Touch-Screen Technology for the Dynamic Display of 2D Spatial Information Without Vision: Promise and Progress

Roberta L. Klatzky [1,*], **Nicholas A. Giudice** [2], **Christopher R. Bennett** [2] **and Jack M. Loomis** [3]

[1] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[2] Spatial Informatics Program, School of Computing and Information Science, Orono, ME 04469, USA

[3] Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA 93106, USA

Abstract

Many developers wish to capitalize on touch-screen technology for developing aids for the blind, particularly by incorporating vibrotactile stimulation to convey patterns on their surfaces, which otherwise are featureless. Our belief is that they will need to take into account basic research on haptic perception in designing these graphics interfaces. We point out constraints and limitations in haptic processing that affect the use of these devices. We also suggest ways to use sound to augment basic information from touch, and we include evaluation data from users of a touch-screen device with vibrotactile and auditory feedback that we have been developing, called a vibro-audio interface.

Keywords

Perception, multi-modal, touch, sensory substitution, touch-screen, blind, graphics, graphs, auditory, vibration

1. Introduction

Graphs, figures, and charts are essential tools in education and workplace success. While sighted individuals take for granted that these materials will be available, the situation is very different for the blind, as graphical information is almost exclusively constrained to visual rendering. Legally blind persons

---

[*] To whom correspondence should be addressed. E-mail: klatzky@cmu.edu

constitute a sizeable population (approximately 3.4 million people in the U.S. and 39 million people worldwide according to the World Health Organization, 2011) that could be helped by technologies for conveying graphical materials, just as text-to-speech has been developed to provide verbal information. If graphical images from textbooks were readily conveyed to blind students in the classroom through portable media, they would be able to assimilate the relevant content along with text and follow a more integrated curriculum. To the contrary, however, many graphical educational materials are simply not translated into a modality accessible without sight, while those that are conveyed to the blind may be difficult if not impossible to interpret. The lack of access to displays intended for visual use may undermine blind people's ability to conceptualize and learn from quantitative data (Walker and Mauney, 2010). Similarly, lack of access to visually-oriented graphical material imposes significant vocational challenges and certainly contributes to the high levels of under-employment and unemployment by this demographic (Kaye *et al.*, 2000).

Efforts to deal with this problem, in the form of raised tangible graphics, can be traced back for centuries (Perkins Museum, 2013). The rise of the computer led to interfaces intended for the blind, using auditory (verbal or non-verbal), haptic, or multi-modal cues. Such displays predominantly rely on exploration of the graphical content *via* keyboard or mouse, although force-feedback devices have also been used (Ferres *et al.*, 2010; Walker and Mauney, 2010; Yu and Brewster, 2002a, b).

As touch-screen technologies have come widely into use, the question naturally arises as to whether devices of this type could be employed to provide blind people with access to graphics. On the positive side, they are widely available at a reasonable cost. The capability of being refreshed means that essentially unlimited archives of data can be displayed on a single portable surface. Features such as rescaling and translation of the images across the surfaces can potentially be used to enhance information content. Graphical text can also readily be linked to speech and nonverbal sound, and augmentation with vibration adds yet further modalities of communication.

While touch-screen displays offer great promise as graphical displays for blind people, they also offer challenges. Because the surfaces are smooth, the approach to displaying graphics must be different from one based on raised tangible elements. In general, touch-screens work by linking features such as symbols, lines, or familiar shapes to screen locations. For use without sight, an external cue (e.g., sound or vibration) representing some graphical element must be emitted when the user's exploring finger contacts the corresponding location. In order to form an integrated representation of the information layout, the user must (a) use kinesthetic sensory cues to keep track of his or her finger position within some frame of reference, as defined by the body or ex-

ternal landmarks such as the edges of the display, (b) interpret the external cue, and (c) associate the cued content with the currently contacted coordinates.

All of these processing components potentially present challenges, but the first is of particular interest here, as it corresponds to forming a 'mental map' of the screen layout. Essentially, non-visual touch-screen graphical displays require that the user form this map by mentally tracking the positions of his or her fingers as they traverse a flat, featureless terrain. Challenges arise in terms of the user's controlling finger position in relation to the display, encoding local elements and contours, and building an enduring layout representation. In the sections that follow, we will describe these challenges in further detail in the context of fundamental perceptual and cognitive capabilities of human users. We first frame the problem by considering what content a 2D, refreshable touch-screen application might aim to deliver.

2. The Content of 2D Information for Perception Without Vision

To a sighted audience, touch-screen technologies can convey spatial information of enormous complexity, limited only by the screen size and its spatial resolution as well as by visual spatial acuity. For visual images and visual scenes too large to fit on the display at any one time, pan and zoom operations can be used to compensate, but such operations are poorly studied and difficult to implement without careful consideration using non-visual displays (Rastogi *et al.*, 2013). When one considers an audience of users who will be using the technology without vision, the potential information content becomes considerably more limited. For present purposes, we will consider four general categories of information to be conveyed: symbols, maps, graphs, and pictures.

Graphical symbols are arguably the simplest of these categories, by convention if not definition, for they consist of simple component features such as lines and curves. A symbol can bear an arbitrary relation to what it represents; for example, a filled dot on a graph can be assigned to any quantitative data source. An icon is a symbol that bears a more direct pictorial or conceptual correspondence to its referent. Tversky (2003) adds to these graphical categories that of the *morphogram*, which includes arrows, lines and line intersections, blobs, and other elements that convey meaning by simple geometry.

Maps and graphs are more complex entities that tend to include symbols, icons, and morphograms. A 2D map uses a local frame of reference to represent 2D spatial layout in some other frame of reference, usually with transformation of scale. There are, of course, many variants of maps, depending on their topological fidelity, what features or content of the represented world are retained in the depiction, and their use of symbolic content. Graphs are

more general than maps, in that they use a 2D frame of reference to represent arbitrary quantitative data, not just spatial in source.

Pictures are the most general category of 2D content. We reserve this term for displays that convey objects and (when present) inter-object relations. The complexity of pictures varies from sparse line drawings to complex scenes.

