar behavioral performance across a range of inputs and spatial abilities including spatial updating (Avraamides et al., 2004), target localization (Klatzky et al., 2003), map learning (Giudice et al., 2011) and forming spatial images in working-memory (Giudice et al., 2013).

The purpose of the current study was to investigate how effectively working age adults could learn graphical-based STEM content information from a universally designed interface that was developed to support functional equivalence across visual and non-visual modalities for representing diagrams. We asked: How effectively do multisensory inputs (high contrast visuals, spatial language, and haptics) convey functionally equivalent spatial information for learning concepts that are represented in diagrams? We investigated this question by developing and testing a multisensory diagram system that was designed to be accessible to all learners.

## 2. Background

### 2.1. Universal design for assistive technologies

Our focus on all learners was motivated by two considerations. One, BVI learners face significant social challenges in school, where impromptu group discussion and peer-to-peer learning are important components of social and behavioral skill development (Smith et al., 2009). Inclusive classrooms are increasingly the most common educational settings among BVI students, with over 80% of this demographic attending local public schools and spending most of their time in inclusive classrooms alongside of their sighted peers (Heward, 2003; American Printing House for the Blind (APH), 2019). Two, a universal design approach is thus advantageous because it reduces barriers for BVI learners to participate in peer-mediated classroom activities (e.g., group work) – when everyone is using the same resources, there is no reason for the BVI learners to receive special accommodation. This is also an important consideration as

whole class discussions that occur naturally in inclusive settings play a crucial role in the development of social, linguistic, and behavioral skills, as well as improve conceptual understanding and overall academic performance (Smith et al., 2009; Voltz et al., 2016).

### 2.2. Multisensory learning

Apprehending information through multiple sensory modalities is beneficial for everyone, not only those for whom a sensory accommodation was initially designed (e.g., Yelland, 2018; Abrahamson et al., 2019). In the 20 years since Mayer's seminal paper "Multimedia Learning" (Mayer, 2002), hundreds of studies have investigated how complementary sensory modalities, such as pictures and text, can enhance the acquisition and retention of information. How closed captioning has been adopted and integrated into educational, professional, and recreational videos is one example. Closed-captioning benefits deaf and non-hard of hearing viewers alike (Kent et al., 2018; Tipton, 2021). The availability of closed captioning across media reflects not only a commitment to accessibility but also provides empirical examples, at scale, that illustrate the redundancy principle.

The redundancy principle hypothesizes that simultaneous presentations of the same information *via* different modes allows modality-independent sensory processing to occur simultaneously: Two cognitive systems can process the same information in parallel (Moreno and Mayer, 2002). Reading closed captions taxes visual working memory, while hearing spoken words taxes auditory working memory. These processes are independent, which means reading captions while simultaneously listening to spoken words allows for both the visual and auditory systems to work synergistically toward apprehending the information that is represented in written (visual) and spoken (auditory) words. Redundancy, when partitioned across independent sensory modalities, helps learners build and retain conceptual (i.e., mental) models (Moreno and Mayer, 2002), which are integrated representations of spatial information about objects and relations.

### 2.3. Spatial mental models

Model theory asserts that people translate a perceived spatial configuration into a mental model and then use this mental representation to problem-solve and make inferences on spatial information (Johnson-Laird, 1983, 2010; Johnson-Laird and Byrne, 1991). Under the best of conditions, spatial reasoning problems are difficult to solve using language alone (Ragni and Knauff, 2013). For example, describing something as simple as how to locate the reception desk within a hotel lobby is both a complex description task (for the person doing the describing) and a difficult non-visual navigation task (for the BVI person who needs directions to navigate) because there are no tools (e.g., a standardized coordinate system) for providing spatial references within the lobby (or other similar indoor environments). As such, the non-visual navigation task for solving what we call the "lobby problem," i.e., independently finding the check-in desk from a hotel's main entrance, or the elevator from the check-in desk, or the hotel restaurant down a long hallway from the lobby can be extremely challenging. Instructional graphics present a

similar challenge where the typical accessibility solution is a poorly structured (and all too often ambiguous) description from a teacher or instructional aide. The adage that ‘a picture is worth a thousand words’ is most certainly true in that humans can process complex visual information in an image to understand spatial configurations, relationships, and be able to make inferences on their meaning far more quickly and efficiently than it would take to verbally describe a complex graphic.

There are several factors that influence spatial information processing using language to form mental models, such as the number of required models to solve a reasoning problem (Johnson-Laird, 2006), presentation order (Ehrlich and Johnson-Laird, 1982), use of transitive/non-transitive relations (Knauff and Ragni, 2011), binary/*n*-place relations (Goodwin and Johnson-Laird, 2005), and the differences in spatial reasoning on determinate/indeterminate problems (Byrne and Johnson-Laird, 1989). In many cases, sighted annotators (and in turn automated image captioning systems) often use non-transitive spatial relations such as “next to” or “contact” or “on the side” instead of transitive relations such as “left of” or “in front of” (Knauff and Ragni, 2011) to describe the spatial arrangements in images. Imagine the difficulties BVI students would face if they had to reason about a 100-point scatterplot that they could not see and instead, were provided with a list of 100 ordered pairs accompanied by a set of vague descriptions about their relative spatial positions (e.g., “point A (3,6) is on the side of point K (3,7) which is below point G”). Instead, we argue that spatial information must be explicitly incorporated into accessible learning systems to reflect current multisensory learning and model theory. This study investigates how high contrast visuals, sonification, vibrotactile haptic feedback, and spatial information descriptions collectively affect information retention, when compared to traditional accessibility solutions.

### 3. Design and development

#### 3.1. The multimodal diagram system

We designed and developed a multimodal diagram system (MDS) to investigate how effectively one platform could provide multisensory representations that would be accessible to all learners, especially those who have visual impairments. The MDS was specifically designed to be widely accessible and practical for diverse user populations, such as the broad spectrum of users with vision impairments and sighted users who require increased multimodal information access. The prototype system was designed on the iOS platform to leverage the many embedded universal design features in the native Apple iPhone UI, which accounts for why the vast majority of BVI smartphone users (72–80%) prefer to use iOS-based devices (Griffin-Shirley et al., 2017; WebAim.org, 2019). The MDS has two components: (1) the MDS *vibro-audio interface* mobile application and (2) an associated website that hosts a diagram library and an online diagram annotation/authoring tool for use by diagram creators.

##### 3.1.1. Vibro-audio interface mobile application

The MDS renders a high-contrast diagram on screen and provides audio and haptic feedback when the screen is touched, making diagram information access possible with or without vision. The MDS was designed to be used *via* single finger screen scanning. A short

vibration is triggered when the user moves their finger over the bounds of an onscreen element (i.e., moving from one element to another, such as from the front paneling of a house to the door). They can move their hand to follow/trace the vibratory element or listen to hear its name (tapping will repeat the auditory label). The dimensions of on-screen elements were informed by prior research into vibro-haptic interface design. For the MDS, the minimum width of lines was 4 mm and the minimum gap between lines was 4 mm (Palani et al., 2020). The vibration feature at the edge of an element was designed to be analogous to the raised lines between features of a traditional embossed tactile diagram and was implemented using the default iOS notification vibration.

Haptic/audio redundancy was integrated with the design: While a user’s finger touches a graphic element, such as the red circle in Figure 1, a constant element specific audio background tone is played. In addition, the name of the element is read *via* text-to-speech, and after a brief pause, a description of the element is read if the user’s finger remains within the red circle element. If the user moves their finger to enter a new element, such as the purple section in Figure 1, a vibration is triggered as their finger changes elements, then the unique audio tone, label and description begins again for this new graphic element. This procedure is based on guidance from earlier multimodal research (Choi and Walker, 2010).

