SUPPORTING LEARNING FOR INDIVIDUALS WITH VISUAL IMPAIRMENT

⋆Francis Quek, ⋆Yasmine Elglaly, and †Francisco Oliveira

⋆Virginia Tech, USA, †Ceará State University, Brazil
⋆quek@vt.edu

ABSTRACT

This paper presents an *embodied perspective* on the development of multimodal and perceptual replacement technologies for *Individuals with Blindness or Severe Visual Impairment* (IBSVI). The premise is that much of the difficulties encountered by IBSVI comes from the fact that our cultural world and means of information exchange and design are inextricably associated with the capabilities of embodied humans endowed with spatial vision. We present a pair of projects associated with supporting mathematics instruction and learning to IBSVI. The first relates to the particulars of embodied human discourse where gesture allows dynamic fusion of spoken and visual instruction content. We developed a system that addresses this challenge using computer vision and haptic technology. In the second project, we address the problems encountered by IBSVI in accessing information through reading. Our e-Reader system allows IBSVI to engage in active reading at their own pace and control.

Index Terms— blindness, embodiment, learning

1. INTRODUCTION

Imagine if everyone else were endowed like Michael Jordan in his prime. What kind of a world would we live in? The entire built-up world and cultural expectations would be designed around 2 m tall individuals with tremendous dexterity. Stair steps would be high enough to require significant effort for most readers of this paper, shelves would be too high for us to reach, and even the dimensions of keyboards, phones, and books would concomitantly enlarged. Furthermore, inter-human interaction would be designed where we would expect to hand things to one another across larger distances (most of us would become klutzes, dropping things that are routinely tossed to one another). The authors and readers of this paper would, in short, be 'disabled' in this world.

The point of the previous paragraph is that we are *embodied beings*, and that the cultural world is designed for this embodiment. *Individuals with Blindness or Severe Visual Impairment* (IBSVI) are 'otherly embodied' along the dimension of visual perception. This paper advances the premise that many of the barriers faced by IBSVI arise, not only from the natural world, but from the cultural expectations designed into world we construct, and the means by which we communicate and inform. This insight is accompanied by implications for technology we develop for IBSVI. With respect to support for learning, these implications extend to understanding of the expectations built into the way we communicate concepts and design information. We shall explore our premise by looking at two systems designed to support learning by IBSVI. The first addresses the communicative deficits encountered by IBSVI in receiving mathematical/science instruction, and the second addresses the design of information and conveyance of knowledge itself.

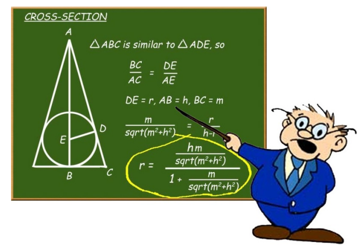

Fig. 1. Illustration of mathematics instruction

2. MATHEMATICS INSTRUCTION FOR IBSVI

The means by which we communicate mathematical concepts in instruction is illustrated in Figure 1 where three channels of communication are evident: 1. The vocal presentation by the instructor; 2. The graphic that carries the mathematical concepts being discussed; and 3. The pointing gesture that allows the instructor and student to share a focus into the illustration co-temporally with the vocal utterance. For IBSVI, the latter two channels require some form of augmentation. For line graphics, it is trivial to render the material as 'raised-line' drawings using such inexpensive technology as paper embossing [1] that are easily accessed by IBSVI [2].

2.1. Requirements for Pointing

To understand the requirements for conveying the pointing gestures of the instructor, we draw on the principles of gesture production, uptake, and communication. We highlight three key features of human use of gestures: 1) Time synchrony is critical for both production and uptake of deixis with respect to speech [3]. This means that the gestural information must be conveyed in real-time to allow the IBSVI to have

