

# Evaluating Haptic and Auditory Directional Guidance to Assist Blind People in Reading Printed Text Using Finger-Mounted Cameras

LEE STEARNS, University of Maryland, College Park

RUOFEI DU, University of Maryland, College Park

URAN OH, University of Maryland, College Park

CATHERINE JOU, University of Maryland, College Park

LEAH FINDLATER, University of Maryland, College Park

DAVID A. ROSS, Atlanta VA R&D Center of Excellence in Vision and Neurocognitive Rehabilitation (CVNR)

JON E. FROELICH, University of Maryland, College Park

The recent miniaturization of cameras has enabled finger-based reading approaches that provide blind and visually impaired readers with access to printed materials. Compared to handheld text scanners such as mobile phone applications, mounting a tiny camera on the user's own finger has the potential to mitigate camera framing issues, enable a blind reader to better understand the spatial layout of a document, and provide better control over reading pace. A finger-based approach, however, also introduces the need to guide the reader in physically navigating a document, such as tracing along lines of text. While previous work has proposed audio and haptic directional finger guidance for this purpose, user studies of finger-based reading have not provided an in-depth performance analysis of the finger-based reading process. To further investigate the effectiveness of finger-based sensing and feedback for reading printed text, we conducted a controlled lab experiment with 19 blind participants, comparing audio and haptic directional finger guidance within an iPad-based testbed. As a small follow-up, we later asked 4 of those participants to return and provide feedback on a preliminary wearable prototype called HandSight. Findings from the controlled experiment show similar performance between haptic and audio directional guidance, although audio may offer an accuracy advantage for tracing lines of text. Subjective feedback also highlights tradeoffs between the two types of guidance, such as the interference of audio guidance with speech output and the potential for desensitization to haptic guidance. While several participants appreciated the direct access to layout information provided by finger-based exploration, important concerns also arose about ease of use and the amount of concentration required. We close with a discussion on the effectiveness of finger-based reading for blind users and potential design improvements to the HandSight prototype.

General Terms: Human Factors

Additional Key Words and Phrases: Accessibility, Visual Impairments, Wearables, Real-Time OCR

## 1. INTRODUCTION

Despite the increased availability of digital information and screen reader software, reading printed text materials remains an important but challenging task for people who are blind or visually impaired. The inability to read menus, receipts and handouts, bills and other mail can negatively impact the daily activities of those living with visual impairment (*e.g.*, [Brady et al. 2013; Haymes et al. 2002]). Although braille has long provided a promising alternative, fewer than 10% of the approximately 2 million adults with severe visual impairment in the United States are braille literate [National Federation of the Blind Jernigan Institute 2009; National Center for Health Statistics 2012], and many materials are not available in braille format.

Although many devices and mobile applications—such as *SARA CE*<sup>1</sup>, *KNFB Reader iOS*<sup>2</sup>, and *OrCam*<sup>3</sup>—attempt to provide access to printed materials through camera capture and optical character recognition (OCR), open questions remain. One

<sup>1</sup> SARA CE: <http://www.freedomscientific.com/Products/LowVision/SARA>

<sup>2</sup> KNFB Reader: <http://knfbreader.com/>

<sup>3</sup> OrCam: <http://www.orcam.com/>



**Figure 1:** HandSight uses a  $1 \times 1 \text{mm}^2$  AWAIBA NanEye 2C camera developed for minimally invasive surgeries (e.g., endoscopies) that can capture  $250 \times 250 \text{px}$  images at 44fps (a). Also shown are two early designs for our finger-based reading system (b) and (c). Future designs can be made much smaller.

challenge is how to help blind readers properly aim the camera so that a target object is completely visible and centered within the camera's field of view (e.g., [Vázquez and Steinfeld 2012; Jayant et al. 2011; Cutter and Manduchi 2015]). To accommodate this issue, the popular KNFB Reader iPhone application, for example, provides a spoken report to describe whether the document is fully visible and rotated correctly. Another challenge is how to interpret and communicate documents with complex layouts such as newspapers or menus. Determining which blocks of text to read, in what order, and what layout details to convey are known issues even with digital content [Lazar et al. 2007; Borodin et al. 2010].

As an alternative to these handheld solutions, recent advances in camera technology have led to new wearable approaches that augment the reader's finger with computer vision capabilities [Shilkrot et al. 2014; Shilkrot et al. 2015; Stearns et al. 2014; Yang et al. 2012]. Specifically for reading printed materials, *FingerReader* [Shilkrot et al. 2014; Shilkrot et al. 2015] and our past work [Stearns et al. 2014] both explore ring-based devices with an embedded camera that allows blind readers to trace their finger over printed text and hear real-time speech output. Compared to mobile applications, this finger-based approach may mitigate overhead camera framing issues, enable a blind reader to better understand the spatial layout of a document, and provide better control over pace and rereading.

A finger-based approach, however, also introduces new challenges that have not been fully investigated. Because the field of view from a finger-mounted camera is limited, the reader must precisely trace along the current line of text so that the image does not get cut off or distorted. Physical navigation through the document is also needed to support reading, such as finding the start of a text passage and moving from one line to the next. Thus, a finger-based reading approach is contingent not only on accurate text capture and OCR, but also on effective finger guidance. To date, however, user studies of a finger-mounted approach have focused on feasibility with small sample sizes (3-4 participants) and have not reported on quantitative performance metrics [Stearns et al. 2014; Shilkrot et al. 2014; Shilkrot et al. 2015]. This prevents an in-depth understanding of finger guidance effectiveness, reading performance, and user reactions. The most recent of these studies underscores the need for further investigation: despite the theoretical advantages of finger-based reading, all three participants found it difficult to read text with *FingerReader* [Shilkrot et al. 2015]. This provokes the question: *why?* To what extent are finger-based cameras a viable accessibility solution for reading printed text? What design choices can improve this viability?

To further investigate the feasibility of a finger-based sensing and feedback system for reading printed text, we conducted a controlled lab experiment to compare audio and haptic directional finger guidance with 19 blind participants using an iPad-based testbed (Study I). The primary goal was to compare the effects of the two guidance methods in terms of line tracing accuracy, reading speed, comprehension (through standardized comprehension questions), and subjective response. We later also randomly selected 4 of those participants to provide feedback on a preliminary wearable prototype, called *HandSight*, so as to help guide its design (Study II). *HandSight* is an extension of our previous work [Stearns et al. 2014]. These participants also provided feedback on the use of a smartphone app (KNFB Reader iOS) to read printed documents, which allowed us to compile a list of some of the relative advantages and disadvantages of each. An overview of preliminary results from these studies was presented in an ASSETS poster [Findlater et al. 2015].

The findings from Study I showed similar performance between haptic and audio directional guidance, although audio may offer an accuracy advantage for line tracing. While a small majority of participants preferred haptic guidance to audio, the overall split reflects contradictions found in previous research [Stearns et al. 2014; Shilkrot et al. 2014; Shilkrot et al. 2015]. Open-ended comments also highlight the tradeoffs of the two types of guidance, such as the interference of audio guidance with speech output and the potential for desensitization to haptic guidance. Finally, while several participants appreciated the direct access to layout information provided with *HandSight*'s exploration mode, and the lower learning curve of *HandSight* as compared to braille, important concerns arose about ease of use and the amount of concentration required. In the follow-up sessions (Study II), while not offering a controlled comparison, participants appreciated that *HandSight* provided immediate access to text content without the need to worry about first capturing the document, but overall they preferred the fast and smooth text-to-speech output of KNFB Reader iOS. Combined, these findings lead to new questions about finger-based reading, who may benefit the most from such an approach, and how to refine the design tested in our study.

The contributions of this paper are: (1) empirical results comparing audio and haptic directional finger guidance for a reading task in terms of user performance and subjective response; (2) the implementation and preliminary evaluation of a real-time proof-of-concept system that combines a small finger-mounted camera and feedback mechanism with efficient computer vision algorithms to read printed text; and (3) design reflections for finger-based reading devices for people who are blind. While our long-term goal is to investigate the many interactions made possible by collocating sensing and feedback on the fingers, our focus here is on the interactions necessary to use such a system to explore and read a physical document.

## 2. RELATED WORK

We survey existing reading approaches, including braille and dedicated reading devices for people who are blind, and research on enabling access to printed text materials.

### 2.1. Braille

Braille is one of the oldest and most widely used tactile writing systems for people who are blind [Millar 2003]. Raised line scripts (*e.g.* Moon) and alternative coded dot or dash scripts (*e.g.* Fishburne) are used occasionally for teaching newly blind adults

or for labeling personal items, but none have been as widespread or successful as braille. Braille literacy has been shown to correlate with stronger reading skills and higher employment rates, especially if taught at an early age [Ryles 1996]. However, braille literacy is steadily on the decline. In the past 40 years, the percentage of legally blind school-age children living in the United States who can read or are learning braille has decreased from more than 50% to fewer than 10% [National Federation of the Blind Jernigan Institute 2009]. Moreover, braille is difficult to learn and relatively slow to read—the average reading speed is 125 words per minute (wpm), as opposed to 250-300 wpm for sighted readers [Millar 2003] and even faster for experienced screen reader users [Asakawa et al. 2003]. Additionally, many physical materials still exist that are not available in digital or braille formats, and improving the accessibility of these materials could substantially improve the quality of life of blind individuals [Brady et al. 2013; Haymes et al. 2002].

Braille is a spatial medium, in much the same way as a printed page. Sighted readers use spatial information to help locate specific regions within a document (e.g., paragraphs, headings, captions, sidebars) [Keefer et al. 2013], and evidence suggests that braille readers do too [Millar 2003]. In braille documents, text is organized in straight lines with indentations to identify distinct blocks of text (*i.e.* paragraphs). Some forms also include embossed illustrations, maps, or graphs, with braille letters printed alongside them for labels and captions. These important spatial attributes, however, are not easily conveyed by a screen reader [Di Blas et al. 2004; Kouroupetroglou and Tsonos 2008].

Braille has other advantages as well. As a touch-based system, readers can easily control their reading speed and reread words or lines of text. The ability to reread text is critical, as it can increase comprehension and retention [Dowhower 1989; Levy 2001]. Additionally, although reading styles vary by individual, braille readers who use both hands have been shown to read much more quickly than one-handed readers [Burklen 1932; Millar 2003]. The reason for this is unclear—some individuals primarily use one finger to recognize and verify braille patterns and the other only for tracking position within the page, while others appear to use both fingers in an “enlarged perceptual window” [Millar 2003].