##### 3.1.2. Multimodal diagrammatic system interface information flow

The MDS conveyed graphical information through images (visual), spoken words (auditory), sounds (auditory), and vibrations (haptic), where there were redundancies among the visual, auditory, haptic modalities along with kinesthetic cues (e.g., hand-movement and gestures). The MDS visually represented points as high-contrast, color-filled circles/vertices, lines/curves as 4 mm width high-contrast, colored line segments, and shapes/regions as high-contrast, color-filled areas. Simultaneously, points, lines, curves, and regions were each represented as distinct audio tones. The tones were programmatically generated for each element in a diagram by incrementally shifting a 180 Hz sine wave tone up in pitch depending on the number of elements in the diagram to assure that each element was represented by a distinguishing tone. Also simultaneously, points/lines/curves were represented haptically through the phones vibration motor that was activated whenever a finger touched that x-y point on the screen.

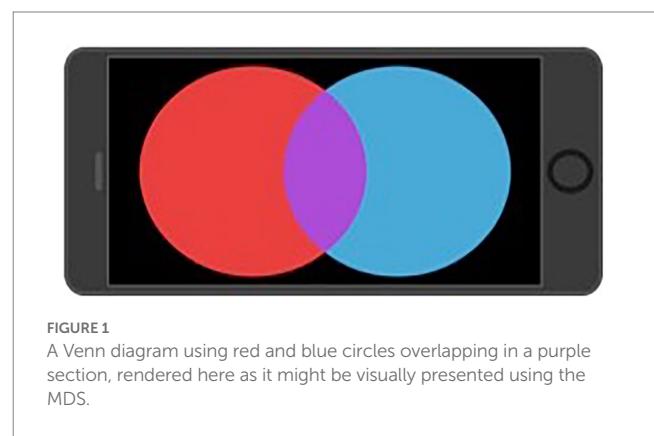


FIGURE 1

A Venn diagram using red and blue circles overlapping in a purple section, rendered here as it might be visually presented using the MDS.

When a diagram is loaded *via* the MDS, the system uses text-to-speech to read (auditory) the diagram title and, if present, an instructors note (e.g., instructing the user to begin their exploration at the bottom of the diagram). Element labels (e.g., “point p,” “line l”) were provided through native iOS text-to-speech and were played whenever a user’s finger entered the bounds of an element. Text-to-speech element descriptions followed 1 second (s) after the element label was read if the user remained within that element.

## 4. Materials and methods

The user study employed a perception-based (rather than a memory-based) information access task, where participants had access to the diagrams while they simultaneously completed worksheets related to the content. To control for pre-existing knowledge (i.e., variability in pre-test scores), pre-test and active-test worksheets were used to calculate normalized information gain scores. A finding of similar information gain between the MDS interface and traditional hardcopy stimuli would indicate that the MDS system is equivalently effective in conveying non-visual information. This design was motivated by previous work in the education and educational gaming literature (Furió et al., 2013). Similar procedures are typically used in education technology research to provide a “consistent analysis over diverse student populations with widely varying initial knowledge states” (Hake, 1998).

Users completed pre-test questions to establish their pre-existing knowledge on the diagram content. They then completed diagram content related worksheets while using diagrams in two different modal conditions: (1) using a traditional diagram (a visual or embossed/tactile diagram, between sighted and blind users, respectively), or (2) using the experimental MDS interface. The test worksheets used in this experiment were identical in format but not content to those used in the pre-test and were designed to emulate worksheets employed in standard STEM curricula. As the evaluation was designed as a perceptual task to determine whether the MDS could provide access for learning new information, not how well it could facilitate recall or mental representation of the information, the active test worksheet was completed with simultaneous access to the diagram. This pre-test/active test design was used so that normalized information gain could be calculated for each diagram.

### 4.1. Hypothesis

We hypothesized that the MDS interface would provide a *functionally equivalent* information access solution, resulting in similar results for worksheet accuracy and time to completion between worksheets completed with diagram access using a control condition: (1) traditional tactile stimuli (BVI control condition), or (2) visually-presented stimuli (sighted control condition). That is, we postulated that the use of the multisensory interface would allow participants with BVI to function at equivalent levels to their sighted peers.

*H0:* There will be no significant difference between groups (BVI, Sighted) completing worksheets in each condition (MDS, Control) providing a *functionally equivalent* information access experience.

This hypothesis is based on pilot user testing and previous work on functional equivalency (Giudice et al., 2013) that demonstrates the efficacy of vibro-audio interfaces in facilitating nonvisual access to spatial information using multimodal maps (Brock et al., 2010) touchscreen haptics (Palani and Giudice, 2016); spatial tactile feedback (Yatani et al., 2012) and related research on the application of multimodal interactive tools in education (Cairncross and Mannion, 2001; Moreno and Mayer, 2007). Previous work evaluating a vibro-audio interface noted slower encoding *via* learning with this type of interface in comparison to visual and traditional tactile graphics for both sighted and BVI users, although behavioral performance on testing did not differ (Giudice et al., 2012). Despite this difference in learning time, the same study found that the overall learning and mental representation of the diagram information (e.g., graphs, figures, and oriented polygons) was not reliably different between types of presentation modalities. Based on previous studies, we anticipated that with increased geometric complexity of the experimental diagrams, there could be increased worksheet completion times when using the MDS experimental interface. Similar testing of vibro-audio interfaces has found that nonvisual tracing of lines (audio or vibrotactile) rendered on a flat surface (e.g., touch screen) can be more challenging than following lines visually or using embossed tactile graphics (Giudice et al., 2012).

While this earlier work dealt with different STEM application domains, e.g., diagrams, shapes, maps, it was critical in the development and evaluation of this new multisensory interface in: (1) determining what parameters led to the most perceptually salient stimuli, (2) showing that using these multisensory stimuli led to accurate learning, mental representations, and other cognitive tasks using the interface, and (3) that it could support similar learning as was possible from existing/established modes of nonvisual information access (i.e., hardcopy tactile renderings). In other words, the early work dealt with design optimization and determining efficacy (e.g., does this system work or can stimulus x be learned using this approach?) By contrast, the pedagogy and motivation in this study is different, as we are now explicitly studying the nature of the learning and comparing this multisensory approach to existing *de facto* approaches using touch or vision between sighted and blind groups. Without this previous work, it would not be possible to use this interface here with any *a priori* knowledge of its efficacy. Our comparisons in this paper extend the previous work in multiple ways: (1) we are using very different STEM stimuli, (2) assessing its use in a knowledge gain task, and (3) comparing its use with both sighted and blind individuals (and their respective controls). An additional unique contribution of this study is the focus on the UDL nature of the system. Not only is the system being considered as an accessibility support for blind people but is conceptualized (and evaluated) as a universal support for all potential users, which has important applications for the use of multisensory devices for supporting generalized learning in a variety of STEM educational and vocational settings.

### 4.2. Methods

#### 4.2.1. Participants

The study included 29 working age adult participants: 20 participants (20) without vision impairment and nine participants

(9) with legally defined blindness. We were able to recruit a reasonably matched sample of BVI and sighted participants across age, gender, and education (Tables 1, 2). As age and gender are not critical factors in the outcomes of this study, we only report these participant data in the aggregate. This initial experiment recruited working age adults evaluate the MDS efficacy across a broad age range of adult users that would be representative of a variety of demographic groups (e.g., college and vocational learners). Recruitment was conducted through direct contact with people who have previously participated in lab studies, via a study recruitment ad distributed on several listservs for blind and vision loss communities, and by posting a bulletin board study recruitment ad in the community grocery store near the University.