3. Graphical Displays as an Example of Content Challenges

Clearly, the capability of touch-screen technology to convey any of the four categories of information that we have described far exceeds what can be extracted from the technology without vision. To make this point clearer, let us focus on the category of graph displays of data. In general, graphs for the blind are intended to aid in quantitative understanding. To frame this challenge, consider a set of general goals for 'graph literacy' that has been proposed by Galesic and Garcia-Retamero (2011), in the form of an assessment tool. Three general competencies of increasing complexity are included: extracting data directly portrayed on a graph, understanding relationships between data values, and drawing inferences. Even with simple graphs, and with vision intact, these processes are demanding (e.g., Carpenter and Shah, 1998; Zacks and Tversky, 1999).

We believe the most important limitations on technologically enabled graphical rendering for the blind will be at the front end: Extracting the basic data, the simplest domain assessed by Galesic and Garcia-Retamero (2011), is likely to be the critical factor. Visual graphs, of course, can incorporate multiple variables and use visually based cues such as 3D perspective and shading. The success of non-visual graphics, in contrast, will depend on setting constraints on the complexity of content; for example, appropriate down-sampling of visual information on maps, presenting at most bivariate relationships *via* bars and lines on graphs, and limiting the number of segments in pie charts.

4. Tangible Graphics on Refreshable *vs.* Non-Refreshable Media

As was noted above, raised graphics for the blind have been created for centuries *via* a variety of media (Perkins Museum, 2013). Considerable research effort has been devoted to design issues (for a web compendium of research on tangible maps, see Perkins Maps, 2013). At the symbol level, there are concerns with legibility, discriminability, and interpretability (e.g., Berla, 1973; Loomis, 1990; Nolan and Morris, 1971). At the global level, there is the problem of balancing information content with clutter; veridicality with simplicity, and so on (e.g., Franks and Baird, 1971; Golledge, 1991).

Another concern is how to support the understanding of spatial layout within a frame of reference. Millar (e.g., 2008) has emphasized the variety

of frames that can be established when vision is absent, including body-based, movement-based, and external reference frames. Her work (Millar, 1981) suggests that visual experience facilitates coding within an external frame of reference, and that lack of such experience leads to reliance on body-based features to locate elements in space. Touch-screens are frequently moved with respect to the body under typical use, reducing the utility of self-referred spatial coordinates. If blind readers have only the edges of the map or graph they are reading to provide a coordinate frame, they may lose track of the spatial relations among elements; thus users tend to prefer tangible grids over empty space (Lederman and Campbell, 1982).

Raised-line graphs offer advantages in that their tangible elements are quickly found by scanning with the hands. Moreover, as will be discussed below in the context of tactile perceptual limitations, continuous lines intrinsically provide guidance to the reader. On the negative side, however, tangible graphics are static and relatively expensive to author and produce, even with modern printing technologies. Moreover, they are single-purpose, and their physical bulkiness and weight reduce transportability and impose demands on storage space. These characteristics clearly point to advantages for refreshable touch-screen displays supported by electronic data storage.

Touch-screen displays, despite their name, offer no local tactile cues. Rather, as was noted in the Introduction, they offer a spatial mapping between stored data and locations on the flat display. To obtain spatial data, the user must link the spatial location that is touched with the data content that is mapped to that location. For a touch-screen display to be useful for the blind, then, there needs to be a way of communicating when the finger is on or near a spatial feature of interest. A principal channel for cuing *via* the touch modality is vibration, which can be produced by embedded electromagnetic motors as in smartphones or tablets (Giudice *et al.*, 2012; Poppinga *et al.*, 2011), electrostatic surfaces (Bau *et al.*, 2010), or piezoelectric elements (Winfield *et al.*, 2007). Refreshable braille is another communication channel, but as the braille elements are spatially displaced from the touch-screen, this would require feeling the braille with one hand while exploring with another.

Each of the current vibratory technologies has positive and negative features. All share the property that the entire fingertip is stimulated once contact with a feature is identified. Embedded motor vibrators tend to be high-amplitude and to produce auditory noise. Surfaces with piezoelectric drivers act by reducing friction, but the range of variation is currently quite limited. These and electrostatic-based displays have the additional feature that the stimulation is motion-dependent. However, the intensely active nature of this research area will no doubt refine technologies considerably in the near future.

In the sections that follow, we review more pervasive and long-term challenges for flat-screen technology, reflecting limitations in the sense of touch. We also suggest how audition might be used to augment and improve on the basic features of touch-screens.

5. Implications of Touch Processing for Display Technology

It is obvious that the sense of touch cannot be a wholesale substitute for vision. What is not so obvious, however, is that haptic perception differs from vision in important features that would be expected to affect graphical understanding. These include access to data, sensory adaptation, spatial resolving power (acuity), temporal integration, spatial localization, and vulnerability to systematic distortion.

Reviewing all these aspects of touch would be beyond the scope of the present article (see, e.g., Lederman and Klatzky, 2009 for a brief review and list of resources). First we briefly consider limitations at the perceptual periphery. One is the limited spatial bandwidth. The threshold separation on the finger needed to resolve closely spaced contact points is on the order of 1–2 millimeters (e.g., Craig, 1999; Loomis, 1981). Because this value is large relative to the available surface on the fingers, the total number of resolvable points is tiny in comparison to that of vision (Loomis *et al.*, 2012). The spatial sensitivity of receptors in fingertip skin is sufficient to permit the perception of small embossed patterns pressed on the fingertip (Loomis, 1981, 1990), although for sighted perceivers, this ability declines with age (Legge *et al.*, 2008). The spatial bandwidth limitation is exacerbated on the back and abdomen, where receptors are relatively sparsely distributed. Other spatial limitations are found in localization of a point contact relative to the body, or multiple discrete contact points relative to one another. These capabilities are subject to the imprecision of the kinesthetic sense, which relies on receptors in muscles, tendons, and joints (Klatzky and Lederman, 2003).

Temporal processing is another limitation: Peripheral neurons exhibit adaptation to continuing stimulation, with some populations decreasing their response rates quite rapidly (Johansson and Vallbo, 1983). One effect is that vibration, which might be used in touch-screen displays to indicate contact with an informative element (see above), can eventually lead to a feeling of tingling or numbness when repeated or used for a prolonged period of time (Raja, 2011).