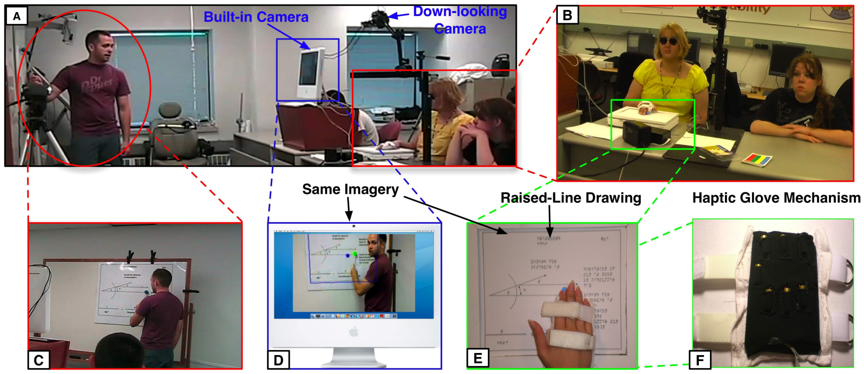

Fig. 2. The Haptic Deictic System – HDS

access to the referent of the pointing gesture simultaneously with the reception of speech. 2) The pointing gestures are performed in the process of speaking [4]. Since gesture and speech are co-produced [3], the instructor she does not think of the words to speak and then consciously think of what to point at to illustrate specific words. This means that pointing must be performed transparently while the speaker is thinking and speaking. 3) The student receives both gesture and speech during instruction. Uptake of deixis cannot be a laborious conscious process. Gesture uptake must be effortless in the process of receiving the multimodal content. Of course, this kind of exchange is lost to IBSVI, possibly accounting in a significant way to the difficulty of such individuals to advance in mathematics education (SBVI are typically one to three years behind their sighted counterparts [5]). Another overarching goal of our solution is to enable inclusive classrooms, where IBSVI attend mainstream classes. It has been argued that inclusive classrooms are beneficial for both disabled [6] and non-disabled students [7]. Furthermore, such inclusive instruction is required by law in the U.S. (Disabilities Education Act Amendments (IDEA, 1997), and the *No Child Left Behind Act* (NCLB, 2001)). A court ruling [8] reinforced the non-segregational approach.

2.2. The Haptic Deictic System – HDS

Our *Haptic Deictic System* (HDS) [9, 10] is designed to meet the aforementioned requirements. The HDS is summarized pictorially in Figure 2. Figure 2A shows a classroom scene with the instructor pointing into a graphic on a poster, and a pair of seated students (one IBSVI, and one sighted) receiving instruction. The instructor's pointing gestures (with a wand in the figure, but the system is capable of tracking an unadorned hand) are tracked via the camera in the iMac placed in front of him (Figure 2 C). Figure 2B shows a close-up picture of the two students from Figure 2A, with the IBSVI on the left reading an embossed raised-line version of the graphic on the poster. The down-looking camera visible in Figure 2A tracks the IBSVI's reading hand (a frame of this tracking video is shown in Figure 2E). In the *instructor's display* on the iMac monitor (see Figure 2D), the instructor can see the video stream from the tracking camera showing the location of his pointing focus (blue dot), and the reading location of the IBSVI (green dot). Since the instructor's graphic and the student's raised-line drawing are exact spatial facsimiles of each other apart from scale, it is trivial to compute the spatial relationship between the instructor's *Point of Instructional Focus* (PIF), and the student's point of *Tactile Point of Access* (TPA) in the student's raised-line drawing. If T_{IS} transforms a point from the instructor's graphic to the student's raised-line drawing, then we can compute the *Focal Disparity* for the student as $\mathrm{FD} = T_{IS}(\mathrm{PIF}) - \mathrm{TPA})$. Since the student is interacting physically with the raised line drawing, T_{IS} has to be updated dynamically to keep the two spaces registered.

The IBSVI wears a haptic glove embedded with eight mechanical actuators (vibrating motors sewn into a glove pad). The glove conveys the FD to the IBSVI to enable her to locate the instructor's deictic focus. All motors come to a complete stop to indicate the SBVI has reached the PIF. The technical details of the HDS design may be found in [11, 9].