The unbalanced design across participants reflects the typical challenges of recruiting research participants that are visually impaired. However, this sample size was sufficiently powered and is a similar size of traditional usability studies aimed at assessing the efficacy of assistive technology interface/device functionality for BVI populations (Schneiderman et al., 2018). The studies were reviewed and approved by the university's Institutional Review Board and all participants provided their written informed consent to participate in this study. All participants self-reported as having at least some college with several participants in both groups reporting they held graduate degrees. All participants in both groups reported as daily smartphone users. All BVI participants reported as being exclusively iPhone users, which is consistent with previous research on smartphone platform preference in the BVI community (WebAim.org, 2019). Among sighted participants smart phone usage was reported at 30% Android, 70% iPhone.

#### 4.2.2. Test interface

All participants used the iPhone-based experimental non-visual interface in the default iOS accessibility mode with a screen curtain on, thus completely disabling the phone's visual display. This was implemented to prevent any possible visual access to the presented diagram.

#### 4.2.3. Test science, technology, engineering, and mathematics content

Diagrams in all conditions were designed to provide equivalent information and be as similar as possible, while still

TABLE 1 Participant demographics summary.

Group	n=()	Sex	Age Range/ Mean	Highest education level
Participants w/ legally defined blindness	9	6 F, 3 M	21–70 M = 49.3	4 some college
				4 undergrad degree
				1 graduate degree
Participants w/o vision impairment	20	8 F, 12 M	19–35 M = 22.6	15 some college
				3 undergrad degree
				2 graduate degree

representing graphical rendering typical of the given modality. These specific diagrams were selected to represent topics normally presented graphically in a STEM curriculum. The two diagrams selected for use in this study were (1) layers of the atmosphere and (2) a helium atom (Figure 2). The images were created in a commercially available presentation slide platform and were based on diagrams of the same subjects from the American Printing House (APH) for the Basic Science Tactile Graphics set (American Printing House for the Blind (APH), 1997), which were used as a benchmark for BVI participants. The traditional visual diagrams used in the study were also based off the APH kit examples and adapted to resemble standard colored visual diagrams with text-based labels and a description key. Response protocol worksheets were designed to incorporate questions that demonstrated the efficacy of the interface in presenting both descriptive and spatial information. For example, each worksheet included questions regarding size and/or relative location of diagram elements, in addition to content questions regarding descriptions or functionality of the elements.

#### 4.2.4. Test procedure

A within-subjects, mixed factorial design was used in the experiment. Within-subjects factors were diagram type and presentation mode, with visual status being a between-subjects factor. All participants completed two pre-study worksheets based on the diagram content used in the study. Participants were then given access to one of the diagrams in each modal condition. The modal conditions differed slightly depending on participants' visual status. All participants completed a common condition using the experimental MDS interface, while their control/benchmark condition varied, with BVI participants using a hardcopy vacuum-formed tactile diagram (the gold standard for tactile-based renderings) and sighted participants using a visual diagram as their control/benchmark. Importantly for such cross-modal comparisons, the diagram elements were identical in each condition and the diagrams were scaled to the same size across condition. Diagram labels were provided using text-to-speech in the experimental MDS diagram condition (e.g., *Protons are positively charged particles in the nucleus of the atom*), using written text in the visual diagram condition, and given verbally by an experimenter acting as a learning assistant in the hardcopy tactile diagram condition.

The two conditions were administered in three phases: a pre-test phase, a practice phase, and an active-test phase, which was followed by a post-study questionnaire. Condition order and diagram presentation were counterbalanced to avoid order effects. In the pre-test phase, participants completed baseline-knowledge pre-tests for both diagrams. The pre-test percent accuracy score was used to represent participants' *a priori* knowledge of the content. Worksheets were given one at a time, and participant completion time was logged for each sheet. Worksheets were scored based on the number of questions answered correctly, out of the total number of questions (i.e., percent accuracy). This was then used to calculate normalized information gain for each participant. To ensure participants were able to demonstrate information gain, only those who initially made two or more errors were deemed eligible to continue in the study. Four sighted participants completed the

TABLE 2 BVI Participant demographics.

Participant number	Etiology of blindness	Onset age	Residual vision	Highest education level	Diagram use frequency
P1	Lebers congenital amaurosis	Birth	Light/ dark perception	Undergrad. degree	None
P2	Lebers congenital amaurosis	Birth	Light/dark perception	Graduate degree	Monthly
P3	Pathological myopia	45	Light/dark perception in right eye, Fuzzy colors	Undergrad. degree	None
P4	Retinitis pigmentosa, atypical, with cone dystrophy	25	Light/dark perception, some functional peripheral	Undergrad. degree	Weekly
P5	Retinitis pigmentosa	11	Light/dark perception	Some college	None
P6	Retinopathy of prematurity	Birth	Light/dark perception	Undergrad. degree	Yearly
P7	Glaucoma	16	Light/dark perception	Some college	Yearly
P8	Unknown	17	Light/dark perception	Some college	None
P9	Congenital cataracts, glaucoma	50	Light/dark perception	Some college	Monthly

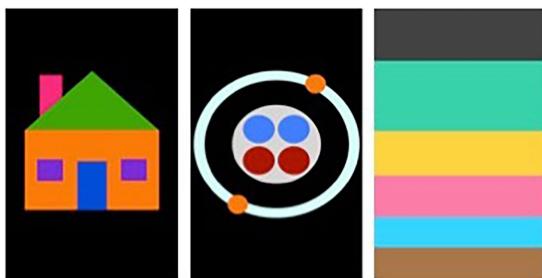


FIGURE 2

Example visualizations of the audio-haptic diagrams from left to right: practice diagram (house), atom diagram, and atmosphere diagram. Each color represents a unique element in the diagram. The phone screen was disabled via screen curtain throughout the study, so participants did not have visual access to the diagrams.

“Atom” diagram pre-test worksheet without any errors (i.e., earning a ceiling score), therefore, they did not continue the study and were replaced with new sighted participants. As we were looking for pre-test/active-test differences, it would not be possible to measure these differences if a participant’s pre-test was already at the ceiling, making the active-test data (if included) irrelevant.

In the practice phase, participants were provided an opportunity to practice with using the house diagram before each condition, either practicing with the experimental MDS interface (experimental condition) or traditional (tactile or visual) diagram (control conditions). All participants were asked simple spatial configuration questions about the sample diagrams, (e.g., for the house sample, *How many windows are present? Where are they relative to the door? What side of the house is the chimney on?*). These questions served as a criterion test to ensure all participants achieved basic competency using the interface before moving on

to the experimental trials and all participants were able to answer these questions during the practice phase. During the active-test phase, participants completed a worksheet with access to the test diagram. Worksheet completion time was measured as the duration of time required for the participants to complete the worksheets.

## 5. Results

### 5.1. Confirmation of learning

Analyses of descriptive statistics were conducted to compare prior (baseline) knowledge (based on pre-worksheet accuracy) and to confirm the presence of learning (by comparing pre- and active-worksheet accuracy) for BVI and sighted participants. These analyses were conducted by collapsing across diagram (atom vs. atmosphere) and mode (MDS vs. control). Pre-Accuracy descriptive results suggest the two groups were significantly different in their prior knowledge of using diagrams evident from the pre-worksheet accuracy mean percentage and the variability represented by the range of scores (Table 3).

Comparing data for each measure in the MDS and control conditions revealed remarkably similar values. For instance, there was only a 4% difference in active-worksheet accuracy between MDS (92%) and control (96%) conditions ( $t(27) = 1.3, p = 0.21$ ). This interaction was examined via Post-hoc comparisons of pre-and active-worksheet accuracy for BVI and sighted participants. Independent samples *t*-tests revealed prior knowledge (based on pre-worksheet accuracy) for BVI participants (13.7%) was significantly less than that of sighted participants (31.2%)  $t(27) = -4.16, p < 0.01$ .