From the perspective of designers of non-visual graphics, even more important are difficulties that are introduced by the active nature of touch encoding. The limited spatial bandwidth of touch means that complex images cannot be simultaneously presented within the span of a single fingertip or even across a larger, but less densely innervated, skin surface such as the back. Data pre-

sented in flat displays must extend beyond the fingertip, and as a result, must be encoded by a sequence of touches, either continuous or discrete. The extended nature of exploration means that spatial elements that are distributed over time must be integrated into whole shapes, in order to be meaningful. Unlike the visual system, which has developed mechanisms to integrate spatial input across fixations (Merriam *et al.*, 2007), the haptic system appears to have no specialized processors for synthesizing successive points of contact. To the contrary, there appear to be fundamental cognitive limitations on the process of building objects over space and time.

These limitations are demonstrated by the difficulty people have with perceiving raised-line drawings of objects by touching them. Pictures that are easily recognized by sight lead to low recognition rates when people must explore their contours with touch (e.g., Lederman *et al.*, 1990). The value of simultaneous exposure is indicated by the finding that visual object recognition falls to the same low level achieved by touch, when the visual displays are similarly exposed piecemeal (Loomis *et al.*, 1991). It might be argued that the problem lies with picking up the contour information. However, after failing to recognize a raised-line drawing from touch alone, a person may successfully recognize his or her drawing of it, indicating that the problem is not with feeling the contours but rather with putting them together (Wijntjes *et al.*, 2008).

The use of perspective in conveying 3D objects may be one source of difficulty in recognizing raised-line pictures by touch. Sighted people in particular have been found to recognize objects conveyed with minimal perspective cues (e.g., spoon seen from above) more effectively than fully 3D depictions (Lederman *et al.*, 1990). However, this difference did not arise in congenitally blind observers, suggesting a difference in the use of visual-imagery mediation in sighted and blind.

A raised-line drawing has the attractive feature that oriented line segments fall within the surface of the fingertip at any point in time. This segment stimulates sensory receptors with sufficient spatial acuity to allow accurate perception of the orientation of the segment beneath the finger. From this information, the perceiver can predict the continued direction of the line segment, and hence he or she knows where to move next in order to maintain the contour continuously under the finger. The guidance from fingertip location is so effective that people can simultaneously track two moving points delivered independently to fingertips of the two hands, tracing out completely different patterns (Rosenbaum *et al.*, 2006). This point raises a caveat, however, that we will consider further when discussing graphical displays for the blind. If vibration from a touch stimulates the whole fingertip, there is no gradient of stimulation across the skin to tell the user when the finger trajectory begins to veer away from the graphical data or when it reaches a change in a contour's direction. Another consideration is that a line rendered by vibration may have

to be wider than a raised line in tangible graphic displays, so that the user does not go past the trailing edge before he or she is able to stop. Preliminary research with line tracing tasks suggests that at least three times the standard line width of a hardcopy tactile embosser is needed to achieve equivalent performance with vibrotactile displays (Raja, 2011). Making all lines thicker means that the effective resolution of the display is reduced.

The demands on the user to track finger position when reading graphs on touchscreen displays should not be underestimated. In the absence of array stimulation on the fingertip to provide natural guidance along informative contours, considerable cognitive resources may be needed to maintain contact using whatever cues are provided by the interface. Moreover, the failure to maintain contact with a contour results in negative consequences. Not only is there transient loss of information and resulting memory load, but in addition, irrelevant searching movements may distort the kinesthetically based representation of the graph outline *per se*.

Systematic errors, such as shape distortions and illusions, may also be introduced by the distributed nature of touch encoding. For example, perceptual errors arising from temporally extended exploration may turn curves into straight lines (Sanders and Kappers, 2007) or non-parallel lines into parallel ones (Kappers, 2007); lines tangential to the body may seem shorter than the same rays traced radially (Cheng, 1968). A wide variety of such haptic illusions has been documented (for a review, see Lederman and Jones, 2011).

6. User Capabilities and Sensory Substitution

Developers of new devices aim to provide blind users with perceptual avenues to spatial information. It is common to find that, on the basis of their enthusiasm for this goal, they tend to focus on what the technology might offer rather than on what the user might be able to perceive and interpret. We have called this approach the 'engineering trap' (Giudice and Legge, 2008; Loomis *et al.*, 2012). User-centered design principles and consideration of basic human processing constraints enter relatively late into development, if they are present at all.

Loomis *et al.* (2012) pointed out that the effectiveness of sensory substitution for the blind is undermined when developers neglect the human user; more specifically, the limits imposed at both peripheral/sensory and central/cognitive levels, reviewed in detail in the previous section. While extensive practice can lead to improvements in performance (e.g., Sampaio *et al.*, 2001), perhaps overcoming some of the central limitations, the limitations of the spatio-temporal bandwidth of the cutaneous sense ultimately determine what can be achieved.

An example in the domain of current interest, 2D displays, is provided by early efforts to deliver video information to pixelated vibratory or electro-tactile displays on the back or abdomen (Bach-y-Rita, 1967, 1972; Collins, 1970; Collins and Madey, 1974). These early displays failed to allow the perception of 2D pattern information of any complexity; for example, block capital letters filling the display were recognized with an accuracy of only 35% (Apkarian-Stielau and Loomis, 1975; Loomis, 1981). More recent devices, like that based on an electronic tongue display, appear to afford slightly better performance (Bach-y-Rita *et al.*, 1998; Bach-y-Rita and Kercel, 2003; Chebat *et al.*, 2007; Sampaio *et al.*, 2001), and research with such displays indicates that cortical areas associated with vision can be activated (e.g., Ptito and Kupers, 2005; Ptito *et al.*, 2005). Even so, the still limited performance can be traced to constraints on spatio-temporal processing.