These strengths of braille inspired HandSight’s design. As with braille, HandSight should allow blind readers to gain an understanding of the location of document elements (margins, text, image placement), as well as to control reading speed and rereading. HandSight also relies on bimanual operation: the left hand tracks a document’s margin while the right hand scans and reads the printed text.

## 2.2. Dedicated Reading Devices

Scientists have long sought to support blind people in reading printed text by developing new technologies (for reviews: [Cooper et al. 1983; Capp and Picton 2000]). Many early so-called “reading machines for the blind” used a sensory substitution approach where the visual signals of words were converted to non-verbal auditory or tactile modalities. These systems were complicated to learn but increased the accessibility of printed text. Two such examples include the *Optophone*, which used musical chords or ‘motifs’ [D’Albe 1914] and the *Optacon* (OPTical to TActile CONverter), which used a vibro-tactile signal [Bliss 1969; Goldish and Taylor 1973]. As with our approach, the Optacon requires the user to move a camera sequentially along each line of text. However, this camera is handheld rather than worn and, more importantly, the user’s other hand receives a translated vibro-tactile signal that

provides some feedback about orientation to the line. The Optacon is well researched (e.g., [Hislop et al. 1983; Hislop et al. 1985]). These studies showed that even with training, reading rates (28.2 wpm on average) were limited to around 10% of the speed of sighted readers. Also, as with wearable finger-based reading approaches, Optacon users must take care in transitioning from the end of one line to the start of the next. However, even with Optacon’s steep learning curve, relatively slow reading speeds, and high cognitive load, the device continues to be used by blind readers to this day, suggesting that these challenges are not necessarily a barrier to use.

With advances in sensing, computation, and OCR, modern approaches attempt to scan, recognize, and read aloud text in real-time. This transition to OCR and speech synthesis occurred first with specialized devices (e.g., SARA CE, the original KNFB Reader<sup>4</sup>, [Gaudissart et al. 2004]), then mobile phones (e.g., *Text Detective*<sup>5</sup>, KNFB Reader iOS), and now wearables (e.g., [Shilkrot et al. 2014; Shilkrot et al. 2015], *OrCam*). While decades of OCR work exist (e.g., [Mori et al. 1992; Shen and Coughlan 2012; Chen and Yuille 2004; Wang et al. 2011]), even state-of-the-art reading systems become unusable in poor lighting, require careful camera framing [Jayant et al. 2011; Manduchi and Coughlan 2014], and do not support complex documents and spatial data [Kane et al. 2013]. Because HandSight is self-illuminating and co-located with the user’s touch, we expect that many of these problems can be mitigated or even eliminated.

As a wearable solution, HandSight is most related to OrCam and FingerReader [Shilkrot et al. 2014; Shilkrot et al. 2015]. OrCam is a recent commercial head-mounted camera system designed to recognize objects and read printed text in real-time. Text-to-speech is activated by a pointing gesture in the camera’s field-of-view. While live demonstrations with sighted users have been impressive<sup>6,7</sup>, there is no academic work examining OrCam’s effectiveness with visually impaired users for reading tasks. A primary distinction between HandSight and OrCam is the use of hand-mounted sensing versus OrCam’s head-mounted sensing. Although more work is needed to study the tradeoffs in camera position and the impact it has on function and overall user experience, HandSight’s finger-based approach allows sensing of finger gestures regardless of body or head position. Secondly, HandSight supports direct-touch scanning compared to OrCam’s indirect approach. Direct scanning should provide the user with increased control over reading speed as well as a better spatial understanding of a page. While our current studies focus on a finger-mounted approach, we plan to explore combining body- and head-mounted approaches in the future (e.g., body-mounted for acquiring global context and finger-mounted for precise information scanning).

More closely related to HandSight is FingerReader [Shilkrot et al. 2014; Shilkrot et al. 2015], a custom finger-mounted device with vibration motors designed to read printed text by direct scanning on a word-by-word basis. As with HandSight, haptic and audio directional finger guidance is provided to help guide the reader’s finger along the text. While the technical capabilities of FingerReader and the performance of its underlying algorithms have been evaluated, the two user studies reported to

<sup>4</sup> KNFB Reader Classic: <http://www.knfbreader.com/products-classic.php>

<sup>5</sup> Text Detective: <http://blindsight.com/textdetective/>

<sup>6</sup> OrCam TED@NYC. YouTube Video: [http://youtu.be/\\_3XVscsscyw](http://youtu.be/_3XVscsscyw)

<sup>7</sup> OrCam at Digital-Life-Design (DLD) in Munich. YouTube Video: <http://youtu.be/3m9ivtJI6iA>

date have been feasibility studies with 3 and 4 participants respectively, focusing primarily on qualitative feedback. As such, it has been difficult to draw generalized conclusions about the effectiveness of finger-based reading and, more specifically, about directional guidance to support that reading. Indeed, user preference for haptic versus audio directional guidance in the two studies was conflicting. Our work addresses the need for a more in-depth investigation by comparing audio and haptic guidance with 19 participants using quantitative speed and accuracy metrics in addition to subjective feedback.

### 2.3. Mobile Applications for Reading Printed Text

As mentioned above, smartphone-based accessibility applications such as KNFB Reader iOS and Text Detective have begun to replace bulkier and more expensive dedicated devices like SARA CE. KNFB Reader iOS, which we consider in Study II, evolved from a dedicated scanner device in the 1970s to a Windows PDA and Nokia phone app in the early 2000s, and is now available as a \$99 iPhone app [Holton 2014]. It uses the phone's built-in camera to detect and scan physical documents, then processes the captured photo and reads detected text aloud using speech synthesis (like a screen reader). The application has been well received and represents the state-of-the-art in smartphone-based document reading for the blind population [Kendrick 2008; Holton 2014]. To help users capture a document, the iPhone app provides vibration feedback based on tilt and a spoken report on document edges and rotation. Text Detective is similar to KNFB Reader iOS but has a lower user rating on the Apple App Store and did not perform as well in our own informal user tests.

Smartphones with pervasive internet access have also enabled crowd-powered accessibility approaches [Bigham et al. 2011] such as VizWiz [Bigham et al. 2010] and *BeMyEyes*<sup>8</sup>. Here, sighted volunteers and/or employees answer questions about visual content, including text recognition. As with mobile applications like KNFB Reader iOS, a target must be appropriately framed in the camera to receive useful feedback. Additionally, a crowdsourced approach incurs a delay for the crowd worker to respond and has negative implications for privacy. Automated applications, while perhaps not as reliable, offer more immediate results and fewer privacy concerns.

In contrast to these smartphone-based approaches, HandSight allows the user to directly touch printed text, control reading speed, and explore the spatial layout of the page. Framing may also be less problematic because the camera distance and angle are fixed to the user's finger.

### 2.4. Interfaces for Directional Guidance

While readers with low vision may be able to precisely trace along a line of text with their finger, blind readers will require additional finger guidance. This directional guidance needs to be responsive enough to use at reading speeds approaching braille, but should not distract the reader from listening to and understanding the text.

Haptic and audio directional guidance cues have been used in numerous projects to guide visually impaired users. Manduchi and Coughlan [Manduchi and Coughlan 2014], for example, combined synthesized speech directions with discrete audio patterns to guide users in properly framing a target for camera capture. Sonification

<sup>8</sup> <http://www.bemyeyes.org/>

has been used to assist blind users in exploring complex indoor maps [Su et al. 2010], to support non-visual learning of touchscreen gestures [Oh et al. 2013], and to convey radial direction with vowel sounds [Harada et al. 2011]. Other researchers have combined sonification and (haptic) force feedback to teach handwriting to blind children [Plimmer et al. 2011] and shapes to users with visual impairments [Crossan and Brewster 2008]. This lattermost study, for example, found that combining audio and force feedback resulted in higher user performance than force feedback alone; however, the force feedback controller moved in 3D space, so is not comparable to the finger-mounted vibration motors used in HandSight. Yet other researchers have explored primarily tactile feedback to transmit directional and shape data [Yatani and Truong 2009; Yatani et al. 2012; Noble and Martin 2006].

HandSight’s audio and haptic directional guidance is designed based on these previous findings. For audio, we use pitch to convey direction, a decision based on work showing that pitch is the most salient for conveying gestural trajectory information among pitch, stereo, timbre, and volume [Oh et al. 2013]. Similar to Shilkrot *et al.* [Shilkrot et al. 2015], our haptic guidance condition varies the intensity and location of vibration on the finger to indicate direction and distance from a target (the center of a line of text).

## 2.5. Interfaces for Spatial Exploration

In addition to sequential text reading capabilities, blind readers may benefit from the ability to explore the layout of a printed page, locating and identifying the sizes and positions of images and blocks of text. For example, a complicated layout (*e.g.*, a two-column menu) can be difficult to interpret when scanned and read aloud by screen reader software, but by exploring the spatial positions of the text and images a user may be able to better understand the content.

Previous researchers have investigated ways to convey the spatial layout and other non-textual information on websites or touchscreens via audio or vibration. For example, Mynatt et al. [Mynatt and Weber 1994] explored methods for conveying visual user interface elements on a computer screen using braille displays and audio cues. Vibro-audio interfaces have also been explored to convey graphical information on a touchscreen such as bar graphs and visual shapes [Toennies et al. 2011; Giudice et al. 2012]. Other researchers have explored the use of sonification and tactile displays to enable blind users to access digital map data [Parente and Bishop 2003; Poppinga et al. 2011]. While our own “exploration” is simpler than these approaches (conveying the presence of text, images, and white space), future iterations can build on this body of work to enable users to more fully explore the visual content and spatial layouts of printed documents.

## 3. STUDY I: AUDIO VS. HAPTIC GUIDANCE FOR FINGER-BASED READING

To investigate the exploration and reading of printed text documents using finger-based interactions, we conducted a controlled lab study with 19 blind participants. The primary goal of this study was to compare audio and haptic directional finger guidance methods in terms of user performance and preference. However, as the first larger-scale study of finger-based reading ( $N=19$  *vs.*  $N=3$  and  $N=4$  [Stearns et al. 2014; Shilkrot et al. 2015]), the study also quantitatively explored to what extent a finger-based reading approach can allow a blind reader to interpret the spatial layout of a document and to read and understand that document.

We simulated the experience of reading a physical document using a touchscreen tablet (an iPad) covered with a sheet of paper (Figure 2c). This approach allowed us to bypass certain technical challenges in implementing a real-time camera and text recognition system, and instead to focus on the user experience of finger-based reading. The iPad also allowed us to collect precise finger traces to enable detailed finger-movement analysis not previously possible.