In other words, the lower overall accuracy in BVI participants was due to significantly lower pre-test (but not active-test) accuracy when compared to sighted participants. Importantly,

TABLE 3 Pre/post accuracy collapsed across diagram and mode.

Visual status	Pre-accuracy mean (%)	SD	Range	95% CI	Active-accuracy mean (%)	SD	Range	95% CI
BVI	13.7	18.5	0–50	4.5–22.8	96.1	6.6	83.3–100	92.8–99.4
Sighted	31.2	34.9	0–75	20.1–42.4	92.0	10.7	66.7–100	88.5–95.4

although each group started with different levels of prior knowledge, their final scores (as measured by the active-test worksheet) were remarkably similar.

Furthermore, weighted mean accuracy data were submitted to a  $2 \times 2$  ((learning: pre- vs. active-worksheet accuracy)  $\times$  (group: BVI vs. sighted)) mixed-model ANOVA. Learning across all participants was evident in the difference between pre-test worksheet accuracy (25%) and active-test worksheet accuracy (94%),  $F(1,27) = 597.1$ ,  $\eta_p^2$  (partial eta<sup>2</sup>) = 0.9,  $p < 0.01$ . Collapsing across pre-test and active-test worksheet performance, overall accuracy was reliably lower for BVI participants (56%) than for sighted participants (64%),  $F(1,27) = 7.2$ ,  $\eta_p^2 = 0.2$ ,  $p < 0.01$ . However, this difference was likely driven by the interaction between participant group and learning mode,  $F(1,27) = 19.0$ ,  $\eta_p^2 = 0.4$ ,  $p < 0.01$ .

## 5.2. Information gain and completion time

In addition to the dependent variables of pre-test and active-test worksheet accuracy, the effect of the MDS interface and traditional hardcopy diagrams were also evaluated on two measures calculated to control for variance in pre-test knowledge (information gain and worksheet completion time). Individual normalized information gain scores reflect the improvement from pre- to post-test divided by the total amount of improvement possible ([gain = %posttest - %pretest] / [100 - %pretest]) and were calculated for each participant's performance in both modal conditions (Hake, 1998).

Worksheet completion time was calculated by dividing the time to complete the worksheet by the number of questions participants needed to answer on the worksheet (this varied based on pre-worksheet performance). Participants completed worksheets (two in total) using both modalities (MDS and control) and both diagrams (atom and atmosphere). Each diagram could only be tested once per participant; therefore, a full  $2$  (mode)  $\times$   $2$  (diagram) within-subjects design was not possible. Therefore, the effect of mode and diagram were each considered separately (collapsing across the other factor).

## 5.3. Information presentation mode

Pre-Accuracy descriptive results suggest the two groups were significantly different in their prior knowledge of using diagrams evident from the pre-worksheet accuracy mean percentage and the variability represented by the range of scores (Table 4). Analyses of descriptive statistics were conducted comparing data for each measure in the MDS and control conditions. These revealed remarkably similar values. There were greater Mean information gains for both groups. Additionally, the average time spent per question to complete the worksheet was also quite similar between MDS and control conditions.

The effect of mode was evaluated via  $2 \times 2$  ((mode: MDS vs. control)  $\times$  (group: BVI vs. sighted)) mixed MANOVA with active-worksheet accuracy, information gain, and worksheet completion time serving as the dependent measures. Neither the multivariate main effects nor the multivariate interaction reached significance (all  $p$ 's  $> 0.05$ ).

## 5.4. Diagram type

Analyses of descriptive statistics were conducted comparing diagram type by condition and these also revealed remarkably similar values (Table 5). Again, there was only a small difference in active-worksheet accuracy in the BVI group between MDS (94%) and control (98%) conditions and in the sighted group between the MDS (90%) and control (94%) conditions. Additionally, the average time spent per question to complete the active worksheet was also quite similar across diagram types between MDS and control conditions.

The effect of diagram type was evaluated via a  $2 \times 2$  ((diagram: atom vs. atmosphere)  $\times$  (group: BVI vs. sighted)) mixed MANOVA with pre-worksheet accuracy, active-worksheet accuracy, gain, and worksheet completion time serving as the dependent measures. Analyses revealed significant multivariate main effects of group, Wilks'  $\lambda = 0.5$ ,  $F(4,24) = 4.3$ ,  $\eta_p^2 = 0.4$ ,  $p < 0.01$  and diagram, Wilks'  $\lambda = 0.2$ ,  $F(4,24) = 25.1$ ,  $\eta_p^2 = 0.8$ ,  $p < 0.01$ , as well as a significant multivariate interaction between the two factors, Wilks'  $\lambda = 0.5$ ,  $F(4,24) = 6.2$ ,  $\eta_p^2 = 0.5$ ,  $p < 0.01$ . Given these results, univariate main effects and interactions are presented below.

### 5.4.1. Pre-worksheet accuracy

There was a significant difference in pre-worksheet accuracy between BVI (14%) and sighted (31%) participants with a greater variance in pre-test accuracy observed in the sighted participants,  $F(1,27) = 16.0$ ,  $\eta_p^2 = 0.3$ ,  $p < 0.01$  (Table 6). There was also a significant effect of diagram,  $F(1,27) = 95.3$ ,  $\eta_p^2 = 0.8$ ,  $p < 0.01$  with greater pre-worksheet accuracy for the atom diagram (43%) than the atmosphere (2%) diagram. The interaction between diagram type and group also reached significance,  $F(1,27) = 25.5$ ,  $\eta_p^2 = 0.5$ ,  $p < 0.01$  (Table 7).

Pre-test worksheet accuracy on the atom diagram was lower in BVI participants (24%) than that of the sighted participants (63%) with similar variability (Table 7). Independent sample  $t$ -tests revealed that was significantly less  $t(27) = -4.6$ ,  $p < 0.01$ . However, there were no reliable differences between BVI (4%) and sighted (0%) participants for the atmosphere diagram,  $t(8.000) = 1.5$ ,  $p = 0.17$  (corrected values reported due to heterogeneity of variance,  $F = 41.7$ ,  $p < 0.01$ ). Additionally, paired-sample  $t$ -tests revealed that BVI participants,  $t(8) = 2.9$ ,  $p < 0.05$ , and sighted participants,  $t(19) = 13.3$ ,  $p < 0.01$  had

TABLE 4 Pre/post accuracy collapsed across condition mode.

Visual status	Diagram	Gain mean (%)	SD	Range	95% CI	Pre-accur mean (%)	SD	Range	95% CI	Active accur. mean (%)	SD	Range	95% CI	Comp time mean (s)	SD	Range	95% CI
<b>BVI</b>	Atom	95.4	7.0	83.0–100.0	90.1–100.8	23.6	21.1	0.0–50.0	7.4–39.7	95.8	6.3	87.5–100	91.0–100.6	46.0	18.5	16.5–80.5	31.8–60.2
	Sky	96.2	7.5	83.0–100.0	90.5–102.0	3.7	7.35	0.0–16.7	−1.9–9.4	96.3	7.4	83.3–100	90.6–101.9	41.6	13.2	27.8–63.4	31.4–51.7
<b>Sighted</b>	Atom	79.2	25.2	33.0–100.0	67.4–90.9	62.5	21.1	0.0–75	52.6–72.4	93.1	8.6	75–100	89.1–97.1	65.7	28.8	35.0–158.0	52.2–79.2
	Sky	90.8	12.6	67.0–100.0	84.9–96.7	0.0.0	0.0	0.0–0.0	0.0–0.0	90.8	12.7	66.7–100	84.9–96.8	31.9	8.12	20.3–50.8	28.1–35.7

TABLE 5 Pre/post accuracy collapsed across diagram.