When sensory-substitution aids for the blind are developed, it is important to consider not only the device *per se* but also the task for which it is intended. A tactile display on the abdomen or tongue in conjunction with the haptic perceptual system might suffice, for example, to trigger attention or to allow travel along paths specified by bright edge markers against a dark background (Chebat *et al.*, 2011; Collins, 1985; Jansson, 1983; Segond *et al.*, 2005). To the extent that a task calls for spatial resolution near that offered by current touch-screen displays, it could not be afforded by a tactile display on these parts of the body (Loomis *et al.*, 2012).

To avoid being trapped by the elegance of technology at the expense of the end user, we advocate that user needs be assessed before and, iteratively, during research and development of new projects. It is important to determine whether the end product will be functional not just in terms of the technology but in the context of the user's capabilities and preferences. Failure to understand these factors will inevitably reduce the functional utility of the device and its likelihood of being adopted by the intended end-users. With these caveats in mind, we next consider sensory substitution in the specific context of touch-screen displays, by means of the auditory modality.

7. Auditory Augmentation of Touch-Based Displays

The limitations of haptic perception with touch-screen displays suggest the possibility of augmentation through other senses that could provide tracking cues to compensate for the smoothness of the physical surface and more directly convey information content. For the blind user, the obvious augmentation modality is audition. For example, a semantically rich auditory label may enhance or replace otherwise coarse information provided by a tactile icon. Our interest here, however, is primarily in augmenting spatial content rather than semantics. As we have pointed to the importance of understanding the pe-

ripheral limitations of any sense used for substitution, we will briefly provide some basic information about audition as a spatial sense (see Blauert, 1997, for review).

The primary cues to directional spatial hearing are binaural: the difference in intensity and time of arrival to the two ears. These cues pertain to the angular position of sound in the ear plane but not its elevation or distance. The distance of a sound source is primarily conveyed by the ratio of the intensity of the direct sound from the source to the intensity of the reverberant sound (Zahorik *et al.*, 2005). Filtering of sound occurs (changing the frequency mix) as it travels past the head and into the auditory canal, and therefore the spectral properties of sound are cues to spatial location as well. We should not forget, however, that sound richly conveys objects and events: Cows moo, bottles break, bells ring, and so on.

One way to add the auditory modality to a 2D screen display is to use touch for querying content at a particular location and to use sound with some spatial quality to tell the user where that location is. A caveat, however: Given that sound is emitted by speakers in the same device that senses touch, its spatial content is not easily manipulated. In particular, the binaural cues to sound localization are of relatively little value, unless the user wears stereo earphones with output coupled to hand position. Without such augmentation, this leaves the frequency and amplitude components of sound as the basic means of mapping from the planar space of the display to spatial auditory cues. With multiple sounds, relative timing and pitch can be used as well. The next question is how to map between 2D space and these cues.

A common, and perhaps intuitive, approach to this mapping is to represent the vertical axis by sound frequency (and its perceptual correlate, pitch). This approach is supported by a common association between pitch (i.e., frequency) and vertical position in the visual field (Pratt, 1930; Walker and Manuey, 2010, but see Eitan and Timmers, 2010, for non-generality of this association).

A proponent of the pitch/vertical association for sensory substitution is Meijer (1992), who has used it as the basis for a device (The vOICe) that substitutes hearing for vision. Following his lead, other researchers have further explored the design and its variants (Auvray *et al.*, 2005; Brown *et al.*, 2011; Cronly-Dillon *et al.*, 2000; Kramer, 1994). The basic design consists of a camera, a visual-to-auditory converter, and headphones. The camera-based input obviates the need for any active exploration with the hand, although head movements may be used. The display sweeps left to right over time, representing the brightness of pixels at a given pictorial height by amplitude at the corresponding frequency. The device has been found useful for displays of some complexity. Auvray *et al.* (2007) showed that it enabled listeners to point at a target with an average error less than 10°, and that after a few minutes of training with each of 10 objects, users could name them from their sound

signature and discriminate them from similarly trained, same-category exemplars. Another variant on image-to-sound mapping, called EyeMusic, has been demonstrated to enable rapid reaching movements similar to those performed under visual guidance (Levy-Tzedek *et al.*, 2012). More generally, promise for this approach is indicated by the finding of measurable cortical changes after training on mapping from visual height and brightness to sound (Pollok *et al.*, 2005; Renier *et al.*, 2005). The vOICe has also been shown to activate an area specialized for visual word representations during processing of letters by the blind (Striem-Amit *et al.*, 2012). However, when the goal is to render the contents of a 2D touch-screen display, it remains an open question as to the level of functionality that a passive time-frequency coupling of this sort could support.

As was noted above, the auditory channel can also be used to convey the sounds emitted by objects, such as vocalizations or percussive impact events. When used as symbols for content, these sounds (as well as beeps, buzzes, etc.) are referred to as 'earcons' or 'auditory icons' — the auditory equivalent of visual icons (Blattner *et al.*, 1989; Gaver, 1986). While earcons and speech have common use in computer interfaces, their lack of intrinsic spatial content brings them outside the scope of this article.

8. Experimental Data from a Multi-Modal Touch-Screen Graphical Display

The relatively recent emergence of touch-screen displays and their novelty as non-visual interfaces for the blind has had the consequence that relatively little empirical research has addressed their efficacy for communicating fundamental graphic content. Poppinga and colleagues (2011) found that sighted, blindfolded participants could use auditory labels and vibratory cues to encode a map on a smartphone screen. A touch-screen display with vibrators was shown to allow users to understand simple graphics (Goncu and Marriott, 2011; Petit *et al.*, 2008). The TeslaTouch, an electrostatic device, has also been offered as a display for the blind (Xu *et al.*, 2011). There has been little systematic analysis of the capabilities and limitations offered by these devices, however.

Above we have referred to the demands on the non-sighted user of tracking data locations on graphics displays, and the considerably increased difficulty when the display intrinsically offers no differentiating tangible features — as is true of touch-screens. We conducted a study specifically intended to investigate whether users without vision could track simple graphical displays (consisting of one or more lines) based on position-coupled vibration alone, and whether auditory position information enhanced their capabilities and/or liking of the display. Accordingly, we developed what we term a vibro-audio

interface (Giudice *et al.*, 2012), designed to facilitate tracking and information extraction.