### 3.1. Method

In this controlled lab study, participants read two types of printed documents with audio and haptic finger guidance. We used a within-subjects design with a single factor of *Directional Guidance* that had two levels (*Audio* and *Haptic*); order of presentation of the conditions was fully counterbalanced. In addition to measuring reading speed and finger movement, we collected subjective feedback and assessed basic document comprehension using standardized questions. Despite similarities to Shilkrot *et al.*'s method [Shilkrot et al. 2014; Shilkrot et al. 2015], our protocol is an extension of our previous work [Stearns et al. 2014], which was underway prior to the first FingerReader publication [Shilkrot et al. 2014]. The final apparatus and method described here were also refined through pilot sessions with 5 additional participants (1 sighted, 1 low vision, 3 blind) who did not take part in the full study.

**3.1.1. Participants.** Twenty participants were originally recruited via campus email lists and local organizations, but one participant's data was discarded because he was unable to complete all of the required tasks. Of the remaining 19 participants, 11 were male and 8 were female, and the median age was 48 ( $SD=12.0$ , range 26–67). All participants were completely blind or had only minimal light perception. Five participants were congenitally blind, while the others had lost their vision later in life (some as recently as two years ago). As shown in Table 1, most participants were

ID	Age	Sex	Vision Level	Braille		Screen Reader		Computer
				Use	Comfort	Use	Comfort	Comfort
P1	54	F	Blind	5	5	4	3	4
P2	33	F	Light	4	5	5	5	5
P3	55	M	Blind	3	5	5	5	4
P4	44	M	Light	2	2	5	5	5
P5	67	M	Blind	3	4	5	5	4
P6	62	M	Light	3	4	5	5	4
P7	40	M	Blind	1	1	5	4	4
P8	27	F	Light	5	5	5	5	4
P9	49	F	Light	5	5	5	5	3
P10	43	M	Blind	5	4	1	1	3
P11	44	M	Light	4	4	1	1	1
P12	39	M	Blind	4	5	5	5	5
P13	67	M	Blind	3	3	1	1	1
P14	50	F	Light	4	4	5	5	5
P15	26	M	Blind	5	5	5	5	5
P16	48	M	Blind	5	4	5	4	4
P17	59	F	Light	2	3	1	1	1
P18	47	F	Blind	4	3	1	1	1
P19	64	F	Light	4	3	4	4	3
<b>Mean (SD)</b>	48.3 (12.0)	N/A	N/A	3.7 (1.2)	3.8 (1.1)	3.8 (1.7)	3.7 (1.7)	3.5 (1.4)

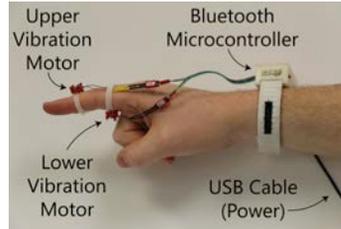
**Table 1:** Study I participants. All participants were either blind or had minimal light perception (denoted "Light"). Frequency of use ("Use") varied from 1 ("never") to 5 ("very often"), while comfort level ("Comfort") varied from 1 ("very uncomfortable") to 5 ("very comfortable").

People have used coins as a means of exchange for thousands of years. Valued for their craftsmanship and purchasing power, coins have been collected in great numbers throughout history and buried for safekeeping. Because coins of some particular nations or cities were in circulation for many years, they can reveal a great deal about a given culture.

Coins are useful in revealing many aspects of a culture. They can provide clues about when a given civilization was wealthy and when it was experiencing a depression. They also indicate how a given nation's technology advanced over time. The distribution of coins can also reflect the boundaries of an empire and the trade relationships within it. Roman imperial gold coins found in India indicate the Roman merchant goods from the East.

The way the coins themselves are decorated sometimes provides key information about a culture. Many coins are stamped with a wealth of useful historical evidence, including portraits of political leaders, important buildings and temples, mythological and religious figures, and useful dates. Some coins, such as many from ancient Greece, can be considered works of art themselves and reflect the artistic achievement of the civilization as a whole.

Information gathered from old coins by historians is most useful when placed alongside other historical documents, such as written accounts or data from archeological digs. Combined with these other pieces of information, coins can help historians reconstruct the details of lost civilizations.



(a) Screenshot of iPad software showing a single-column document.

(b) Haptic feedback device, with actuator mounted on the finger.

(c) Test setup, with physical paper covering the iPad.

**Figure 2:** Study I test apparatus.

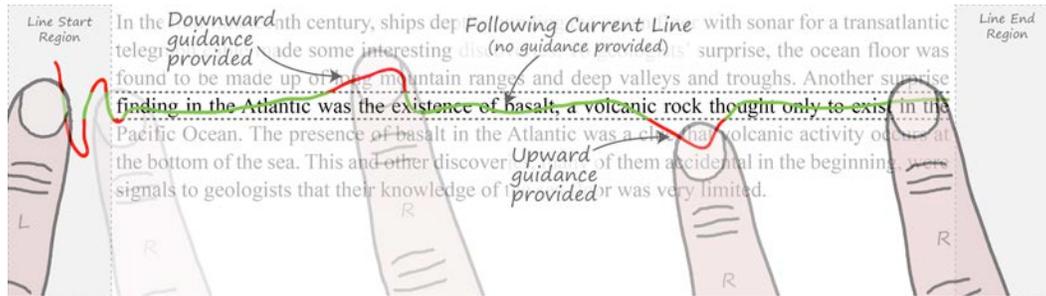
frequent users of braille, although 6 were just learning to read it and rated their comfort level as lower. All but 5 participants used screen readers at least some of the time and only 4 were not comfortable with computers and/or mobile devices. Participants were compensated for time and transportation.

**3.1.2. Apparatus.** The test apparatus consists of an Apple iPad running custom software and connected via Bluetooth to a custom-built finger-worn haptic device (Figure 2). The source code is available on GitHub<sup>9</sup>. As noted previously, the iPad was used to provide a dynamic test environment that could precisely track finger movement in response to our directional guidance conditions. To simulate the feel of a physical document and reduce friction from the screen, a thin, blank paper covered the iPad. In addition, because there is no tactile border between the iPad screen and bezel, we added our own physical border made of 1/16" flexible foam (Figure 2c). The software displays documents and provides two modes of interaction: exploration and reading. All touch events (down, up, and move) on the screen are logged with x, y coordinates and timestamps.

**Exploration mode.** In this mode, audio cues allow users to gain a spatial sense of the document layout (e.g., locations of images, columns, paragraphs) before transitioning to reading mode. As the user traces their finger over the document, they hear either a high-pitched flute sound when on a block of text or a low-pitched cello sound when on a picture. These sounds were selected and refined via pilot testing to be easily distinguishable by their pitch and timbre. When over whitespace, such as between paragraphs or columns, no sound plays.

**Reading mode.** In this mode, the user traces their finger from left to right along each line of text, while the system generates text-to-speech output using Apple's default iOS speech synthesis engine and provides directional finger guidance (haptic or audio depending on the condition). Reading is bimanual: the left hand, which is uninstrumented, serves as a line anchor (see "line start region" in Figure 3) while the right index finger traces the line. To begin reading, the user moves their right finger to the line start region shown in Figure 3 and an audio cue of ascending xylophone notes plays. If the finger is not already at the first line of text, audio or haptic feedback guides the user's finger up or down. Once the right finger is properly positioned over the "line start region" of the first line, the left hand joins the right hand and subsequently serves as a line anchor.

<sup>9</sup> <https://github.com/HCIL/HandSight>

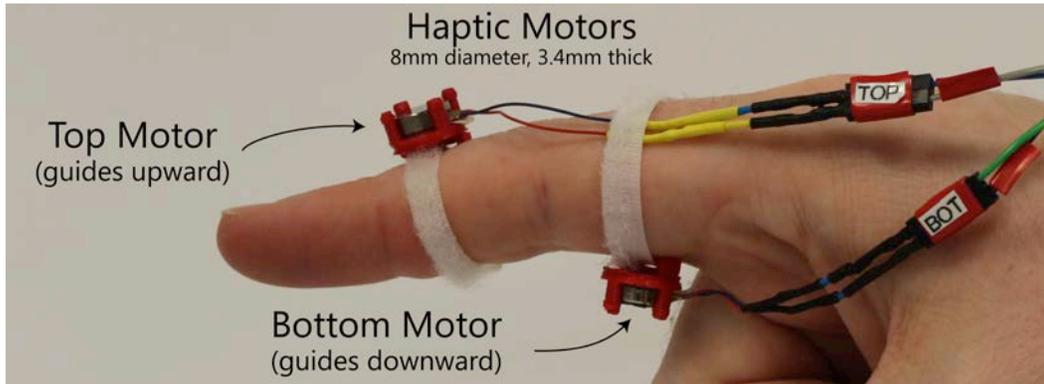


**Figure 3:** Reading mode interaction is bimanual. The user (1) places the right index finger in the “line start region” and moves vertically to find the start of the current line; (2) places the left index next to the right finger as an anchor; (3) traces the right finger along the line until it reaches the “line end region”; (4) returns the right index finger to beside the left finger before moving down to the next line. When the right finger is directly on the line (green trace) no directional guidance is provided, but when the finger moves too high or low (red trace), audio or haptic guidance indicates which direction to move to return to the line.

The user then traces his/her finger along the line to the right, while the system speaks each word aloud and provides audio or haptic guidance whenever the finger strays above or below the line (Figure 3). The speed of the text-to-speech output adapts to match the speed of the finger movement. Speech is provided only for the current line, and only when the user’s finger is within 73 pixels (0.7cm) of the middle of the line (simulating a finger-mounted camera’s field of view). At the end of the line (“line end region”; Figure 3), another audio cue plays, this time with descending xylophone notes, and the text-to-speech stops. The user then moves their finger left again to find the line start region and read the next line in the same manner. Finally, at the end of a paragraph, a new audio chime plays. The audio cues for the start and end of line and end of paragraph were selected to be easily distinguishable, which we again verified using early feedback from pilot participants.