2.3. Summary of Discourse Support Studies

The HDS is designed to support 'natural' speaking and pointing in an inclusive classroom situation where IBSVI learn alongside sighted students. The instructor would simply speak using a pre-prepared graphic for which raised-line versions are available to the IBSVI in the class. The instructor would be able to point at the graphic while engaging the class just as she would in an all-sighted classroom, except that the IBSVI in the class would follow the pointing using the HDS, and the instructor can monitor if the students are following the instruction. For the sighted students, she would employ the natural capacity for gaze awareness, and for the IBSVI, she would use the *instructor's display* in the iMac monitor.

We summarize the steps taken to realize the HDS and test its operation. The design of the haptic glove involved a number of design and test iterations to arrive at a design that was able to convey direction that the IBSVI can perceive in realtime. This entails a series of cycles of design, prototype building, and perception-action tests. Because IBSVI need a reference hand to anchor the reading, the glove was worn on the student's reading hand. Our design had to ensure: 1. Real-time perception of direction and navigation to the target; 2. That the glove use does not interfere with the student's fingertip reading of the raised-line drawing; and, 3. That the student had to be able to manage the cognitive load of listening to speech while navigating and reading.

Our studies showed that even when this level of perception was achieved, it still does not mean that the HDS can facilitate fluid discourse. Our studies with the system in a discourse context (a joint problem solving task) showed that 30% of discourse turns were dedicated to the technology and the act of pointing itself. This focus on the operational aspects of communication interfered with the functional goal of discourse. We determined that this is because the cognitive load of maintaining discourse as far above that of even listening, reading, and navigating in tandem. IBVSI require far more familiarity with the system before they could use it as part of a discourse-support system. Our solution was to develop a computer game to enable the users to gain the necessary 'embodied skill' [12]. The game was placed in the disabilities support center of the institute where the IBSVI were situated, and they could play the game at their leisure.

After our IBSVI participants engaged the game for an academic semester, we tested the system for discourse support again. This time, the discourse task was completed in a third of the time of our first study, and no discourse turns were expended on either the technology or the act of pointing. In essence, embodied skill enabled the technology to become transparent, and the IBSVI and their seeing partners were simply talking and problem solving.

We constructed a pair of 3-class session mathematics instruction mini-courses and distributed our IBSVI participants into constructed inclusive learning situations (an IBSVI with three sighted students), and conducted the classes in counterbalanced with- and without-HDS conditions.

2.4. HDS Results and Discussion

To assess the impact of the HDS on classroom dynamics and learning, we designed three evaluation approaches. Although the details of this study are beyond the scope of this paper, we summarize them here to aid in understanding of the benefits provided by the HDS. A detailed description may be found in [10]. First, we developed a pair of psycholinguistically-informed analysis approaches to determine if the HDS presented the IBSVI with the opportunity to learn. Second, we employed a battery of analyses to determine if the HDS impacted instructional discourse fluency. Third, we analyzed the experience of all the participants in the inclusive learning scenarios (the instructor, SBVI, and the sighted students).

Our psycholinguistically-grounded analyses indicated that the IBSVI in the with-HDS condition were afforded the 'opportunity to learn' by having access to the mathematical concepts being conveyed. Our studies showed that the IBSVI could combine the vocal utterances of the instructor with the co-temporal graphical focus through awareness of pointing . Our second test involved discourse analysis of the instructors' utterances to see how instructional style is mediated by the technology. We determined that the system allowed the instructor to employ more deictic references (e.g., 'this' and 'that'), and to be more economical in expression over against the non-HDS case where every referent had to be explicitly called out by name. There were also fewer class interruptions to single out the IBSVI for assistance.

Finally, the introduction of the HDS was found to be beneficial by all stakeholders in the inclusive classroom. The IBSVI had access to the material taught, and did not feel 'singled out' for special attention. The sighted students felt that the class flowed better, and was not unnecessarily impeded by the presence of the IBSVI. The instructors felt that the HDS helped them to pace the class better by giving them awareness of the IBSVI's attentional status.