8.1. Participants

Eighteen sighted, blindfolded participants (nine female), ages 18–27 ($M =$ 20.9, SD $= 2.8$), took part for monetary compensation, with informed consent.

8.2. Apparatus and Stimuli

Stimuli were presented on a Samsung Galaxy Tablet with a 7.0 in (17.8 cm) touchscreen running Android OS version 3.2, Target version 13. A pair of Sennheiser USA HD 201 circumaural ear headphones was used to present audio from the device. The display presented three types of vibratory stimuli: continuous lines, dashed lines, and information points. The continuous line was 0.35 in (0.9 cm) wide and was rendered as a vibration based on the UHL effect 'Engine_100', which repeats at 250 Hz at full power. The dashed line was a slowly pulsing vibration based on the UHL effect 'ENGINE3', which uses an infinite repeating loop at 143 Hz with full power. The information points were 0.35 in (0.9 cm) diameter circles that when contacted, vibrated with the UHL effect 'Weapon_1', a wide-band 0.01 s pulse with a 50% duty cycle and a 0.02 s period.

Auditory analogues of these signals were also constructed. The line stimuli consisted of a position/frequency coupling. The vertical (y) axis, corresponding to the shorter dimension of the tablet, was divided into 600 rows of pixels (the screen range), and the frequency at each row was then linearly varied between 220 Hz at the lowest and 660 Hz at the highest. Auditory pulsing of dotted lines occurred with the same duty cycle as for vibration. The information point signal was a series of clicks based on the vibratory analogue. For all stimuli, the amplitude was clearly audible for all users.

Two categories of line graphs with data points were presented: One was a single zigzag line with either four or five changes of direction. Seven information points were located at the endpoints, vertices, or, in the four-change graphs, midpoint of longer segments. By tapping on a vertex, the user could hear its horizontal-axis coordinate spoken. The second category of data presented a graph with one continuous and one dashed line, each having five information points evenly spaced along it (including the endpoints). The lines might or might not intersect. The information points did not display the x-coordinate.

8.3. Design and Procedure

The design was within-participant with a single factor of display type: vibratory only, auditory only, or bimodal. The user's hand was placed on the

highest leftmost position in the graph, and he or she then explored with a single finger while receiving a position-dependent auditory and/or vibratory signal, as described above, to indicate line, information point, and (when available) spoken coordinate. Each display type was implemented in a block with three exemplars of four graph types: four- and five-bend zigzag figures and two-line graphs with and without intersections. Each graph was unique, but variations were matched for complexity. Block orders were counterbalanced across participants, and the 12 graphs randomized within each block. A practice trial preceded each block, using a figure consisting of two lines, each with only three information points and one line having a single bend.

The user began the trial by freely exploring a graph. The interval was timed by the experimenter with the user's knowledge, and location of contact with the screen was recorded at 10 Hz. Then the user answered a series of three questions. The questions for the zigzag figures were: How many bends (turns) are present; where is the minimum y-axis value, and where is the maximum y-axis value? The three questions for the line graphs were: Which line has the steepest slope, at which point are these lines furthest apart on the y-axis, and do these lines intersect?

In addition, participants sketched each display after exploration, with blindfold removed. The sketch maps were analyzed by the Gardony Map Drawing Analyzer (Gardony *et al.*, 2013), which performs a bi-dimensional regression comparing the sketch to the original data and produces an index of distortion that is a combined measure of scale, rotation, and $x-y$ error. An exemplar stimulus and superimposed sketch are shown in Fig. 1. After all trials were completed, participants filled out a five-item survey to assess their preference for the interface modality, as follows: Which interface (i) did you prefer? (ii) did you feel you performed best? (iii) best provided the information? (iv) was the easiest to use? (v) would you use, given the proper scenario?
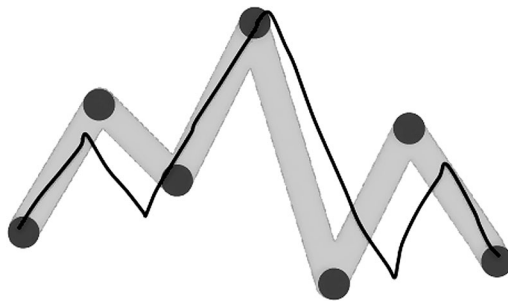

Figure 1. A five-bend zigzag and one participant's sketch, superimposed so that endpoints match.

8.4. Results

The results of a one-way ANOVA across the three display types are shown in Table 1 for accuracy of the question responses (based on number of correctly answered questions out of three), mean time for exploring and interpreting the graph, and the subsequent display sketching time. All were statistically equivalent across the displays. An ANOVA on the distortion index with factors block, display, and figure type showed only that distortion was reliably less for the two-line than the zigzag figures, means 12.23 *vs.* 20.25, $F(1, 17) = 41.66$, $p < 0.001$ (F-test for display shown in Table 1). From the finger tracking data collected as participants explored the figures, the percentage of contact time that the finger was on a graph element could be computed. This measure showed similar tracking performance across conditions, with the percentage of contact time spent touching a graph element of 58.2% for the vibration condition, 54.3% for auditory, and 55.6% for bimodal. These means did not differ significantly by a one-way ANOVA, $F(2, 17) = 2.92$, $p = 0.056$.

We conclude from these results that for all display conditions, the process of extracting the data from the touch-screen was quite slow, with considerable time spent off the target lines. It is possible that in a task with instructions to explore rapidly, time differences between displays might have been demonstrated. Nevertheless, the high accuracy level in all conditions indicates that given sufficient time to explore, users could extract graphical data sufficiently for the questions asked.

Although the interface modality did not affect the objective results, the first survey question revealed a clear preference for the use of auditory information for tracking, whether alone (8/18 preference choices) or combined with vibration (9/18). When asked which interface was most informative or would be the user's choice, not a single participant chose vibration alone. Vibration was preferred for ease of use by only three participants, compared to the preference for bimodal by seven and auditory alone by eight.