For *audio directional guidance*, the system provides a continuous tone that varies in pitch. A low pitch indicates that the finger should move down and a high pitch indicates that the finger should move up. If the finger is properly positioned over the current line, no audio plays. If the user’s finger moves above the line, an audio tone at frequency 300Hz begins playing. If the user’s upward movement continues, the frequency linearly decreases based on distance, down to a minimum of 200Hz at 127 pixels (1.2 cm). The 200Hz tone continues for any movement more than 2.4cm above the line. Similarly, if the user’s finger strays below the line, the audio frequency begins at 500Hz and increases to a maximum of 600Hz at 127 or more pixels away. The choice to vary audio frequency to indicate distance and direction was motivated by our prior work [Stearns et al. 2014; Oh et al. 2013], and the exact pitches and distance thresholds were selected after pilot sessions.

The *haptic directional guidance* includes two disc vibration motors (8mm diameter, 3.4mm thick) controlled by an Arduino Pro Micro that communicates with the iPad via Bluetooth. The motors are attached to the user’s right index finger with separate Velcro rings (Figure 4), one on top of the finger on the intermediate phalange and one below the finger on the proximal phalange. The lower motor indicates that the finger should move downward and the upper motor indicates the opposite. Neither motor vibrates while the user’s finger is directly over the current line of text. Vibration intensities off the line range from a minimum perceptible



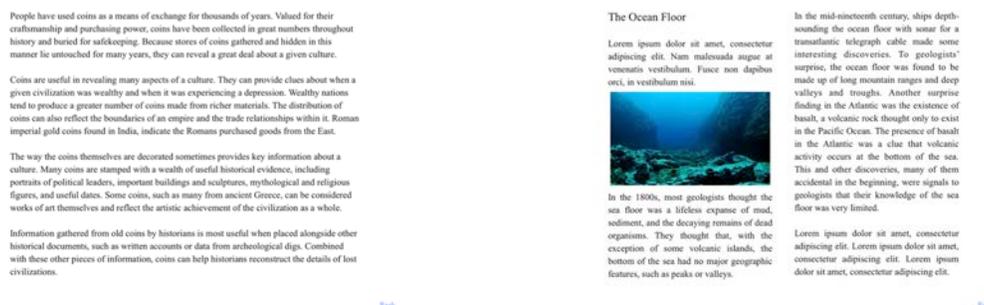
**Figure 4:** Close-up view of the haptic motors mounted on the finger via Velcro rings. The top motor vibrates when the user's finger moves below the line, providing upward guidance; the bottom motor vibrates when the user's finger moves above the line, providing downward guidance. The intensity of vibration depends upon the distance to the line, achieving maximum intensity at 127 pixels (~1.2 cm).

strength to the maximum strength the motors can provide, using the same distance thresholds as the audio condition. The choice to vary the position and intensity of vibration to indicate direction and distance was also motivated by our prior work [Stearns et al. 2014] and validated in pilot sessions.

In early testing within our research lab and with external pilot participants, we tested multiple mappings for audio and haptic cues and intended finger direction (e.g., higher pitch to indicate up *vs.* the opposite). Users were split in terms of which mappings were most intuitive, a point we revisit in the Discussion section.

**3.1.3. Procedure.** Each study session lasted 1.5–2 hours. Throughout, we employed two document types (Figure 5): single-column plain text, and two-column magazine-style with a figure and an article heading. For the reading tasks described below, we adapted four test documents from a Grade 8 Iowa Test of Basic Skills practice book [Princeton Review 2000]. The original text was modified slightly for length and to ensure clarity with our speech synthesis engine (e.g., removing unnecessary proper nouns); see Appendix. The documents were thus all at similar reading levels and had multiple-choice comprehension questions. We also created training documents that were similar in length to the four test documents.

Following a background questionnaire, participants first learned how to use the



**Figure 5:** Examples of our test documents: plain text (left), and magazine (right).

document exploration mode as a precursor to the more complex task of both exploring and reading a document. The experimenter demonstrated the audio cues for text and images in exploration mode, then asked participants to explore one plain document and one magazine document for up to three minutes each. To ensure that participants understood the exploration mode, participants were asked about the structure and layout of each document (*i.e.*, how many paragraphs and columns, are there pictures or headings and if so where are they located). To avoid biasing participants toward a particular exploration strategy or interpretation, we initially provided very little direction aside from demonstrating the audio cues and warning participants of the questions they would be expected to answer. After recording the answers for a document, the experimenter then guided participants to find the correct answers to ensure that they could later use the mode correctly prior to each reading task.

After the introduction of exploration mode, participants explored and read documents with each of the directional guidance conditions (audio and haptic). The order of presentation for these conditions was fully counterbalanced. Document order was identical across all participants so that the documents themselves were matched an equal number of times with each guidance condition. To ensure similar physical experiences across conditions, participants wore the Arduino wristband and finger rings with the haptic motors throughout the study session.

The procedure for each directional guidance condition was identical, with training using a plain document (~10 minutes) followed by testing with two documents (one plain and one magazine). For the training document, the experimenter demonstrated the feedback cues and participants incrementally learned to follow a line, find the next line or paragraph, and listen to the speech feedback while moving their finger. For each test document, participants were allowed up to 90 seconds in exploration mode to assess the layout before the experimenter switched the system to reading mode. For the plain document, the reading task was to locate the first line of text and read the entire document. For the magazine document, participants read the last paragraph in the first column and the first paragraph in the second column. Exploration mode was used to locate the start of text for each document, as well as the start of the second column for the magazine document. After each test document, two multiple-choice comprehension questions provided in the Grade 8 Iowa Test of Basic Skills practice book were administered. At the end of each guidance condition, participants were asked about subjective ease of use. Finally, at the end of the study participants were asked to compare the two directional guidance conditions. See Appendix B for the full text of the subjective questionnaires.

Before conducting this study, we validated our selection of test documents and comprehension questions in a simple baseline study. Ten sighted college-age participants listened to synthesized speech of the four test documents and the comprehension questions. All 10 participants answered the questions correctly.

**3.1.4. Data and Analysis.** Collected data included log files from the iPad, participant responses to close- and open-form questions, and experimenter observations. To compare reading performance with haptic and audio guidance, we examined the following subtasks separately:

- *Line finding:* Finding the start of the current line. A line finding instance began with the first right-handed touch within the line start region (Figure 3)

and ended with the finger exiting that region. Sometimes participants' search paths resulted in more than one exit from the start region, so we included all data up to the final exit. For each line finding instance, we calculated elapsed time and, as an error measure, the length of the movement path traced.

- *Line tracing:* Tracing left-to-right along the current line. A line tracing instance included all touch points after a successful line finding subtask until the right index finger entered the line end region (Figure 3). For each line tracing instance, we calculated reading speed in words per minute (wpm), and, as an error measure, the average absolute distance of the finger from the vertical center of the line across all  $x$ -coordinates in that line trace.
- *Full document:* Reading the full document from the start of the first line to the end of the final line. This comprehensive analysis includes all line finding and line tracing subtasks for a single document, as well as the time to transition between columns for the magazine documents. For each document, we calculated the average reading speed in words per minute (wpm) as well as the number of skipped words that were not read aloud.

Across the 19 participants, we collected data for 1513 lines. We identified outlier samples that were more than 3 standard deviations away from the mean for a given participant and condition, removing 31 samples (2.0%) of line tracing subtask samples and 49 (3.4%) of line finding subtask samples.

We used paired t-tests to compare line tracing speed between haptic and audio guidance. However, other measures violated the normality assumption of a t-test (determined using separate Shapiro-Wilk tests for each measure,  $p < 0.05$ ). For these measures, we conducted non-parametric Wilcoxon signed rank tests to compare haptic and audio. For all posthoc pairwise comparisons, we applied Holm's sequential Bonferroni adjustments to protect against Type I error [Holm 1979].

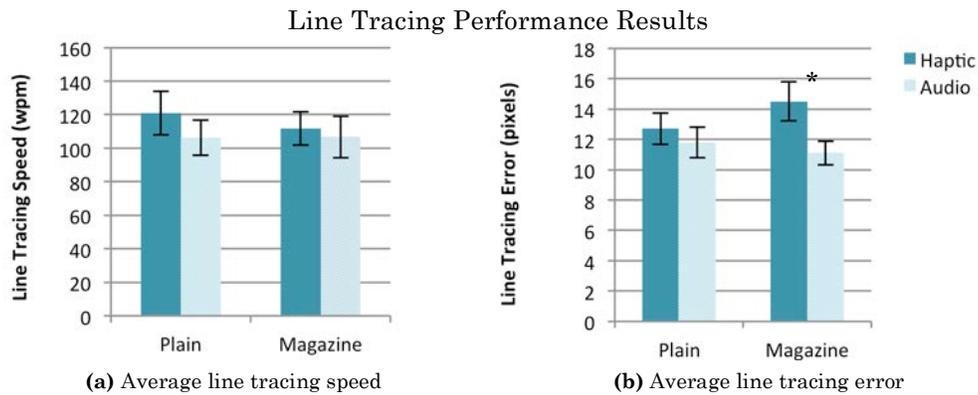
### 3.2. Findings

Our findings include performance results derived from the log data and exploratory descriptions of how participants responded to and interacted with the finger-based reading approach (*e.g.*, initial use of exploration mode, potential advantages of such an approach).

**3.2.1. Reading Mode – Line Tracing.** Figure 6 shows line tracing performance. For plain documents, the average reading speed with haptic guidance was 120.9 wpm ( $SD=57.0$ ), compared to only 106.3 wpm ( $SD=46.2$ ) with audio; however, a paired t-test comparing the two types of guidance was not statistically significant. A similar trend followed for magazine documents, at 111.8 wpm ( $SD=43.3$ ) and 106.7 wpm ( $SD=54.1$ ) for haptic and audio, respectively, with a paired t-test revealing no statistically significant difference between the two.

In terms of error, audio guidance was significantly more accurate than haptic guidance for the magazine documents, with an average distance of 11.2 px ( $SD=3.5$ ) to the center of the line versus 14.6 px ( $SD=5.7$ ). A Wilcoxon signed rank test was statistically significant on this measure, with a large effect size ( $Z_{19}=-2.374$ ,  $p=.018$ ,  $r=.54$ ). Figure 7 shows a representative finger trace that illustrates this performance difference. For the plain documents, however, the two guidance conditions resulted in more similar distances, at 11.9 pixels for audio ( $SD=4.6$ ) and 12.8 pixels for haptic ( $SD=4.6$ ). This difference was not significant using a Wilcoxon signed rank test.