Visual status	Condition	Gain mean (%)				Pre-accur				Active accur				Comp time			
			SD	Range	95% CI		Mean (%)	SD	Range		Mean (%)	SD	Range		Mean (s)	SD	Range
<b>BVI</b>	MDS	93.6	7.8	83.0–100.0	87.5–99.6	6.0	9.6	0.0–25.0	−1.3–13.4	94.0	7.3	83.3–100.0	88.4–99.6	38.2	12.0	16.5–52.0	29.0–47.4
	Control	98.1	5.7	83.0–100.0	93.8–100.0	21.3	22.4	0.0–50.0	4.1–38.5	98.1	5.6	83.3–100.0	93.9–102.4	49.4	17.7	31.5–80.5	35.8–62.9
<b>Sighted</b>	MDS	85.8	17.1	50.0–100.0	77.8–93.8	27.5	34.1	0.0–75.0	11.6–43.4	90.2	11.5	66.7–100.0	84.8–95.6	51.6	30.1	21.7–158.0	37.5–65.6
	Control	84.2	23.9	33.0–100.0	73.0–95.3	35.0	36.2	0.0–75.0	18.1–51.9	93.8	9.9	66.7–100.0	89.1–98.4	46.1	24.0	20.3–101.0	34.8–57.3

better pre-worksheet accuracy for the atom diagram as compared to the sky/atmosphere diagram.

#### 5.4.2. Active-worksheet accuracy and information gain

For active-worksheet accuracy (Table 8), results suggest that with access to the MDS, the active worksheet accuracy for both groups were similar. In other words, neither of the main effects of group,  $F(1,27)=1.7$ ,  $\eta_p^2=0.1$ ,  $p=0.2$ , nor diagram,  $F(1,27)=0.1$ ,  $\eta_p^2=0.01$ ,  $p=0.7$ , nor the interaction,  $F(1,27)=0.3$ ,  $\eta_p^2=0.01$ ,  $p=0.6$  reached significance.

For information gain, neither of the main effects of group  $F(1,27)=3.7$ ,  $\eta_p^2=0.1$ ,  $p=0.1$ , nor diagram  $F(1,27)=2.4$ ,  $\eta_p^2=0.1$ ,  $p=0.1$ , nor the interaction  $F(1,27)=1.9$ ,  $\eta_p^2=0.1$ ,  $p=0.2$  reached significance.

#### 5.4.3. Worksheet completion time

There were small differences between groups in completing the worksheet tasks, however, the main effect of group on worksheet completion time,  $F(1,27)=0.7$ ,  $\eta_p^2=0.02$ ,  $p=0.4$  did not reach significance. The BVI participants Mean time in seconds (46 s) for the atom diagram was faster than the sighted participants Mean time (66 s) (Table 7) with the main effect of diagram  $F(1,27)=14.7$ ,  $\eta_p^2=0.3$ ,  $p<0.01$  reaching significance. This finding is not surprising given our *a priori* prediction that more geometrically complex diagrams would result in slower non-visual diagram access with the MDS.

The interaction between diagram and participant group (see Table 4) was also significant,  $F(1,27)=8.7$ ,  $\eta_p^2=0.2$ ,  $p<0.01$ . Independent samples *t*-tests revealed worksheet completion time for the sky/atmosphere diagram was significantly longer for BVI (42 s) participants than for sighted (32 s) participants (Table 7),  $t(27)=2.4$ ,  $p<0.05$ ; however, there were no differences in worksheet completion time between groups for the atom diagram,  $t(27)=-1.9$ ,  $p=0.07$ . Paired-samples *t*-tests did not reveal a significant difference in worksheet completion time for BVI participants,  $t(8)=0.6$ ,  $p=0.6$ . However, sighted participants took significantly longer for the atom (66 s) compared to the atmosphere (32 s) diagram (Table 7),  $t(19)=5.8$ ,  $p<0.01$ .

The differences in the completion time results may be attributed to the lack of familiarity with non-visual learning among the sighted participants, as well as the use of inefficient tactile scanning strategies by people who are not accustomed to learning through this modality. This interpretation is consistent with other studies showing that differences in tactile scanning strategies can impact the efficiency and accuracy of information acquisition and participant performance on spatial search tasks (Ungar et al., 1996).

#### 5.4.4. Visual status

A sub-analysis of the descriptive statistics for the BVI participant data was conducted to investigate the potential impact of any residual vision on the pre/post worksheet accuracy and completion time (Table 9). Looking at the raw data for performance of BVI participants, there does not appear to be any noteworthy differences based on visual status. Participants 3 and 4 reported having a small amount of residual vision, however their performance when compared to the other BVI participants does not suggest this improved their performance in terms of worksheet accuracy or completion time.

## 6. Discussion

This study began with the question: *Can multisensory spatial inputs (high contrast visual, spatial language, and haptics) lead to the same level of learning for concepts that are conveyed through diagrams?* To investigate this question, we designed a multisensory learning system to evaluate its ability to deliver functionally equivalent spatial information (configuration and relationships) to communicate diagrammatic content. The solution addressed two primary considerations: (1) the multisensory system is based on a universal design approach providing spatial information in diagrams for all learners (including BVI learners) to participate in classroom activities (e.g., groupwork) with their peers – thus reducing barriers presented in the need to create separate, specialized materials for accommodations; and (2) a significant body of research has confirmed that information presented in complementary sensory modalities can enhance the acquisition and retention of information for all learners (i.e., benefit of multisensory information). We hypothesized that the multisensory interface would provide a highly similar (*functionally equivalent*) spatial information access experience for both sighted and BVI participants. The results corroborate this prediction suggesting that all participants received a similar level of spatial information through multisensory input channels that facilitated functionally equivalent communication and interpretation of the diagrams' content and meaning.

### 6.1. Worksheet accuracy

Our hypothesis that there would be no significant difference in *worksheet accuracy* performance between the MDS interface and control stimuli was supported by the null results, as there were no statistically significant differences observed in active-test worksheet accuracy based on participant groups. While there was a significantly lower overall accuracy performance for BVI participants as compared to sighted participants, this was only due to their significantly lower pre-test accuracy (prior knowledge), not active-test accuracy (learning gain). Comparing accuracy results across conditions (MDS and control) revealed a numerically small and statistically insignificant difference (4%) in active-worksheet accuracy between MDS and control conditions, with the mean information gain across conditions only differing by 1%. Therefore, although each group started with different levels of prior knowledge, their final scores (as measured by the active-test worksheet) were remarkably similar, supporting our *a priori* prediction in the ability of the MDS interface to provide the necessary spatial information using multisensory channels to lead to similar learning gains. A possible alternative hypothesis, where gains were only found for the control condition, would suggest that learning was possible but with differential performance between the experimental MDS condition and the standard haptic/visual modes. The absence of this finding, based on the highly similar performance on final active worksheets between conditions argues against this outcome and suggests the MDS was as effective in supporting knowledge gain. There was a significant effect of diagram type (atom diagram vs. atmosphere diagram), however, this was found across both participant groups, further supporting the similarity performance between the MDS and control conditions.

## 6.2. Worksheet time to completion

We hypothesized there would be no significant differences between participants for *worksheet completion* when using the MDS system and this assertion was also supported by the results. Time spent per question to complete the worksheet was not significantly different between MDS (45 s) and control (48 s) conditions or between groups. We hypothesized that increased geometric complexity of the individual diagrams (atmosphere v. atom) would increase worksheet completion times for all participants when using the MDS interface and that was indeed validated by the observed results, with the more complex diagram (atmosphere) taking significantly longer for both groups to interpret and answer questions.