Table 1.
Performance on four measures by display type, with F statistic and p value for the effect of the display factor

	Question accuracy (3 max)	Exploration/ Interpretation time (s)	Sketch time (s)	Distortion index
Auditory/Vibratory	2.71 (0.57)	271.7 (129.1)	46.9 (34.6)	16.14 (10.4)
Vibratory	2.74 (0.48)	253.5 (131.7)	46.3 (24.9)	16.86 (11.1)
Bimodal	2.85 (0.43)	267.0 (135.2)	45.9 (25.7)	15.71 (8.6)
F-test (df = 2, 17)	1.59, $p = 0.21$	0.37, $p = 0.69$	0.02, $p = 0.98$	0.29, $p = 0.75$

It is important to note, in interpreting the preference for auditory input, that the auditory display was more informative than the vibration alone, since frequency coupling was added to kinesthesis in order to indicate the vertical position of the arm. This information was not specified with the vibrotactile interface, which had kinesthesis alone. We cannot be sure from these data whether the auditory frequency-height coupling facilitated tracking, information extraction, or both. The failure to find a significant difference between conditions in exploration time or percentage of time on target suggests that to the extent that sound facilitated tracking the lines in the graph, it did not profoundly affect the nature of exploration. Moreover, the reconstructions done from memory were equivalent in similarity to the originals across displays, suggesting the development of highly similar spatial representations, despite any differences in preference. (Further results with another vibro-audio interface can be found in Giudice *et al.*, 2012.)

9. Concluding Comments

We began this article by highlighting the promise of new touch-screen technologies for addressing the long-standing problem of providing graphical information to users without sight. Our review clearly points to needs for further research. While considerable empirical research on accessible graphical displays has been directed toward design and user preferences at the device level (MacLean, 2008; Nees and Walker, 2009), or conversely, the optimization of the contents of the display (Hoggan *et al.*, 2008; Walker, 2002), relatively little attention has been directed toward the processes by which information is accessed, extracted, integrated, and used. Here we emphasized the need to understand fundamental capabilities and limitations of the haptic system and other modalities that might be brought in to augment the sense of touch, particularly when screen surfaces offer no intrinsic tactile features.

There is clear momentum to direct new technologies toward expanding understanding and training on graphics among the blind community. Although we emphasize here that scientific understanding of fundamental sensory and cognitive processing is critical for the success of these laudable developments, it is also necessary to consider other limitations that frequently plague projects in assistive technology more generally. One is the tendency toward developmental as opposed to translational research; assistive projects often stop at the prototype stage rather than undergoing the final push toward commercialization. When products are commercialized, there is often little consideration of whether the implementing platform has the potential for multi-purposing, which can greatly mitigate the expense of assistive technology and reduce the need for learning multiple interfaces. Touch-screens are particularly attractive displays in this regard, as they can be implemented on multi-purpose devices

such as digital readers, smartphones, or tablets — all of which include various accessibility and universal design features in the native interface.

Finally, although technical researchers are rewarded for describing the technological advances of their designs and algorithms, there is fundamental and ongoing need for rigorous evaluation, focusing not only on system functionality but also on user-directed design. Development with the user in mind must ensure not only that a device can be made to deliver necessary information, but also that it is aesthetically acceptable and does not have such a steep learning curve that it limits end-user adoption (Giudice and Legge, 2008). For instance, the long cane has become the gold standard of blind mobility, as it is inexpensive, easy to use, and solves a real problem — the same principles should be applied to the development of non-visual graphical displays. Similarly, the braille alphabet and text-to-speech-based screen readers have succeeded owing to their information richness and broad access to textual information content. Emerging technologies for graphical access should build on this tradition and offer something not only new, but measurably better than what preceded them.

References

Apkarian-Stielau, P. M. and Loomis, J. M. (1975). A comparison of tactile and blurred visual form perception, *Percept. Psychophys.* **18**, 362–368.

Auvray, M., Hanneton, S., Lenay, C. and O'Regan, K. (2005). There is something out there: distal attribution in sensory substitution, twenty years later, *J. Integr. Neurosci.* **4**, 505–521.

Auvray, M., Hanneton, S. and O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: localisation and object recognition with 'the vOICe', *Perception* **36**, 416–430.

Bach-y-Rita, P. (1967). Sensory plasticity: applications to a vision substitution system, *Acta Neurol. Scand.* **43**, 417–426.

Bach-y-Rita, P. (1972). *Brain Mechanisms in Sensory Substitution*. Academic Press, New York, NY, USA.

Bach-y-Rita, P., Kaczmarek, K. A., Tyler, M. E. and Garcia-Lara, J. (1998). Form perception with a 49-point electrotactile stimulus array on the tongue: a technical note, *J. Rehabil. Res. Dev.* **35**, 427–430.

Bach-y-Rita, P. and Kercel, S. (2003). Sensory substitution and the human machine interface, *Trends Cogn. Sci.* **7**, 541–546.

Bau, O., Poupyrev, I., Israr, A. and Harrison, C. (2010). TeslaTouch: electrovibration for touch surfaces, in: *Proc. 23rd Ann. ACM Symp. User interface Software and Technology*, pp. 283–292. ACM, New York, NY, USA.

Berla, E. (1973). Strategies in scanning a tactual pseudo-map, *Educ. Vis. Handicap.* **5**, 8–19.

Blattner, M. M., Sumikawa, D. A. and Greenberg, R. M. (1989). Earcons and icons: their structure and common design principles, *Hum. Comput. Interact.* **4**, 11–44.

Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, MA, USA.

Brown, D., Macpherson, T. and Ward, J. (2011). Seeing with sound? Exploring different characteristics of a visual-to-auditory sensory substitution device, *Perception* **40**, 1120–1135.

Carpenter, P. A. and Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension, *J. Exp. Psychol. Appl.* **4**, 75–100.

Chebat, D. R., Rainville, C., Kupers, R. and Ptito, M. (2007). Tactile–'visual' acuity of the tongue in early blind individuals, *Neuroreport* **18**, 1901–1904.

Chebat, D. R., Schneider, F. C., Kupers, R. and Ptito, M. (2011). Navigation with a sensory substitution device in congenitally blind individuals, *Neuroreport* **22**, 342–347.

Cheng, M. F. (1968). Tactile–kinesthetic perception of length, *Am. J. Psychol.* **81**, 74–82.