Participants tended to drift frequently, spending on average 29.7% ( $SD=13.2$ ) of their line tracing time off of the line for the audio condition and 37.8% ( $SD=14.7$ ) for the haptic condition. Reflecting the distance accuracy results above, this difference was statistically significant with a Wilcoxon signed rank test ( $Z_{19}=-2.57, p=.010, r=.59$ ). In addition, participants tended to drift consistently above or below the line. Figures 7a, for example, illustrates downward drift whereas Figure 7c shows upward drift. We observed 11 participants who drifted consistently upward, 4 who drifted consistently down, and 4 who varied by document or did not tend toward either direction. This tendency may have been affected by how each participant’s arm was positioned relative to the iPad—participants were instructed to rotate the screen as needed, but few chose to do so.



**Figure 6:** Average line tracing speed (higher is better), and average error—vertical distance offset from the center of the line (lower is better). Error bars indicate standard error ( $N=19$ ). Performance was generally similar between the audio and haptic conditions, but audio resulted in significantly lower line tracing error for the magazine document (\*).

family, he smiled, shrieked, pounded the ground, and looked from one member of the family to the next. Still smiling and shrieking, Nim went around hugging each member of the family. He played with and groomed each member of the family for almost an hour before the family had to leave. People who were familiar with Nim’s

(a) Audio and magazine document (P8)

made up of long mountain ranges and deep valleys and troughs. Another surprise finding in the Atlantic was the existence of basalt, a volcanic rock thought only to exist in the Pacific Ocean. The presence of basalt in the Atlantic was a clue that volcanic activity occurs at the bottom of the sea. This and other discoveries, many of them

(b) Haptic and magazine document (P8)

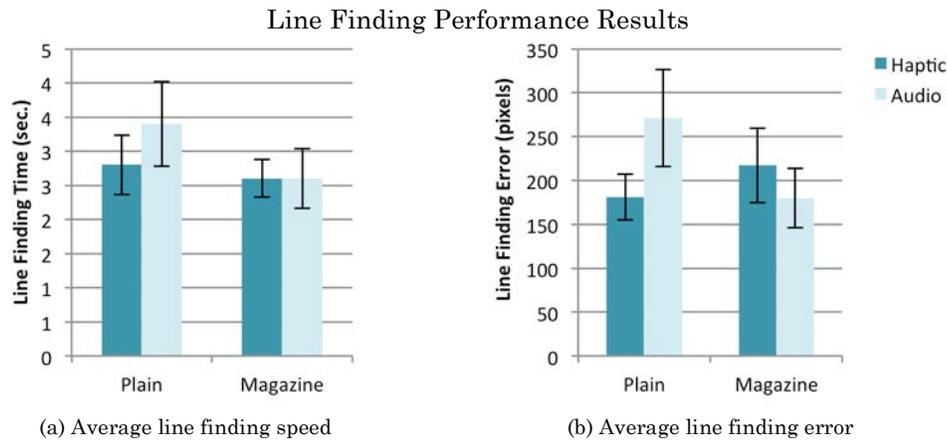
People have used coins as a means of exchange for thousands of years. Valued for their craftsmanship and purchasing power, coins have been collected in great numbers throughout history and buried for safekeeping. Because stores of coins gathered and hidden in this manner lie untouched for many years, they can reveal a great deal about a given culture.

(c) Audio and plain document (P7)

**Figure 7:** Example finger traces. Solid (green) indicates that the finger was on the line, while dotted (red) indicates that the finger was off the line and directional guidance was being provided. (a) and (b) illustrate the difference in accuracy between the audio and haptic guidance conditions for P8. Participants frequently reacted more immediately to audio guidance, but tended to ignore small amounts of vibration with haptic guidance. This observation may explain the significant difference in error between the audio and haptic conditions. Participants also tended to drift consistently above or below a line as they read, as seen in (a), (b) and (c).

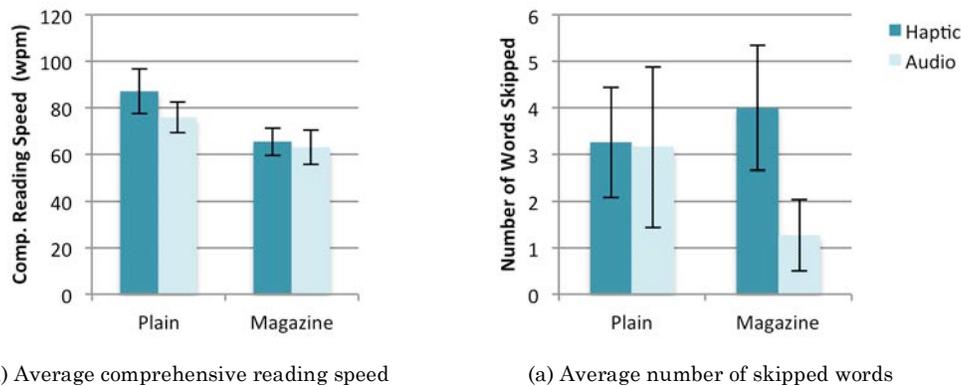
**3.2.2. Reading Mode — Line Finding.** As shown in Figure 8, line finding performance was similar across all directional guidance conditions and document types. No significant differences were found between haptic and audio guidance for either document type or performance measure using Wilcoxon signed-rank tests. Across all conditions, it took participants on average 2.6–3.4 seconds to find the next line in a document (plain: haptic  $M=2.8$  seconds,  $SD=1.9$ , and audio  $M=3.4$ ,  $SD=2.7$ ; magazine: haptic  $M=2.6$ ,  $SD=1.7$ , and audio  $M=2.6$ ,  $SD=1.9$ ). For error, measured as the average path length while searching for the start of a line, haptic averaged 181.2 pixels ( $SD=114.0$ ) with plain documents, while audio averaged 270.9 pixels ( $SD=241.1$ ). In contrast, for magazine documents, haptic averaged 217.1 pixels ( $SD=184.4$ ), compared to 179.8 pixels for audio ( $SD=147.5$ ). Again, however, these differences were not found to be statistically significant.

**3.2.3. Reading Mode — Overall Performance and Comprehension.** Figure 9 shows comprehensive performance over the documents, including the total reading time and



**Figure 8:** The average time elapsed (left) and error (right) in finding the next line; lower is better for both graphs. The error bars indicate standard error ( $N=19$ ). Performance differences between the two conditions were not significant.

### Comprehensive Performance Results



**Figure 9:** The comprehensive reading speed for an entire document (higher is better) and total number of skipped words (lower is better) by document. The error bars indicate standard error ( $N=19$ ). Performance differences between the two conditions were not significant.

number of skipped words. Reading speeds ranged from 63–81 wpm (plain: haptic  $M=81.1$  wpm,  $SD=42.1$  and audio  $M=75.8$  wpm,  $SD=29.0$ ; magazine: haptic  $M=65.4$ ,  $SD=25.6$  and audio  $M=63.0$ ,  $SD=31.9$ ). Overall, the number of skipped words, that is, words that were not read aloud by the text-to-speech engine, was uniformly low across conditions. The four documents contained an average of 211.5 words, but only 1–5 of those words were skipped on average for any given document. The number of skipped words was also similar between conditions for the plain documents (plain: haptic  $M=3.3$ ,  $SD=5.2$  and audio  $M=3.2$ ,  $SD=7.5$ ; magazine: haptic  $M=4.0$ ,  $SD=5.9$  vs. audio  $M=1.3$ ,  $SD=3.3$ ). Using Wilcoxon signed-ranks tests, no significant differences were found between haptic and audio guidance for either measure (speed, number of skipped words) with either document type.

While further investigation is needed to determine to what extent audio and haptic guidance impact comprehension, overall, participants answered the comprehension questions with high accuracy. Across all participants and conditions, 85% of the questions were answered correctly (Table 2).

**3.2.4. Overall Subjective Response.** Overall preference was split, with a small majority of participants (11 out of 19) preferring haptic feedback, 7 preferring audio, and 1 reporting equal preference. Participants also rated the two types of guidance in terms of comprehension and line tracing ease, from 1 – *very difficult* to 5 – *very easy*. The ratings, shown in Table 3a, support the overall preference patterns. Both guidance conditions were rated somewhat positively for both measures (3.1 or higher on average), and the differences between the two conditions were not statistically significant with Wilcoxon signed rank tests. Some challenges with the HandSight approach were seen as common to both types of guidance. For example, P12 said,

Guidance	Document	2/2 Correct	1/2 Correct	0/2 Correct
Audio	Plain	14 participants	3 participants	2 participants
Haptic	Plain	17	2	0
Audio	Magazine	12	5	2
Haptic	Magazine	14	5	0

**Table 2:** Number of participants who answered the set of two comprehension questions correctly in each experimental condition ( $N=19$ ). Most questions were answered correctly regardless of condition.

	Question	$N$	Mean	$SD$
(a)	Reading comprehension with audio guidance	19	3.2	1.3
	Reading comprehension with haptic guidance	19	3.7	1.2
	Line tracing with audio guidance	19	3.3	1.3
	Line tracing with haptic guidance	19	3.1	1.4
(b)	Start of text detection	19	4.5	0.7
	Start of line detection	19	4.2	0.6
	End of line detection	19	4.7	0.5
	End of paragraph detection	19	4.6	0.8
(c)	HandSight vs. braille	18	3.0	1.0
	HandSight vs. screen readers	14	2.9	1.2
	HandSight vs. other reading aids	12	2.4	1.2

**Table 3:** Study I subjective ratings from 1 to 5 where 5 is the best. (a) Reading comprehension and line tracing for each guidance condition. (b) Experience with subtasks common to both guidance conditions. (c) Overall comparison (better/worse) of HandSight versus braille, screen readers, and other reading aids. A score of 5 indicates that HandSight was perceived as much better than the existing technology, while a score of 1 indicates that it was much worse.

*“The haptic feedback only tells you when you’re not in line, not where the next thing would be”*, and made a similar comment for audio guidance.

The 11 participants who preferred haptic guidance generally felt that it was more intuitive, easier to use or faster than the audio. For example, P13 stated: *“It gave me a clearer indication of which way, up or down”*. P9 also commented, *“The vibrations kind of helped as a prompt, so that I automatically would go in the right direction, and I was able to read faster”*. Six of the participants who preferred haptic guidance also mentioned that the audio guidance was more distracting, and that made it harder to focus on the speech feedback: *“You could focus on the audio of the text, and not be listening for other sounds”* (P7), or *“I missed a couple words because I was being distracted by the [audio]”* (P15). Even 4 of those who preferred audio guidance mentioned that the overlapping sounds could be somewhat distracting.