The most notable outcomes of this study are the remarkably similar data in participant accuracy and completion time between modalities (MDS vs. Tactile and MDS vs. Visual Control), which provides compelling evidence of the effectiveness of this interface compared to the gold standard diagram rendering techniques and suggests that there was a high level of similarity in information gain. These results are especially promising given sighted users' lack of experience with vibro-audio information access. This outcome suggests that the multisensory channels of information can provide a functionally equivalent learning experience for students who may need different types of information to understand a complex diagrammatic register exchange (Dimmel and Herbst, 2015). The MDS used multisensory input to provide redundant content information about the diagrams' meaning and interpretation through different channels simultaneously. Thus, participants in both groups could interpret the diagram's spatial information (i.e., spatial configuration and relationships) using vibro-audio input with similar performance as they could using more familiar modalities (tactile or visual). The MDS was able to successfully communicate the type of information needed to complete the diagrammatic register interchange using a combination of information input.

Our findings suggest that given a well-developed multisensory system, such as the MDS prototype, most participants were able to interpret spatial information within a diagrammatic representation well enough to make sense of the graphics using the vibro-audio interface. While the findings support this is true for the simple stimuli used in this study, we acknowledge that further research is needed to investigate if this finding of equivalent performance would hold for

more complex stimuli. In addition, the MDS system was effective for conveying non-visual information for working age adult learners with and without vision. This is an important finding as it represents a new universal design for learning approach for learners in a variety of STEM settings (e.g., college, vocational training, workplace professional development) to work cooperatively using the same reference materials and content platform, regardless of their visual status. The results also suggest this multisensory approach is viable as a multipurpose, affordable, mobile diagram display interface and as an accurate non-visual STEM graphical content learning tool. In the future, this type of MDS application could work in conjunction with diagram creation tutorials and an upload interface. As such, this multisensory approach addresses the long-standing challenge of providing consistent and timely access to accessible educational materials. With further development and testing, this type of system could have the important benefit of helping many learners with diverse learning needs who require additional multisensory supports from being left out of future STEM labor market opportunities due to a lack of adequate and accessible learning materials. These types of accessible STEM materials could help to improve low rates of STEM participation and career success by BVI students creating more accessible pathways for educational, employment, and lifestyle outcomes (Cryer et al., 2013; American Foundation of the Blind (AFB), 2017).

Our results provide further empirical support corroborating the growing body of evidence from multisensory learning demonstrating functionally equivalent performance. That is, when information is matched between inputs during learning, it provides a common level of access to key content, and the ensuing spatial image can be acted upon in an equivalent manner in the service of action and behavior, independent of the input source. Importantly, this study showed functional equivalence in two ways, similarity between learning inputs (i.e., the MDS vs. haptic and visual controls), and between participant groups (i.e., blind and sighted learners).

The finding of functional equivalence between our learning modalities is consistent with comparisons of these inputs (see Loomis et al., 2013 for review of this literature) but extends the theory to similar results in a new domain—interactions with STEM diagrams. The finding of equivalent performance between sighted and blind participants is also important as it supports the notion that when sufficient information is made accessible to these adult learners, they can perform at the same level as their sighted peers (Giudice, 2018). This outcome, as we observed here, speaks to the importance of providing accessible diagrams. However, as we discussed in the introduction, this access is not meant to support a specific population, providing information through multiple sensory modalities benefits all learners and is the cornerstone of good inclusive design. Indeed, we are all multisensory learners as this is how our brain works, taking in, learning, representing, and

TABLE 6 Summary of pre/post accuracy collapsed across diagram and mode.

Visual status	Pre-accuracy	Active-accuracy
BVI	13.7 [4.5–22.8]	96.1 [92.8–99.4]
Sighted	31.2 [20.1–42.4]	92.0 [88.5–95.4]

TABLE 7 Summary of pre/post accuracy collapsed across mode.

Visual status	Diagram	Gain	Pre-accuracy	Active- accuracy	Completion time
BVI	Atom	95.4 [90.1–100.8]	23.6 [7.4–39.7]	95.8 [91.0–100.6]	46.0 [31.8–60.2]
	Sky	96.2 [90.5–102.0]	3.7 [−1.9–9.4]	96.3 [90.6–101.9]	41.6 [31.4–51.7]
Sighted	Atom	79.2 [67.4–90.9]	62.5 [52.6–72.4]	93.1 [89.1–97.1]	65.7 [52.2–79.2]
	Sky	90.8 [84.9–96.7]	0 [0–0]	90.8 [84.9–96.8]	31.9 [28.1–35.7]

TABLE 8 Summary of pre/post accuracy collapsed across diagram.

Visual status	Condition	Gain	Pre-accuracy	Active-accuracy	Completion time
BVI	MDS	93.6 [87.5–99.6]	6.0 [-1.3–13.4]	94.0 [88.4–99.6]	38.2 [29.0–47.4]
	Control	98.1 [93.8–100.0]	21.3 [4.1–38.5]	98.1 [93.9–102.4]	49.4 [35.8–62.9]
Sighted	MDS	85.8 [77.8–93.8]	27.5 [11.6–43.4]	90.2 [84.8–95.6]	51.6 [37.5–65.6]
	Control	84.2 [73.0–95.3]	35.0 [18.1–51.9]	93.8 [89.1–98.4]	46.1 [34.8–57.3]

TABLE 9 Summary of BVI participant results by diagram and visual status.

ID #	Etiology of blindness	Residual vision	Mode	Diagram	Gain	Pre-accuracy	Active-accuracy	Completion time
1	Lebers congenital amaurosis	Light/ dark perception	MDS	Atom	100.0	12.5	100.0	52.0
				Sky	100.0	0.0	100.0	32.2
2	Lebers congenital amaurosis	Light/dark perception	MDS	Sky	100.0	16.7	100.0	27.8
				Atom	100.0	37.5	100.0	37.0
3	Pathological myopia	Light/dark perception in right eye, Fuzzy colors	MDS	Sky	83.3	0.0	83.3	36.7
				Atom	100.0	50.0	100.0	59.0
4	Retinitis pigmentosa, atypical, with cone dystrophy	Light/dark perception, some functional peripheral	MDS	Atom	87.5	0.0	87.5	16.5
				Sky	100	0.0	100.0	31.5
5	Retinitis Pigmentosa	Light/dark perception	MDS	Atom	83.3	25.0	87.5	51.8
				Sky	100.0	16.7	100.0	47.2
6	Retinopathy of Prematurity	Light/dark perception	MDS	Sky	100.0	0.0	100.0	39.8
				Atom	100.0	50.0	100.0	80.5
7	Glaucoma	Light/dark perception	MDS	Sky	100.0	0.0	100.0	33.7
				Atom	100.0	37.5	100.0	31.6
8	Unknown	Light/dark perception	MDS	Atom	87.5	0.0	87.5	50.6
				Sky	83.3	0.0	83.3	63.4
9	Congenital Cataracts, Glaucoma	Light/dark perception	MDS	Atom	100.0	0.0	100.0	34.8
				Sky	100.0	0.0	100.0	61.8

acting upon information from multiple inputs in a seamless and integrative manner. The key role of multimedia and complementing sensory modalities has been shown to enhance the acquisition and retention of information in dozens of contexts and situations (Mayer, 2002). The current work builds on this literature. Our findings not only support the possibility of functional equivalence for STEM learning outcomes when an inclusive, universal-designed system is available, they also show that such multisensory interfaces

benefit all learners and have the potential for many applications beyond traditional accessibility.