Collins, C. C. (1970). Tactile television: mechanical and electrical image projection, *IEEE Trans. Man–Mach. Syst.* **MMS-11**, 65–71.

Collins, C. C. (1985). On mobility aids for the blind, in: *Electronic Spatial Sensing for the Blind*, D. H. Warren and E. R. Strelow (Eds), pp. 35–64. Martinus Nijhoff, Boston, MA, USA.

Collins, C. C. and Madey, J. M. (1974). Tactile sensory replacement, *Proc. San Diego Biomed. Symp.* **13**, 15–26.

Craig, J. C. (1999). Grating orientation as a measure of tactile spatial acuity, *Somatosens. Mot. Res.* **16**, 197–206.

Cronly-Dillon, J., Persaud, K. C. and Blore, R. (2000). Blind subjects construct conscious mental images of visual scenes encoded in musical form, *Proc. R. Soc. Lond. B Biol. Sci.* **267**, 2231–2238.

Eitan, Z. and Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: cross-domain mappings of auditory pitch in a musical context, *Cognition* **114**, 405–422.

Ferres, L., Lindgaard, G. and Sumegi, L. (2010). Evaluating a tool for improving accessibility to charts and graphs, in: *Proc. 12th Int. ACM SIGACCESS Conf. Computers and Accessibility* (*ASSETS '10*), pp. 83–90. ACM, New York, NY, USA.

Franks, F. and Baird, R. (1971). Geographical concepts and the visually handicapped, *Except. Child.* **38**, 321–324.

Galesic, M. and Garcia-Retamero, R. (2011). Graph literacy: a cross-cultural comparison, *Med. Decis. Making* **31**, 444–457.

Gardony, A. L., Brunye, T. T. and Taylor, H. A. (2013). Gardony map drawing analyzer: a novel approach to the analysis of sketch maps, poster at the Annual Convention of the Association for Psychology Science, Washington, D.C., May 25. http://www.aarongardony.com/tools/

Gaver, W. W. (1986). Auditory icons: using sound in computer interfaces, *Hum. Comp. Interact.* **2**, 167–177.

Giudice, N. A. and Legge, G. E. (2008). Blind navigation and the role of technology, in: *Engineering Handbook of Smart Technology for Aging, Disability and Independence*, A. Helal,

M. Mokhtari and B. Abdulrazak (Eds), pp. 479–500. John Wiley and Sons, New York, NY, USA.

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: *Proc. 14th Int. ACM SIGACCESS Conf. Computers and Accessibility* (ASSETS '12), pp. 103–110. ACM, New York, NY, USA.

Golledge, R. G. (1991). Tactual strip maps as navigational aids, *J. Vis. Impair. Blind.* **85**, 296–301.

Goncu, C. and Marriott, K. (2011). GraVVITAS: generic multi-touch presentation of accessible graphics, *Lect. Notes Comput. Sci.* **6946**, 30–48.

Hoggan, E., Brewster, S. A. and Johnston, J. (2008). Investigating the effectiveness of tactile feedback for mobile touchscreens, in: *Proc. SIGCHI Conf Human Factors in Computing Systems (CHI '08)*, pp. 1573–1582. ACM, New York, NY, USA.

Jansson, G. (1983). Tactile guidance of movement, *Int. J. Neurosci.* **19**, 37–46.

Johansson, R. S. and Vallbo, A. B. (1983). Tactile sensory coding in the glabrous skin of the human hand, *Trends Neurosci.* **6**, 27–32.

Kappers, A. M. L. (2007). Haptic spatial processing: allocentric and egocentric reference frames, *Can. J. Exp. Psychol.* **61**, 208–218.

Kaye, H. S., Kang, T. and LaPlante, M. P. (2000). Mobility device use in the United States. *Disability Statistics Report* (14). Dept. of Education, National Institute on Disability and Rehabilitation Research, Washington, D.C., USA.

Klatzky, R. L. and Lederman, S. J. (2003). Representing spatial location and layout from sparse kinesthetic contacts, *J. Exp. Psychol.: Hum. Percept. Perform.* **29**, 310–325.

Kramer, G. (1994). An introduction to auditory display, in: *Auditory Display*, G. Kramer (Ed.), pp. 1–78. Addison-Wesley, Reading, MA, USA.

Lederman, S. J. and Campbell, J. I. (1982). Tangible graphs for the blind, *Hum. Fact.* **24**, 85–100.

Lederman, S. J. and Jones, L. A. (2011). Tactile and haptic illusions, *IEEE Trans. Haptics* **4**, 273–294.

Lederman, S. J. and Klatzky, R. L. (2009). Haptic perception: a tutorial, *Atten. Percept. Psychophys.* **71**, 1439–1459.

Lederman, S. J., Klatzky, R. L., Chataway, C. and Summers, C. (1990). Visual mediation and the haptic recognition of two-dimensional pictures of common objects, *Percept. Psychophys.* **47**, 54–64.

Legge, G. E., Madison, C., Vaughn, B. N., Cheong, A. M. Y. and Miller, J. C. (2008). Retention of high tactile acuity throughout the life span in blindness, *Percept. Psychophys.* **70**, 1471–1488.

Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S. and Amedi, A. (2012). Fast, accurate reaching movements with a visual-to-auditory sensory substitution device, *Restor. Neurol. Neurosci.* **30**, 313–323.

Loomis, J. L., Klatzky, R. L. and Giudice, N. A. (2012). Sensory substitution of vision: importance of perceptual and cognitive processing, in: *Assistive Technology for Blindness and Low Vision*, R. Manduchi and S. Kurniawan (Eds), pp. 162–191. CRC Press, Boca Raton, FL, USA.

Loomis, J. M. (1981). Tactile pattern perception, *Perception* **10**, 5–27.

Loomis, J. M. (1990). A model of character recognition and legibility, *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 106–120.

Loomis, J. M., Klatzky, R. L. and Lederman, S. J. (1991). Similarity of tactual and visual picture recognition with limited field of view, *Perception* **20**, 167–177.

MacLean, K. E. (2008). Haptic interaction design for everyday interfaces, in: *Reviews of Human Factors and Ergonomics*, M. Carswell (Ed.), Vol. 4, pp. 149–194. Human Factors and Ergonomics Society, Santa Monica, CA, USA.