With respect to our basic premise, the HDS project shows how an understanding of the communicative dynamics of embodied human discourse helped us to identify an understudied aspect of perceptual deficit encountered by IBSVI. This allows us to employ sensing technology (vision systems that track the deictic gestures of the instructor and the reading activity of the IBSVI), and output technology (the haptic glove, and the heads-up display for the instructor) in an effective manner for aid instruction for, and learning by IBSVI.

3. DESIGN OF E-READERS FOR IBSVI

Our second project is in the design of e-Readers for IBSVI. The underlying premise of this project is that books are information designed for embodied humans with sight. In fact, the very form of the information itself is organized for visual access. Walter Ong, in his highly influential work "Orality and Literacy: The Technologizing of the Word" [13] argued that before the invention of printing and widespread literacy, all information was designed for aural consumption. Alliteration, rhyme, and orally-oriented organization permeated all literature. After the advent of printing, information became increasingly organized with the expectation of visual access in its very fabric. When paper became the dominant form of storage, transfer, and representation of knowledge, spatial organization, prosaic form, and visual markings (e.g., varying fonts and typestyles) have become the coin of the realm for information organization.

3.1. Understanding the problem

To understand this, consider reading this paper without the ability to look back a paragraph or two, or even to skip back to the introduction to see what the fundamental premise of this paper is. The problem for IBSVI is that Braille, invented in 1824, is still the only information format that supports spatial reading. Unfortunately Braille literacy is declining in the USA for various reasons, and the current Braille literacy level stands at 10% [14]. Various forms of refreshable Braille are limited to very small arrays of no more than eight lines of 40 characters each at maximum. Just as importantly, Braille is cumbersome. "Harry Potter and the Order of the Phoenix" is 870 pages in the paperback version, and the Braille it is thirteen 14" × 11" volumes that makes a two-foot high stack.

Currently, the most viable alternative to Braille is the audio book that renders text into audio streams. The IBSVI has limited control to play, pause, and rewind the audio, but do not have the ability to read and reread in the random order that is trivial for the sighted. Various solutions that translate printed books using OCR technology essentially proffer the same solution. They obliterate space, and render information as linear streams of ephemeral information borne by vibrating molecules of air. The problem is that the information was designed for visual access, and IBSVI is left with the task of maintaining all context, organization, and prior information in memory. While audio books may be suitable for leisure reading, few sighted people would entertain studying for an examination using an audio book [15]. This may explain why very few IBSVI advance through high school and beyond.

The introduction of slate-type devices and large form-factor smart phones threatens to widen the information divide for the IBSVI as information becomes more readily available to the sighted more rapidly. Conceptually, we address the spatial nature of informational media itself. Informational resources are not conceived or designed in a vacuum. They are optimized for consumption by embodied beings, and in this case, typically endowed with vision. An apropos analogy may be a world endowed with staircases designed because we are bipedal (as opposed to wheeled or arboreal) beings. Such a world is biased against individuals with paraplegia. The differences in the information/media world from this analogy are twofold. First, the spatial nature of the media bias is not as obvious, and so, not as well studied or understood. Consequently, the importance of visuo-spatiality for information access is undervalued. Second, unlike the staircase analogy, the solution is not as easy as building ramps between floors. The entire design of floors would be an impediment if the analogy were to hold. Individuals deprived of a visuo-spatial sense do not just have an input-output problem. The entire conceptualization of information is permeated with this spatial formulation. It is not just an information repackaging issue, because the repackaging would have to involve a complete information redesign. Rather, one has to think of providing individuals without vision with an alternate means to access space and not just information bits.

3.2. STAAR Solution Overview

Figure 3 pictures our *Situated Touch Audio Annotator and Reader* (STAAR) e-Reader designed to provide spatial access to textual information for IBSVI. Figure 3-left shows an iPad with text rendered in it. The STAAR reader can handle any PDF document. As the IBSVI moves her finger across the text, as she touches each word the reader sounds the word. The system tracks the finger, and speaks the word at the anticipated rate at which the IBSVI moves across a line of text