## 7. Limitations and future work

As this was a prototype designed for this study, there are limitations in the design of the current MDS system that could be improved with

additional technical development. For instance, while the MDS application was designed for use with both vibration /haptic feedback, this component could be augmented and enhanced in future incarnations. New user interface (UI) elements being developed by our group and collaborators support new haptic profiles that would allow for a greater array of patterns, vibration styles, and haptic interactions with the MDS. Incorporating this development into future MDS design would allow for improved mapping of different diagram elements to haptic feedback. This would provide enhanced stimulus–response pairings that would likely both increase the type of information that could be presented through this modality and the overall efficiency of information encoding and learning strategy when using the MDS. In addition, work by our group and others on automating natural language descriptions could improve how key visual elements are conveyed through speech description when such annotations are created through an automatic vs. human-generated process. We also recognize the fact that the MDS may not be able to communicate other types of diagrams (e.g., charts, graphs, maps, etc) with the same level of effectiveness as the ones used in this study. We are in the process of running additional studies with new MDS features to explore the multisensory system's effectiveness with these additional types of visual representations. An additional consideration not addressed in this study is that while the MDS system was designed to support creation of accessible content, it still involves a significant amount of human intervention. Automating this process is a long-term goal of this project that would greatly streamline the creation of accessible content. In addition to the design limitations, our ability to differentiate among the groups was limited by the ceiling effect of our measurements. It is important to note that there was no reduction in performance across the comparison groups, but in future studies we plan to use more sensitive measures to investigate how variations in modality affect diagrammatic perception. Finally, future studies will need to evaluate the system with specific demographics (e.g., school/college aged people for classroom use, people in vocational settings for supporting work contexts, etc) to fully validate its use across learning settings.

## 8. Conclusion

The Multimodal Diagram System was designed with both sighted and BVI learners in mind. The goal of the MDS design was to create a STEM graphical content learning tool that could be used by all students to help facilitate communication and discussion between peoples with different visual abilities in a classroom. The results of this experiment provide clear support for the efficacy of our approach and of the MDS as a new, universally designed tool for providing inclusive STEM access for all.

## Data availability statement

The datasets presented in this article are not readily available under IRB constraints. Requests to access the datasets should be directed to the corresponding author.

## References

Abrahamson, D., Flood, V. J., Miele, J. A., and Siu, Y. T. (2019). Enactivism and ethnomethodological conversation analysis as tools for expanding universal Design for

## Ethics statement

The studies involving human participants were reviewed and approved by University of Maine IRB, under application number: 2014-06-12. The participants provided their written informed consent to participate in this study.

## Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## Funding

This research was also funded by NSF grant IIS1822800 A Remote Multimodal Learning Environment to Increase Graphical Information Access for Blind and Visually Impaired Students (PI Giudice, Co-PIs Dimmel and Doore) and by NSF grant 2145517 CAREER: A Transformative Approach for Teaching and Learning Geometry by Representing and Interacting with Three-dimensional Figures (PI Dimmel).

## Acknowledgments

We thank the University of Maine Center for Undergraduate Research (UMaine CUGR) and UMaine ASAP Labs for early funding and technical assistance in the development phase of the MDS system.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Learning: the case of visually impaired mathematics students. *ZDM* 51, 291–303. doi: 10.1007/s11858-018-0998-1

American Foundation of the Blind (AFB). (2017). *Interpreting bureau of labor statistics employment data*. Louisville, KY. Available at: <http://www.afb.org/info/blindness-statistics/adults/interpreting-bls-employment-data/234>

American Printing House for the Blind (APH) (1997). *Guidelines for design of tactile graphics*. Louisville, KY: APH.

American Printing House for the Blind (APH) (2019). *Annual Report 2019*. Louisville, KY: American Printing House for the Blind, Inc.. Available online at: <http://www.aph.org/annual-reports>

Avraamides, M. N., Loomis, J. M., Klatzky, R. L., and Golledge, R. G. (2004). Functional equivalence of spatial representations derived from vision and language: evidence from allocentric judgments. *J. Exp. Psychol. Learn. Mem. Cogn.* 30:804. doi: 10.1037/0278-7393.30.4.804

Blaser, B., and Ladner, R. E. (2020). "Why is data on disability so hard to collect and understand?" in *2020 research on equity and sustained participation in engineering, computing, and technology (RESPECT)*, vol. 1 (Portland OR: IEEE), 1–8.

Brock, A., Truillet, P., Oriola, B., and Jouffrais, C. (2010). Usage of multimodal maps for blind people: why and how in ACM International Conference on Interactive Tabletops and Surfaces (November 2010) 247–248. Atlanta, GA: ACM. doi: 10.1145/1936652.1936699

Byrne, R., and Johnson-Laird, P. N. (1989). Spatial reasoning. *J. Mem. Lang.* 28, 564–575. doi: 10.1016/0749-596X(89)90013-2

Cairncross, S., and Mannion, M. (2001). Interactive multimedia and learning: realizing the benefits. *Innov. Educ. Teach. Int.* 38, 156–164. doi: 10.1080/14703290110035428

Chandler, D. (2007). *Semiotics: The basics*. New York, NY: Routledge.

Choi, S. H., and Walker, B. N. (2010). "Digitizer auditory graph: making graphs accessible to the visually impaired" in *CHI'10 extended abstracts on human factors in computing systems*, (Atlanta, GA: ACM), 3445–3450.

Cryer, H., Cryer, H., Home, S., Wilkins, S. M., Cliffe, E., and Rowe, S. (2013). *Teaching STEM subjects to blind and partially sighted students: Literature review and resources*. Birmingham, UK: RNIM Centre for Accessible Information.

Dimmell, J. K., and Herbst, P. G. (2015). The semiotic structure of geometry diagrams: how textbook diagrams convey meaning. *J. Res. Math. Educ.* 46, 147–195. doi: 10.5951/jresmatheduc.46.2.0147

Doore, S. A., Dimmell, J. K., Xi, R., and Giudice, N. A. (2021). "Embedding expert knowledge: a case study on developing an accessible diagrammatic interface" in *Proceedings of the forty-third annual meeting of the north American chapter of the International Group for the Psychology of mathematics education*. eds. D. Olanoff, K. Johnson and S. Spitzer (Philadelphia, PA: Psychology of Mathematics Education), 1759–1763.

Ehrlich, K., and Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *J. Verbal Learn. Verbal Behav.* 21, 296–306. doi: 10.1016/S0022-5371(82)90626-0

Furió, D., González-Gancedo, S., Juan, M. C., Seguí, I., and Costa, M. (2013). The effects of the size and weight of a mobile device on an educational game. *Comput. Educ.* 64, 24–41. doi: 10.1016/j.compedu.2012.12.015

Giudice, N. A. (2018). "Navigating without vision: principles of blind spatial cognition," in *Handbook of behavioral and cognitive geography* (Northampton MA: Edward Elgar Publishing) doi: 10.4337/9781784717544.00024

Giudice, N. A., Betty, M. R., and Loomis, J. M. (2011). Functional equivalence of spatial images from touch and vision: evidence from spatial updating in blind and sighted individuals. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 621–634. doi: 10.1037/a0022331

Giudice, N. A., Klatzky, R. L., Bennett, C. R., and Loomis, J. M. (2013). Combining locations from working memory and long-term memory into a common spatial image. *Spat. Cogn. Comp.* 13, 103–128. doi: 10.1080/13875868.2012.678522

Giudice, N. A., Palani, H. P., Brenner, E., and Kramer, K. M. (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, 103–110. Boulder, CO: ACM. doi: 10.1145/2384916.2384935

Goodwin, G. P., and Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychol. Rev.* 112, 468–493. doi: 10.1037/0033-295X.112.2.468

Griffin-Shirley, N., Banda, D. R., Ajuwon, P. M., Cheon, J., Lee, J., Park, H. R., et al. (2017). A survey on the use of mobile applications for people who are visually impaired. *J. Vis. Impair. Blindn.* 111, 307–323. doi: 10.1177/0145482X1711100402

Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66, 64–74. doi: 10.1119/1.18809

Hasty, L., Milbury, J., Miller, I., O'Day, A., Acquinas, P., and Spence, D. (2011). *Guidelines and standards for tactile graphics: Technical report*, braille Authority of North America, 2011. Available online at: <http://www.brailleauthority.org/tg/>

Heward, W. L. (2003). *Exceptional children: An introduction to special education*, New Jersey: Pearson Education.