Of the 7 participants who preferred audio, almost all ( $N=6$ ) found haptic guidance to be confusing: *“Sometimes when I use the vibrations I would forget which direction I was going based on where the vibration was”* (P5), or *“I had to analyze more what the vibrations meant”* (P14). Two participants also mentioned concerns about comfort, especially for prolonged use, for example: *“If you’re reading longer your finger might get numb and it might get more difficult to figure out where the vibration was”* (P14).

Participants found the audio cues common to both guidance conditions relatively easy to use. Using these audio cues to detect the start of the text, line start/end areas, start of a column, and end of paragraph were all rated above 3.8 on a 5-point scale (Table 3b). Detecting the start of a column received the lowest score ( $M=3.8$ ), perhaps reflecting the challenge of reading text with a more complex layout. This challenge can be non-trivial for some users. It should be noted that the participant whose data we discarded (described in Section 3.1.1), had been blind since early childhood and was thus unfamiliar with the concept of a two-column document, an issue that requires further consideration in future work. He said: *“Can a document be structured this way, with a paragraph just taking half part of the page?”* Other participants also found the magazine document to be more difficult, especially those who were congenitally blind, but all were able to successfully complete the task.

**3.2.5. Comparison to Other Technologies.** As shown in Table 3c, the overall experience of HandSight was rated similarly compared to braille ( $M=3.0$ ,  $SD=1.0$ ), and somewhat negatively compared to other aids such as cell phone apps or scanner hardware ( $M=2.9$ ,  $SD=1.2$ ), and screen reader software ( $M=2.4$ ,  $SD = 1.2$ ).

Seven participants who were not comfortable with braille or existing reading technologies generally liked the lower learning curve and flexibility of our reading approach. For example, P11, who was currently learning braille said: *“With braille you gotta always constantly remember which dots are for which letters [...]. this will tell you what the word is. Less stress.”* (P11). P7 also commented on the utility of being able to directly control reading speed with our approach: *“A [screen] reader you get like one speed, it doesn’t slow down for any reason, and sometimes it’s a lot harder to go back and get your place from where you stopped.”*

However, nine participants who were more familiar with braille and other reading devices raised concerns about ease of use and cognitive load. P14, for example, preferred braille: *“Reading braille I can read at a steadier pace and I can know where the punctuation is, and it’s easier for me to find the next line”* (P14). Both P16 and P18 commented on cognitive load: *“There’s the need to concentrate on staying within lines”*

(P16), and, *“I’m so focused on trying to read the document, I’m not necessarily retaining the information the way I want to”* (P18).

**3.2.6. Initial Use of Exploration Mode.** The analyses above focus on reading mode, but at the start of the study, participants first used exploration mode to receive feedback on the presence of text, images and whitespace in both plain text and magazine documents. Even with this initial use, all but one participant correctly identified the presence or absence of a picture in both documents and described the picture’s location. Determining whether audio breaks represented a gap between two paragraphs or two columns was more difficult, such that 11 participants initially identified multiple columns in the plain text document. However, between the two documents, the experimenter revisited how to distinguish between paragraphs and columns, and almost all participants (17 out of 19) were able to report the correct number of columns for the magazine document. Precisely counting paragraphs was still difficult, with only 9 and 3 participants reporting the correct number for the plain and magazine documents, respectively. For the magazine document the primary source of error was confusion over the definition of a paragraph in a multi-column document—the majority ( $N=15$ ) did not count the paragraphs in the two columns separately. Additionally, 7 participants mistook the heading in the magazine document for another paragraph, and only 9 answered questions about it correctly.

We observed a few exploration strategies, with some participants using multiple strategies: 8 initially moved their fingers quickly but in no discernible pattern, searching out the locations of images and text within the document; 8 followed a procedure similar to reading braille, exploring left to right sequentially down the page; 12 explored sequentially left to right then top to bottom, counting breaks in the sound to identify paragraphs and columns. Though we only told participants that we would ask them about the number of paragraphs, columns, and the presence/location of certain features (*e.g.*, headings, pictures), 6 participants provided additional details such as the width of the margins and the size and locations of the images and blocks of text.

Four participants provided unprompted feedback that they liked the document layout knowledge provided by the exploration mode. P6, for example, compared this advantage of the finger-based approach to a traditional screen reader:

*“You have a perspective of the document layout—how many columns, where the graphics are located, the heading, and things like distribution of the text itself. [...] When you use screen readers, you don’t have any idea about that, you just get the text, you just get the content, but you don’t have any direct access or idea of the document layout or things like that.”* (P6)

P15 was particularly excited about the idea, using the exploration mode to identify the size and locations of images and blocks of text, and speculating based on their relative positions that *“maybe [this block of text is] a description of the picture. I always wonder things like that.”* In contrast, P12 stated that he didn’t see a use for spatial information in most situations: *“Not for blocks of text, but [...] for diagrams or for maps it might be, because that’s the only time that you actually need spatial orientation on a page.”* He felt that a system that could automatically process a page and abstract the layout would be preferable. Further investigation is needed to evaluate how much this additional spatial information impacts comprehension or

document understanding, as well as how to best present that information to the user via audio or haptic feedback. We return to this point in the Discussion section.

**3.2.7. Summary of Study I Findings.** Audio and haptic guidance resulted in relatively similar user performance, although audio may offer an accuracy advantage for line tracing with some documents (it was significantly better than haptic for the magazine document). Although the majority of participants preferred haptic guidance, the overall split in preference reflects contradictions found in previous research [Stearns et al. 2014; Shilkrot et al. 2014; Shilkrot et al. 2015]. Open-ended comments also highlight the tradeoffs of the two types of guidance, such as the interference of audio guidance with speech output and the potential for desensitization with haptic guidance. Finally, while several participants appreciated the direct access to layout information provided with HandSight’s exploration mode, and the lower learning curve of HandSight compared to braille, important concerns arose about ease of use and the amount of concentration required.

#### 4. STUDY II: PRELIMINARY USE OF A PROOF-OF-CONCEPT PROTOTYPE

Following the in-depth comparison of audio and haptic finger guidance in Study I, we recruited 4 participants to return and provide qualitative feedback on a proof-of-concept wearable prototype. These follow-up sessions allowed us to collect preliminary evidence of: (1) the extent to which a blind reader can use a finger-mounted camera and directional guidance system to explore and read a printed document, and (2) the strengths and weaknesses of finger-reading versus a mobile scanner and screen reader.

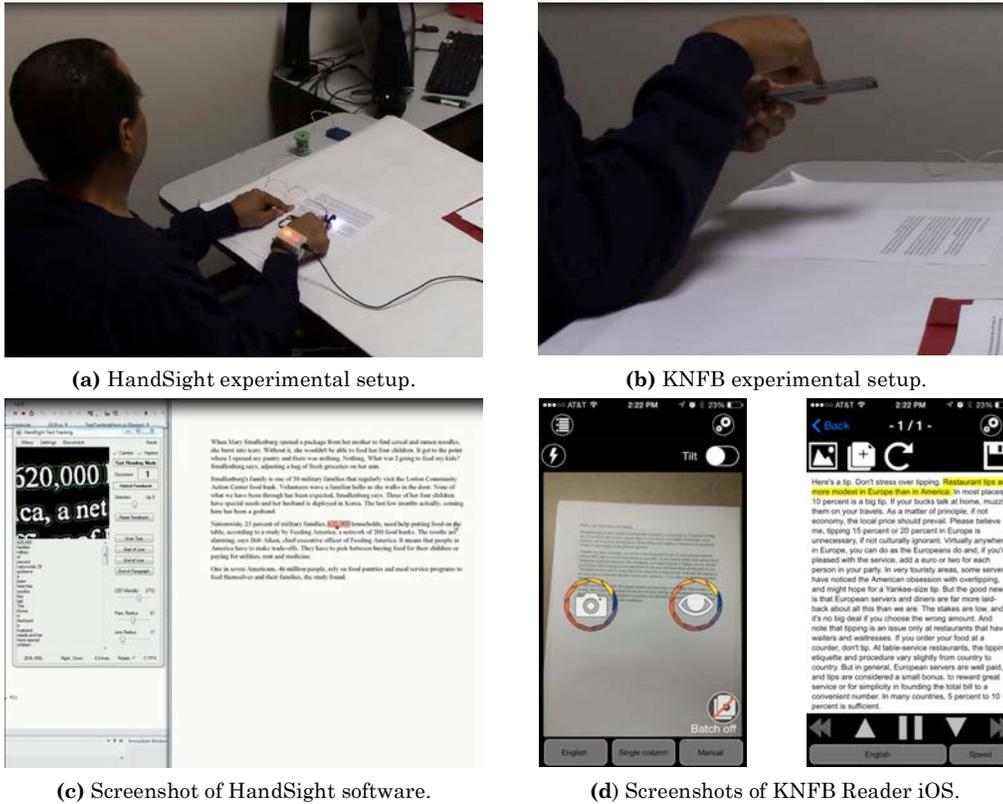
##### 4.1. Method

Participants explored and read printed documents using a proof-of-concept finger-mounted camera system, followed by KNFB Reader iOS, a popular mobile document reader. This was not intended to be a controlled comparison of the two technologies, but instead allowed for preliminary user experience feedback.

**4.1.1. Participants.** We randomly selected 4 participants from Study I to return for this follow-up study, with the constraint that they represent a mix of preferences for haptic and audio directional guidance. Study II was conducted shortly after Study I was completed, with participants returning between 1 and 3 weeks after their initial session. Participants’ durations of blindness varied from 2 to 30 years, but none were congenitally blind. Only one participant (P12) had experience with KNFB Reader iOS. Refer to Table 1 for demographic information and to Table 4 for experience with specific technologies, including KNFB Reader iOS. As with Study I, participants were compensated for their time and transportation costs.

ID	Study I Feedback Preference	Frequency of Braille Use	Frequency of Screen Reader Use	Familiar with KNFB Reader iOS?
P10	No Preference (Tested Audio)	5	1	No
P11	Haptic	4	1	No
P12	Audio	4	5	Yes
P19	Haptic	4	4	No

**Table 4:** Study II participants; IDs are carried over from Study I. Comfort levels ranged from 1-5, with 1 indicating “very uncomfortable” and 5 indicating “very comfortable”.