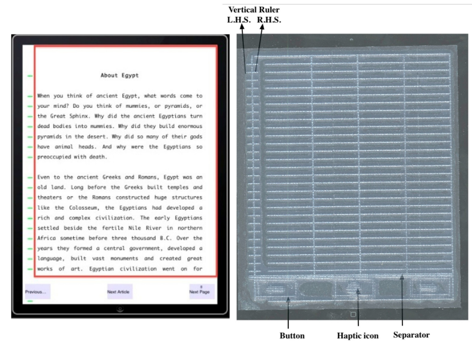

Fig. 3. STAAR Reader with touchable text and overlay

– in essence giving her control of her own reading rate. Figure 3 -right shows an embossed overlay with a tactile pattern that provides the IBSVI with landmarks and the haptic feedback needed to stay on line and to maintain place across the iPad [16]. A set of tactile buttons allows for a limited set of page and reading controls. The vertical lines allow the IBSVI to more rapidly find a position on a line, and the vertical ruler on the left side of the overlay allows the IBSVI to locate lines of text. A minimal set of sonifications are designed to provide the IBSVI with some awareness of page structure (e.g., where the white space is).

The STAAR system thus provides a spatial tactile landmark grid for textual reading, and an audio system that renders the information on the page by location of touch. The IBSVI is able to fuse both of these modes of sensing to gain an understanding of the structure of the page, and to read and reread at her own pace and under her control.

3.3. STAAR Design

Figure 4 provides a graphical overview of the subsystems that make up the STAAR e-Reader. The left of the figure shows the *Basic System Components* of the system architecture, page layout and description, and overlay design. The top right of the figure shows the design for *Intelligent Runtime Reading Support* that helps the IBSVI to stay on a horizontal line during reading. The bottom-left of the figure shows the *sub-systems for audio support*. While the detailed design is beyond the scope of this paper, we shall provide a high-level overview of these system components to elucidate the functioning of the e-Reader system, and how it addresses the requirements and provides the functionality discussed earlier.

3.3.1. Basic System Components

The system architecture that has four major subsystems that 1. Interprets standard PDF documents; 2. Models the page/document to be rendered spatially on the iPad; 3. Tracks the actions of the IBSVI on the touch surface of the iPad; and, 4. Produces the audio to render the page dynamically.

We include the basic document presentation and overlay designs in as *Basic System Components*. A series of careful studies with IBSVI participants allowed us to determine effective document scales and layout of the overlay [17]. After

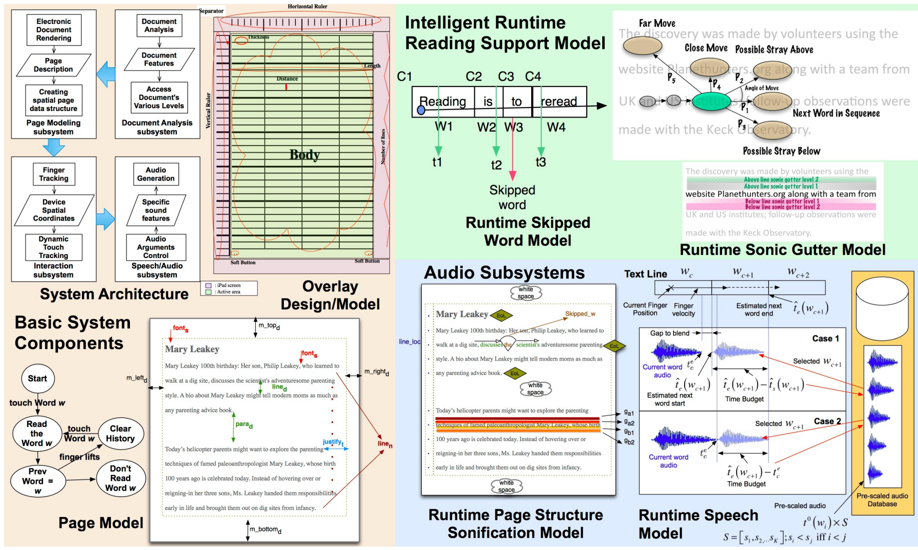

Fig. 4. Situated Touch Audio Annotator and Reader (STAAR)

some prototyping-testing iterations we determined that a material known as 'embossables' [18] with a grid produced by a standard high-end Braille embosser produced a usable overlay that was perceptible to the IBSVI and allowed the iPad touch surface to function. The key finding concerning the document scale and overlay layout to be related here is that IBSVI are able to use the tactile overlay as a landmarking grid even when there is no one-to-one correspondence between the grid lines and text lines. There is a requirement, however, that the grid be denser than the text lines.