Meijer, P. B. L. (1992). An experimental system for auditory image representations, *IEEE Tran. Biomed. Eng.* **39**, 112–121.

Merriam, E. P., Genovese, C. R. and Colby, C. L. (2007). Remapping in human visual cortex, *J. Neurophysiol.* **97**, 1738–1755.

Millar, S. (1981). Self-referent and movement cues in coding spatial location by blind and sighted children, *Perception* **10**, 255–264.

Millar, S. (2008). *Space and Sense*. Psychology Press, New York, NY, USA.

Nees, M. A. and Walker, B. N. (2009). Auditory interfaces and sonification, in: *The Universal Access Handbook*, C. Stephanidis (Ed.), pp. 507–521. CRC Press, New York, NY, USA.

Nolan, C. Y. and Morris, J. E. (1971). *Improvement of Tactual Symbols for Blind Children: Final Report*, American Printing House for the Blind, Louisville, KY, USA.

Perkins Maps (2013). http://www.perkins.org/resources/scout/education/geography/tactile-maps.html

Perkins Museum (2013). http://www.perkinsmuseum.org/section.php?id=188

Petit, G., Dufresne, A., Levesque, V., Hayward, V. and Trudeau, N. (2008). Refreshable tactile graphics applied to schoolbook illustrations for students with visual impairment, in: *Proc. 10th Int. ACM SIGACCESS Conf. Computers and Accessibility* (ASSETS '08), pp. 89–96. ACM, New York, NY, USA.

Pollok, B., Schnitzler, I., Mierdorf, T., Stoerig, P. and Schnitzler, A. (2005). Image-to-sound conversion: experience-induced plasticity in auditory cortex of blindfolded adults, *Exp. Brain Res.* **167**, 287–291.

Poppinga, B., Magnusson, C., Pielot, M. and Rassmus-Gröhn, K. (2011). TouchOver map: audio-tactile exploration of interactive maps, in: *Proc. 13th Int. Conf. Human Computer Interaction with Mobile Devices and Services* (*MobileHCI '11*), pp. 545–550. ACM, New York, NY, USA.

Pratt, C. C. (1930). The spatial character of high and low tones, *J. Exp. Psychol.* **13**, 278–285.

Ptito, M. and Kupers, R. (2005). Cross-modal plasticity in early blindness, *J. Integr. Neurosci.* **4**, 479–488.

Ptito, M., Moesgaard, S. M., Gjedde, A. and Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind, *Brain* **128**, 606–614.

Raja, M. K. (2011). The development and validation of a new smartphone based non-visual spatial interface for learning indoor layouts, *Unpublished Masters thesis*, University of Maine, Orono, MN, USA.

Rastogi, R., Pawluk, T. V. D. and Ketchum, J. (2013). Intuitive tactile zooming for graphics accessed by individuals who are blind and visually impaired, *IEEE Trans. Neural Syst. Rehabil. Eng.* **21**, 655–663.

Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., Veraart, C. and De Volder, A. G. (2005). Cross-modal activation of visual cortex during depth perception using auditory substitution of vision, *NeuroImage* **26**, 573–580.

Rosenbaum, D. A., Dawson, A. M. and Challis, J. H. (2006). Haptic tracking permits bimanual independence, *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 1266–1275.

Sampaio, E., Maris, S. and Bach-y-Rita, P. (2001). Brain plasticity: 'visual' acuity of blind persons via the tongue, *Brain Res.* **908**, 204–207.

Sanders, A. F. J. and Kappers, A. M. L. (2007). Haptically straight lines, *Perception* **36**, 1682–1697.

Segond, H., Weiss, D. and Sampiao, E. (2005). Human spatial navigation via a visuo-tactile sensory substitution system, *Perception* **34**, 1231–1249.

Striem-Amit, E., Dehaene, S., Cohen, L. and Amedi, A. (2012). Reading with sounds: sensory substitution selectively activates the visual word form area in the blind, *Neuron* **76**, 640–652.

Tversky, B. (2003). Some ways graphics communicate, in: *Mobile Communication: Essays on Cognition and Community*, K. Nyiri (Ed.), pp. 143–156. Passagen Verlag, Wien, Austria.

Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification, *J. Exp. Psychol. Appl.* **8**, 211–221.

Walker, B. N. and Mauney, L. (2010). Universal design of auditory graphs: a comparison of sonification mappings for visually impaired and sighted listeners, *ACM Trans. Access. Comput.* **2**, Article 12, 16 pages.

Wijntjes, M. W. A., van Lienen, T., Verstijnen, I. M. and Kappers, A. M. L. (2008). Look what I have felt: unidentified haptic line drawings are identified after sketching, *Acta Psychol.* **128**, 255–263.

Winfield, L., Glassmire, J., Colgate, J. E. and Peshkin, M. (2007). T-PaD: tactile pattern display through variable friction reduction, in: *Proc. 2nd Joint EuroHaptics Conf. Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 421–426. IEEE, New York, NY, USA.

World Health Organization (2011). Visual Impairment and Blindness Fact Sheet. http://www.who.int/mediacentre/factsheets/fs282/en/

Xu, C., Israr, A., Poupyrev, I., Bau, O. and Harrison, C. (2011). Tactile display for the visually impaired using TeslaTouch, in: *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, pp. 317–322. ACM, New York, NY, USA.

Yu, W. and Brewster, S. A. (2002a). Comparing two haptic interfaces for multimodal graph rendering, in: *Proc. 10th Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 3–9. IEEE, New York, NY, USA.

Yu, W. and Brewster, S. A. (2002b). Evaluation of multimodal graphs for blind people, *Univers. Access Inform. Soc.* **2**, 105–124.

Zacks, J. and Tversky, B. (1999). Bars and lines: a study of graphic communication, *Mem. Cogn.* **27**, 1073–1079.

Zahorik, P., Brungart, D. S. and Bronkhorst, A. W. (2005). Auditory distance perception in humans: a summary of past and present research, *Acta Acustica United w. Acustica* **91**, 409–420.