Honey, M., Alberts, B., Bass, H., Castillo, C., Lee, O., and Strutches, M. M. (2020). *STEM education for the future: a visioning report*. National Science Foundation Technical Report, 36.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*, Cambridge, MA: Harvard University Press, 6.

Johnson-Laird, P. N. (2006). "Mental models, sentential reasoning, and illusory inferences" in *Advances in psychology*, vol. 138 (North-Holland), 27–51. doi: 10.1016/S0166-4115(06)80026-9

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proc. Natl. Acad. Sci.* 107, 18243–18250. doi: 10.1073/pnas.1012933107

Johnson-Laird, P. N., and Byrne, R. M. (1991). *Deduction*, Hove, East Sussex, United Kingdom: Lawrence Erlbaum Associates, Inc..

Kent, M., Ellis, K., Latter, N., and Peaty, G. (2018). The case for captioned lectures in Australian higher education. *TechTrends* 62, 158–165. doi: 10.1007/s11528-017-0225-x

Khine, M. (2013). *Critical analysis of science textbooks. Evaluating instructional effectiveness*. Netherlands: Springer. doi: 10.1007/978-94-007-4168-3

Klatzky, R. L., Lippa, Y., Loomis, J. M., and Golledge, R. G. (2003). Encoding, learning, and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision. *Exp. Brain Res.* 149, 48–61. doi: 10.1007/s00221-002-1334-z

Knauff, M., and Ragni, M. (2011). Cross-cultural preferences in spatial reasoning. *J. Cogn. Cult.* 11, 1–21. doi: 10.1163/156853711X568662

Kress, G., and van Leeuwen, T. (2006). *Reading images: The grammar of visual design (2nd ed.)*. London, England: Routledge. doi: 10.4324/9780203619728

Loomis, J. M., Klatzky, R. L., and Giudice, N. A. (2013). "Representing 3D space in working memory: spatial images from vision, hearing, touch, and language" in *Multisensory imagery: Theory & Applications*. eds. S. Lacey and R. Lawson (New York: Springer), 131–156. doi: 10.1007/978-1-4614-5879-1\_8

Lundgard, A., and Satyanarayan, A. (2021). Accessible visualization via natural language descriptions: a four-level model of semantic content. *IEEE Trans. Vis. Comput. Graph.* 28, 1073–1083. doi: 10.1109/TVCG.2021.3114770

Mayer, R. E. (2002). "Multimedia learning" in *Psychology of learning and motivation*, vol. 41 (New York, NY: Academic Press), 85–139. doi: 10.1016/S0079-7421(02)80005-6

McDonnell, M. C., and Sui, Z. (2019). Employment and unemployment rates of people who are blind or visually impaired: estimates from multiple sources. *J. Vis. Impair. Blindn.* 113, 481–492. doi: 10.1177/0145482X19887620

Moon, N. W., Todd, R. L., Morton, D. L., and Ivey, E. (2012). *Accommodating students with disabilities in science, technology, engineering, and mathematics (STEM)*. Atlanta, GA: Center for Assistive Technology and Environmental Access, Georgia Institute of Technology.

Moreno, R., and Mayer, R. E. (2002). Verbal redundancy in multimedia learning: when reading helps listening. *J. Educ. Psychol.* 94, 156–163. doi: 10.1037/0022-0663.94.1.156

Moreno, R., and Mayer, R. (2007). Interactive multimodal learning environments. *Educ. Psychol. Rev.* 19, 309–326. doi: 10.1007/s10648-007-9047-2

National Center for Education Statistics (NCES) (2022). "Students with disabilities" in *Condition of education* (U.S. Department of Education, Institute of Education Sciences. Available online at: <https://nces.ed.gov/programs/coe/indicator/cgg>

National Federation of the Blind (NFB). *Blindness statistics*. (2009). Available at: <https://nfb.org/blindness-statistics>

Palani, H. P., Fink, P. D. S., and Giudice, N. A. (2020). Design guidelines for schematizing and rendering haptically perceptible graphical elements on touchscreen devices. *Int. J. Hum. Comput. Interact.* 36, 1393–1414. doi: 10.1080/10447318.2020.1752464

Palani, H. P., and Giudice, N. A. (2016). *Usability parameters for touchscreen-based haptic perception*. In *2016 IEEE Haptics Symposium (HAPTICS)*. (Philadelphia, PA, USA: IEEE).

Ragni, M., and Knauff, M. (2013). A theory and a computational model of spatial reasoning with preferred mental models. *Psychol. Rev.* 120, 561–588. doi: 10.1037/a0032460

Schles, R. A. (2021). Population data for students with visual impairments in the United States. *J. Vis. Impair. Blindn.* 115, 177–189. doi: 10.1177/0145482X211016124

Schneiderman, B., Plaisant, C., Cohen, M., and Jacobs, S. (2018). *Designing the user interface: Strategies for effective human-computer interaction*, 6th ed.; Addison Wesley: Boston, MA, United States.

Sharif, A., Chintalapati, S. S., Wobbrock, J. O., and Reinecke, K. (2021). Understanding screen-reader users' experiences with online data visualizations in In Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21), Association for Computing Machinery, New York, NY, USA, 1–16.

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., et al. (2009). Why peer discussion improves student performance on in-class concept questions. *Science* 323, 122–124. doi: 10.1126/science.1165919

Tipton, K. (2021). *Video captions as a universal design technique: impact on student success*. California State University: Northridge.

Trickett, S. B., and Trafton, J. G. (2006). "Toward a comprehensive model of graph comprehension: making the case for spatial cognition" in *International conference on theory and application of diagrams* (Springer, Berlin, Heidelberg: Springer), 286–300.

U.S. Census Bureau (2020). *2019 American community survey 1-year estimates: Employment status by disability status and type, Table B18120*. Available online at: <https://data.census.gov/cedsci/tableq=B18120&tid=ACSDT1Y2019.B18120>

Ungar, S., Blades, M., and Spencer, C. (1996). "The construction of cognitive maps by children with visual impairments" in *The construction of cognitive maps* (Dordrecht: Springer), 247–273. doi: 10.1007/978-0-585-33485-1\_11

Vickers, P., Faith, J., and Rossiter, N. (2012). Understanding visualization: a formal approach using category theory and semiotics. *IEEE Trans. Vis. Comput. Graph.* 19, 1048–1061. doi: 10.1109/TVCG.2012.294

Voltz, D. L., Brazil, N., and Ford, A. (2016). What matters most in inclusive education. *Interv. Sch. Clin.* 37, 23–30. doi: 10.1177/105345120103700105

W3C (2019). Web Content Accessibility Guidelines (WCAG) [http://www.w3.org/TR/WCAG/W3C\\_is\\_the\\_publisher](http://www.w3.org/TR/WCAG/W3C_is_the_publisher). *WAI web accessibility tutorials: Complex images*

WebAim.org (2019). *Survey on screen reader Usage #9*. Available online at: <https://webaim.org/projects/screenreadersurvey9/>

Yatani, K., Banovic, N., and Truong, K. (2012). SpaceSense: representing geographical information to visually impaired people using spatial tactile feedback in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin, TX: ACM. 415–424.

Yelland, N. J. (2018). A pedagogy of multi-literacies: young children and multimodal learning with tablets. *Br. J. Educ. Technol.* 49, 847–858. doi: 10.1111/bjet.12635

Zhang, M. R., Zhong, M., and Wobbrock, J. O. (2022). “Ga11y: an automated GIF annotation system for visually impaired users,” in *Proceedings of the ACM conference on human factors in computing systems (CHI’22)* (New Orleans, New York, United States: ACM), 1–16.