3.3.2. Intelligent Runtime Reading Support

In a usability study with 16 IBSVI, we found a high occurrence to "Wandering Between Lines" incident despite the existence of the horizontal landmarks on the overlay [19]. Hence STAAR is designed with dynamic reading support features [20]. If the system determines that the reader is in the process of reading a line of text, the line-reading support is activated. This support comprises two components. First, to augment the tactile grid, a '*sonic gutter*' is activated. This is a kind of sound fence that produces a rustling sound when the reader strays off the center of the current text line. Two levels of audible signals are used on each side of the text for closer, and farther deviations. At the same time, a probabilistic decision system estimates if the user intended to stay on line, and is merely straying, or if she intended to move to a different line. In the first case, the system reads the 'intended next word' even if the user's finger had strayed onto an adjacent line (still within the range of the *sonic gutter*). The sonic gutter provides feedback that the deviation has occurred and allows the user to self-correct. If the system determines that the user intended to leave the line, or if she moves beyond the bound of the *sonic gutter*, the line reading support feature disengages and the user is free to explore the document. What this essentially does is 'fatten the line' that is currently being read to account for motoric uncertainty.

Two other reading support features are designed into STAAR. The first is the 'skipped word notification system' that produces a click if the reader moves so fast that the audio system cannot keep up [19]. Each word skipped results in an audible click to alert the IBSVI to the fact. The second other support is that STAAR estimates which is the intended 'reading finger' and only renders words touched by that finger. This allows the IBSVI to use a trailing touch point as a physical reference, or to touch the iPad surface with the heel of the palm while reading.

3.3.3. Audio Subsystem

The final component block in STAAR handles the production of audio. The runtime page structure sonification produces audible cues concerning page spatial structure. This is a minimal set of sonifications are designed to provide the IBSVI with spatial information on the page. These include whitespace (rustling sound when whitespace is touched), end-of-line (a typewrite-style 'ding' when the last word is read), skipped word (a click is heard if the IBSVI reads to quickly for the system to keep up, and a word is skipped), and line location (as the user moves her finger along the left vertical margin, a click is heard when the finger crosses a line of text – this allows the reader to know where to begin reading a line).

Finally, to enable self-paced reading, we had to implement a variable-speech-rate reading sub-system. This makes it impractical to use a text-to-speech (TTS) system directly because it is very hard to vary the vocalization duration of each word dynamically within a sentence or range words being rendered. To address this, we developed a system where TTS output is pre-sampled and pre-scaled for playback.

3.4. STAAR e-Reader Results and Discussion

The STAAR e-Reader was developed in collaboration with a cadre of IBSVI participants who served as our early testing participant designers and advisors [17]. The components described in Figure'4 were developed over multiple cycles of 'design–prototype–test with IBSVI–glean lessons learnt–redesign' cycles. As of this writing, we are in the process of analyzing our results for the *intelligent runtime reading support* subsystem with the *sonic gutter*. To give a sense of the functionality of the STAAR system, we summarize the results of an extended *Experience-Sampling Method* [21] study [19]. This was after the third cycle of our design loop. We seeded 7 IBSVI with the STAAR system to use for two weeks in their own homes and workplaces. Each participant was given 27 pages to read, and a set of reading and comprehension benchmark tests were administered at the beginning and end of the period. The results were very promising, with all readers completing the reading and able to answer questions on content and comprehension with minimal error. In fact three of the participants completed all the readings in seven days, and requested more. There was marked improvement over all similar benchmarks administered in the second cycle tests.

With respect to this paper's basic premise, the STAAR project shows how an understanding nature of information as being designed intrinsically for embodied/sighted individuals determined how we approached the problem. This allows us to employ modern slate technologies with multimedia and touch capabilities to address a fundamental problem encountered by IBSVI.