**Figure 10:** Study II experimental setup. (a) The HandSight test apparatus consisted of a desktop computer running a custom reading program, stereo speakers, a finger-mounted camera system, and the haptic feedback device from our first study. Participants were asked to read through two documents using our prototype system. (b) The KNFB experimental setup consisted simply of an iPhone with the KNFB Reader iOS app. Participants were asked to read three documents using the app. (c) A screenshot of HandSight’s OCR interface (this was not shown to the participant and used only by the experimenter). (d) Two screenshots of KNFB Reader iOS: (left) the ‘capture’ interface helps users orient the phone’s camera to take a photo of the target document; (right) the digitized document screen-reading interface.

**4.1.2. Apparatus.** The proof-of-concept HandSight prototype consisted of a desktop computer running custom software, external speakers, a finger-mounted camera, and the haptic device from Study I (Figure 10). The camera was a self-illuminated Awaiba NanEye 2C CMOS camera and LED ring (~40 fps, 90° square field of view, 250x250 pixels, 2.4mm diameter), embedded in an adjustable ring and positioned above the finger to point down at the page (Figure 1b). The camera was positioned 1–2cm above the page, and had a field of view approximately 1.5cm across (2–3 lines of text). These numbers varied somewhat depending on the participant’s hand position.

As with Study I, the software provided two modes of interaction: *exploration* and *reading*. Exploration mode provided the same feedback as in Study I, except that the prototype system did not detect images; as such, documents used in Study II did not include images. To ensure that the flute sound did not stop between individual characters or lines of text, the system first blurred the text using a blur radius that was manually calibrated prior to beginning the exploration tasks. The audio and

haptic cues in reading mode were identical to those in Study I, with text-to-speech output using the IVONA Voice for Windows speech synthesis engine.<sup>10</sup> Exploration and reading events were logged with timestamps, but we could not log precise finger-trace data as we had done with the iPad in Study I.

The software processed each video frame from the camera using OpenCV<sup>11</sup>, an open-source computer vision and image-processing library. With each frame, we applied four preprocessing algorithms. First, to correct radial distortion from the camera lens, we used standard camera calibration algorithms [Hartley and Zisserman 2003]. Second, to reduce noise, perform binarization necessary for OCR, and adapt to uneven lighting from the LED, we filtered each frame using an adaptive threshold in a sliding window. Third, to reduce false positives, we performed a connected component analysis and removed components with areas too small or aspect ratios too narrow to be characters. Finally, to correct for finger rotation, we blurred the image to efficiently group the components into likely lines of text, then extract the minimum-area bounding rectangle for each new component. We used the estimated orientation of this rectangle to correct for camera rotation, inverting it so that the lines of text were parallel to the  $x$ -axis. This process is similar to that described in [Safabakhsh and Khadivi 2000].

To simplify sensing for this proof-of-concept prototype, we assumed that a complete image of the page was available to the system in advance. The software then estimated the current finger location by performing OCR on the visible text and matching it to the known content of the page. We used the Tesseract OCR library<sup>12</sup> for text detection and recognition of each preprocessed frame, then compared the results to the pre-computed document text. For efficiency, we tracked character motion between frames and only performed OCR when sufficient motion had occurred or when the system was unable to reliably estimate the current location (allowing us to achieve an average processing rate of 20–30 fps). Because the camera’s field of view was large enough to encompass multiple partial words across 2-3 lines of text (Figure 10c), the system did not generally encounter difficulty distinguishing the locations of repeated words. The likelihood of this potential problem was further reduced using recent location estimates and the motion of the user’s finger to resolve conflicts. We tracked the current line of text using the camera’s estimated motion and the known content of the page, and only provided text-to-speech feedback when the user advanced on the current line. In order to provide a smooth reading experience, it was not possible to skip or repeat words. Although this enforced sequential reading of the text, it mitigated several potential sources of confusion that would have arisen had we allowed rereading or moving between lines. The software detected that the user had reached the start or end of a line or paragraph using the known content of the page, and provided the same audio cues as in Study I. Also, as with Study I, the speed of the text-to-speech feedback was adjusted to match the user’s finger speed.

The test apparatus for the second part of the study consisted of the KNFB Reader iOS application running on an iPhone 5S with the VoiceOver feature enabled. To take a picture, users tapped on the left side of the screen to select the “Take Picture”

<sup>10</sup> <http://www.ivona.com/us/for-individuals/voices-for-windows/>

<sup>11</sup> <http://opencv.org/>

<sup>12</sup> <https://code.google.com/p/tesseract-ocr/>

button, and then double-tapped the button to capture an image. The software played a shutter sound to inform the user that the picture was captured successfully, and then immediately began reading any recognized text.

**4.1.3. Procedure.** These exploratory study sessions lasted 1–2 hours. The participant first used HandSight with his or her preferred directional guidance method from Study I. As with Study 1, training and testing documents were selected from the *Iowa Test of Basic Skills*. For training, the experimenter first re-introduced exploration mode and asked the participant to explore a plain document for up to three minutes. Participants were directed to count the number of paragraphs and columns, and to note the size and position of the margins. The experimenter then re-introduced reading mode's audio cues and directional guidance, and helped the participant read the training document, providing verbal or physical guidance if necessary. The training tasks lasted 10–15 minutes. After training, participants explored and read one single-column test document, with the experimenter providing verbal assistance only if the participant was unable to proceed. Afterward, participants answered questions about the layout of the document, three multiple-choice questions to judge comprehension, and subjective questions about the experience. We did not use a magazine-style document because HandSight does not currently support two-column documents.

Following the use of HandSight, the experimenter introduced KNFB Reader iOS: how to position the phone's camera over a page, take a picture, and listen to the recognized text. Although the KNFB Reader iOS application included a spoken field of view report to assist with framing a document, we did not evaluate this feature due to time constraints and because it was not the focus of this study. Participants were allowed to repeat this process up to three times with a single-column training document, with verbal or physical guidance as needed. This training task lasted 10–15 minutes. Participants then read two test documents unassisted: a single-column document (from the Iowa Test of Basic Skills) and a two-column magazine document (from USA Today) similar to those read in Study I but without images. KNFB Reader iOS advertises support for multicolumn formats, and the procedure for capturing and reading the two types of document was identical. If the participant was unsatisfied with the reading result, they were allowed one additional attempt per document. Participants answered multiple-choice comprehension questions after the single-column document and summarized the content of the two-column document. Finally, participants reported on their experience using the application. See Appendix B for the full text of the subjective questionnaires for both HandSight and KNFB Reader iOS.

## 4.2. Findings

Our findings are exploratory, including general observations about how participants approached the reading tasks, and subjective responses to both our proof-of-concept implementation and KNFB Reader iOS. While the focus is on qualitatively describing experiences with the technologies, we include performance statistics such as reading speed, line finding time, and number of skipped words.

**4.2.1. Overall Experience.** All four participants completed the reading tasks, but with varying levels of success (Table 5). P10 read slowly and required frequent verbal and physical intervention by the tester to adjust hand position and answer questions

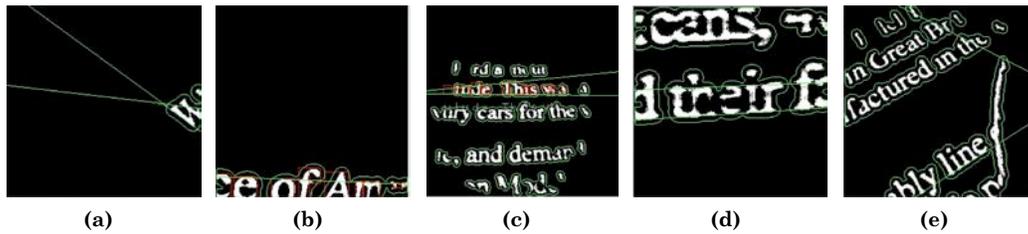
Participant Identifier	P10	P11	P12	P19	Mean
<b>Guidance Type</b>	Audio	Haptic	Audio	Haptic	N/A
<b>Start of Text</b>	5	2	5	1	3.3
<b>Start of Line</b>	5	2	2	2	2.8
<b>End of Line</b>	5	5	5	5	5.0
<b>End of Paragraph</b>	3	5	5	5	4.5
<b>Line Tracing</b>	2	2	3	2	2.3
<b>Understanding Cues</b>	5	5	5	3	4.5
<b>Reading and Understanding</b>	3	3	5	4	3.8
<b>Mean Ease of Use Rating</b>	4.0	3.4	4.3	3.1	3.7
<b>Average Reading Speed per Line (wpm)</b>	18.4 (SD=5.5)	56.6 (SD=16.4)	60.2 (SD=11.1)	44.9 (SD=17.1)	45.0
<b>Average Line Finding Time per Line (s)</b>	30.5 (SD=24.4)	8.8 (SD=5.6)	7.3 (SD=5.0)	18.0 (SD=12.7)	16.15
<b>Time to Read Full Document (s)</b>	1493	469	409	717	772
<b>Comprehension Questions Score</b>	2/3	3/3	3/3	3/3	2.75/3

**Table 5:** Top: Ease of use responses while using the HandSight prototype. Responses range from 1 - very difficult to 5 - very easy. Bottom: Performance metrics from the HandSight reading task. The document for this task consisted of 282.6 words (normalized to 5 character length) across 17 lines.

about the directional audio cues. P11 and P19 read more quickly, and only needed infrequent verbal reminders (P11 was reminded once about hand position, and P19 was reminded once about hand position and the procedure for finding the start of a line). P12, who was very comfortable with both braille and screen readers, read the fastest, at 60.2 wpm, and did not require any assistance. Only P10 failed to answer all three comprehension questions correctly, likely due to decreased attention to the content while struggling to complete the task.

Comments were similarly mixed. P19 was enthusiastic about the concept, stating: *“I’m very pleased and excited about the system. I think it could make a great difference in my life.”* P12 was more critical, finding the approach to be slower than expected: *“It seems like a lot of effort for reading text.”* P10, P11, and P19 were all learning to read braille at the time of the study, and P11 and P19 found the reading experience using HandSight to be easier than braille for reasons similar to those expressed in Study I (e.g., lower learning curve, less to remember). P10 stated that braille and finger-reading were both difficult at times, requiring too much concentration to read quickly or fully comprehend the text. P12, who had the most braille experience, found HandSight to be “much worse” than braille and “somewhat worse” than other technologies for reading printed documents. In addition to commenting on the ease of following a line of braille text due to the tactile feel of the dots and the lack of layout issues such as multiple columns, P12 said that he typically scans printed documents to read on his computer or mobile device, an approach he finds faster compared to HandSight and one that does not require use of both hands.