4. CONCLUSION

We have presented a basic embodiment premise that many of the impediments faced by IBSVI arise from the cultural expectations designed into our constructed world, and the means by which we communicate. This lens has allowed us to focus our development of multimodal and alternative perception technologies to support IBSVI in the task of learning. We believe that this perspective is critical for developers of such technologies to address real, and pressing problems for this important population.

5. ACKNOWLEDGEMENTS

This research has been supported by National Science Foundation grants IIS-0451843IIS-1117854 and IIS-1059398.

6. REFERENCES

[1] Universal Low Vision Aids Inc., "Emprint," http://www.ulva.com/Online-Store/Braille-Embossers/emprint.htm, 2009.

[2] S.A. Wall & S. Brewster, "Sensory substitution using tactile pin arrays: Human factors, technology and applications," *Sig. Proc.*, **86**:12, pp. 3674–3695, 2006.

[3] David McNeill, *Hand and Mind: What Gestures Reveal about thought*, U. of Chicago Press, 1992.

[4] S. Goldin-Meadow, *Hearing Gesture: How Our Hands Help Us Think*, Belknap Press, 2003.

[5] J. Williams, "Nationwide shortage of teachers for blind students must be corrected," *Nat'l Fed. of Blind: Advocates for Equality, Canadian Blind Monitor*, 2002.

[6] D. Thomas & E. Kubiak, "Issues and aids for teaching mathematics to the blind," *Math. Teacher*, **90**, 1997.

[7] D Staub and C Peck, "What are the outcomes for nondisabled students?," *Educational Leadership*, vol. 52, no. 4, pp. 36–40, 1994.

[8] Oberti v. Clementon, "Court decision," 1993, 995 F.2D 1204, (3rd Cir. 1993).

[9] F. Oliveira, F. Quek, et al, "The haptic deictic system - HDS: Bringing blind students to mainstream classrooms," *IEEE T. on Haptics*, **5**:3, pp. 172-183, 2012.

[10] F. Quek & F. Oliveira, "Enabling the blind to see gestures," *ACM T. of CHI*, **20**:1, to appear 2013.

[11] F. Oliveira & F. Quek, "A multimodal communication with a haptic glove: On the fusion of speech and deictic over a raised line drawing," in *1st Int'l Conf. PErvasive Tech.s Related to Assistive Environments*, 2008.

[12] F. Oliveira, et al, "Fun to develop embodied skill: how games help the blind to understand pointing," in *Proc. 3rd Int'l Conf. PErvasive Tech. Related to Assistive Environments*. ACM, 2010, pp. 1–8.

[13] Walter J. Ong, *Orality and Literacy: The Technologizing of the Word*, London: Methuen and Co., 1982.

[14] C. Stephanidis, *Universal Access in HCI. Apps & Services: 5th Int Conf UAHCI 2009*.

[15] F. Robinson, *Effective study*, Harper & Row, NY, 1970.

[16] Y. Elglaly, F. Quek, T. Smith-Jackson, & G. Dhillon, "Touch-screens are not tangible: Fusing tangible interaction with touch glass in readers for the blind," in *Int'l Conf. Tangible, Embedded & Embodied Int.*, 2013.

[17] G. Dhillon, et al, "Use of participatory design to enhance accessibility of slate-type devices," in *Proc. 56th Hum. Factors & Ergonomics*, 2012.

[18] American Thermoform Corporation, "Embossables," http://www.americanthermoform.com/brailon.htm.

[19] Y. El-Glaly, F. Quek, T. Smith-Jackson, and G. Dhillon, "Audible rendering of text documents controlled by multi-touch interaction," in *Proc. 14th ACM Int'l Conf. Multimodal Int.*. 2012, pp. 401–408.

[20] Y. El-Glaly, *Spatial reading system for individuals with blindness*, Ph.D. thesis, Virginia Tech., 2013.

[21] R. Larson & M. Csikszentmihalyi, "The experience sampling method.," *New Directions for Methodology of Soc. & Beh. Sci,*, 1983.