**4.2.2. Cognitive Load.** Although they were able to complete the reading task, all participants expressed concern about the level of concentration required to interpret the directional guidance and other audio cues while listening to synthesized speech. P11, for example, commented on the difficulty of remembering how to map the haptic guidance to up/down movement: *“it gets you a little confused sometimes, especially if you was [sic] into reading the story and you forget which one was the vibration for moving up and which one was for moving to the bottom.”* P11 also commented on the focus and practice required, concluding that it would be difficult to use, *“if you’re tired, if you’ve had a long day.”* More practice with the device may address some of



**Figure 11:** Examples of situations where HandSight was unable to provide feedback. All images have been preprocessed to emphasize text and highlight baselines for the current line. (a), (b) Not enough text is visible in the margins to provide directional guidance. (c) The camera position changed after calibration, and too far from the page to reliably recognize text. (d) The camera is moving too quickly, blurring the text and reducing the frame rate of the recognition algorithms. (e) The user’s middle finger is in the camera’s field of view, preventing correct segmentation of the lines of text.

these issues, though interaction design changes are also likely needed (e.g., more intuitive and responsive directional cues to reduce required concentration on line tracing task, efficient rereading to enhance comprehension).

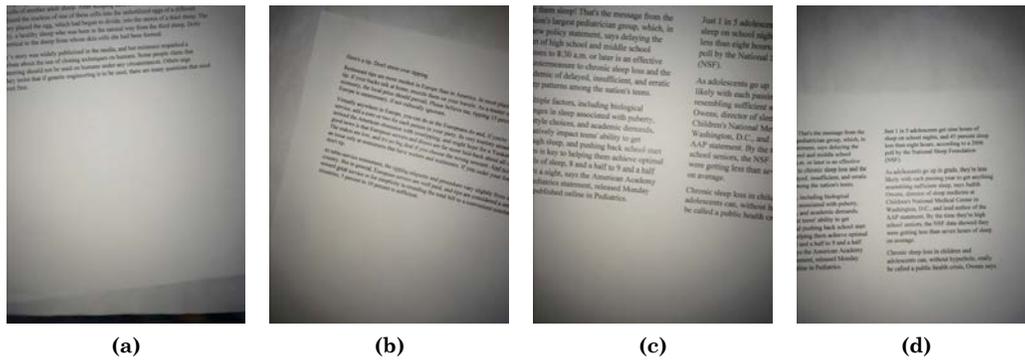
Technical limitations with the prototype may have exacerbated cognitive load issues. Although our algorithms ran at approximately 30 fps on average, they tended to run more slowly after rapid finger movements. This limitation caused a noticeable lag at times, which P11 and P19 reported required more concentration. P19, for example, commented that after the start-of-line audio cue there was sometimes a delay before the speech began, causing problems: *“I wasn’t getting that in my head to just wait for the delay. I started moving my finger”*.

**4.2.3. Physical Design.** Three participants identified limitations with the prototype’s physical design. The primary issue stemmed from the camera placement: for the text to be an appropriate size and orientation within the camera’s field of view, participants’ hands needed to be held at a specific angle. Although the camera’s placement on the finger was adjusted at the start of each study session, it could not easily be readjusted. Participants thus had to hold their hand at nearly the same angle throughout the study. Two participants reported that this position was too uncomfortable for extended use, suggesting that the physical design will need to improve in future versions and/or the camera location should be easily adjustable. We also identified the need for feedback when the system loses its position in the text or is unable to recognize visible text in reading mode (Figure 11 shows examples); this occurred when the hand position changed too much or, more commonly, when the participant moved into the upper or lower margins of the document.

**4.2.4. HandSight vs. KNFB Reader iOS.** While the study did not offer a controlled comparison of HandSight and KNFB Reader iOS, we can draw preliminary conclusions about tradeoffs between the two. Even without KNFB Reader iOS’s document-framing guidance enabled, participants unanimously preferred it to HandSight, with three participants rating it as 5 – *much better* and one as 4 – *somewhat better*. The primary reason was the fluidity of the reading experience after capturing an image with KNFB—the application read the full document quickly and participants were able to concentrate solely on the content of the passage. For example: *“It just did it all for you, that way you just listen to what it’s saying and then take in the details”* (P11). P12, who had previous experience with KNFB, also stated:

Participant Identifier	P10	P11	P12	P19	Mean
Document 1: Number of Attempts	2	2	1	2	N/A
Document 2: Number of Attempts	2	2	1	2	N/A
Document 1: Total Time (s)	230	198	93	225	187
Document 2: Total Time (s)	138	219	89	137	146
Document 1: Text Lost (%)	29.7%	48.6%	0.4%	10.4%	22.3%
Document 2: Text Lost (%)	51.5%	0.0%	0.0%	51.5%	25.8%
Comprehension Questions Score	1/3	3/3	3/3	3/3	2.5/3

**Table 6:** Performance metrics from the KNFB Reader iOS reading tasks. The amount of text lost includes both cropped and misrecognized words, and the percentages indicate the best performance out of the two attempts participants were allowed for each document.



**Figure 12:** Examples of cases where the KNFB Reader iOS application failed to fully capture the content of a document due to partial visibility or excessive rotation.

*“I like that the text is immediately available to use for other purposes [...] I can go back and review the text letter by letter if I need to.”* The average reading time was only 187 seconds for the first document and 146 seconds for the second, even with the two attempts that participants were allowed, as compared to an average time of 772 seconds to complete the reading task with HandSight (Tables 5 and 6).

Although participants preferred KNFB Reader iOS overall, the process of capturing an image was not always straightforward without the document-framing guidance. P11, for example, said: *“It was easy to read it once you got it right, but it was difficult to center [the camera] in order to get the whole text”* (P11); see Figure 12 for examples of images captured by participants during this study. With the exception of P12, all participants required a second attempt to capture each document, and even with a second attempt part of the document was frequently omitted. Although accuracy varied across participants and attempts, approximately one quarter of the documents’ content was missed on average (see Table 6).

Document comprehension appeared to be similar to reading with HandSight, with three participants answering all comprehension questions correctly (Tables 5 and 6). However, even when participants were able to understand the main points of a document, the reading experience was not always smooth due to missing text: *“It’s not always easy to know if I have the entire page. That was a problem with the first test document. While it was still understandable, I clearly lost some of the text”* (P12).

**4.2.5. Summary of Study II Findings.** While this study was not meant to be a controlled comparison of finger-based reading versus a mobile scanner, it offers some guidance for future studies. HandSight provided more immediate access to text content than KNFB Reader iOS, but was much slower and was perceived as requiring a greater level of concentration. Once the document was successfully scanned, KNFB Reader iOS offered a faster and smoother reading experience, and was preferred by all participants. HandSight provided additional information about the spatial layout of documents, but further investigation is needed to determine the impact that may have on document understanding.

## 5. DISCUSSION

Our findings highlight tradeoffs between haptic and audio directional guidance for finger-based reading. We also reflect on the feasibility of finger-based reading compared to existing methods, and outline ideas for iteration on HandSight's design.

### 5.1. Audio versus Haptic Directional Guidance

For blind users, effective finger guidance is critical for line-by-line reading, and therefore directly impacts the feasibility of the finger-based reading approach. Although there were few statistically significant differences between audio and haptic finger guidance in Study I, some tradeoffs emerged. For the magazine documents, audio guidance resulted in significantly more accurate line tracing than haptic guidance. The exact cause is unclear. To scaffold participants in learning how to do finger-based reading, we always presented plain documents before the more complex magazine documents. That audio was more accurate than haptic for the magazine document thus suggests that haptic may have a steeper learning curve, participants have become desensitized to the haptic vibration over time, or that haptic is somehow not as effective with complex document layouts.

In terms of subjective responses, our findings reflect the conflicting results seen in prior work [Stearns et al. 2014; Shilkrot et al. 2014; Shilkrot et al. 2015]. Out of 19 participants, 11 preferred haptic, 7 preferred audio, and one was undecided. One downside of the audio guidance is that it occupies the same channel as the speech output, which made it difficult for some participants to concentrate on the text-to-speech synthesis. Twelve participants, half of whom even preferred audio guidance, commented on this issue. For haptic guidance, the potential issue of desensitization or numbness arose even in this short study, suggesting that a longer-term evaluation will be important.

We also encountered disagreement over how audio and haptic cues should map to up/down direction, which could have impacted results. The mappings used in our studies were the result of pilot testing and our experiences in [Stearns et al. 2014]. For audio, we used high pitch to indicate that the finger should move up and low pitch for down. For haptic, the vibration motor on the underside of the finger indicated downward movement, while the top vibration motor indicated upward movement (in essence, pulling the finger). While the majority of Study I participants were satisfied with these mappings, 4 felt audio should be reversed and 3 felt haptic should be reversed. More work is needed to identify which mapping is best for both audio and haptic, or whether additional training time would mitigate the issue. Ultimately, this setting may need to be user-configurable. Future work should also investigate alternative feedback approaches (*e.g.*, the pitch of the speech synthesis could provide directional guidance).

## 5.2. Feasibility of a Finger-Based Reading Approach

We had expected a finger-based reading approach such as HandSight or Shilkrot *et al.*'s FingerReader [Shilkrot et al. 2014; Shilkrot et al. 2015] to offer many advantages over mobile-based scanners for reading printed text: access to spatial layout information, direct as opposed to sequential access to text on the page, reduced camera framing issues, and, compared to crowdsourced approaches (*e.g.*, [Bigam et al. 2010], BeMyEyes), real-time OCR. However, while we observed some of these advantages, important concerns also arose. Here, we reflect on the feasibility of finger-based reading, incorporating ideas for future work throughout.

**5.2.1. Document Layout and Spatial Awareness.** A primary motivation for investigating HandSight's finger-based reading approach was to provide users direct access to spatial layout information. Our exploration mode provided audio cues to indicate text, pictures, or white space beneath the user's finger. While 4 participants in Study I commented positively and unprompted on this information, one participant was strongly against the idea, feeling that software that could automatically process a document's layout to extract content would be preferable in most situations. The difficulties encountered by the participant who was removed from our dataset in Study I also highlight an unexpected but important potential for confusion: some users, particularly those who are congenitally blind, may have an inaccurate or incomplete understanding of basic document structures (*e.g.*, columns, margins) simply because they have never encountered them. In that participant's case, he was not familiar with the notion of columns, which led to confusion. Future work should explore the relationship between a user's spatial abilities and their proficiency in exploring a document or responding to finger guidance.

While exploration mode helped participants understand a document's layout (*e.g.*, number of images), distinguishing a gap between paragraphs versus columns was particularly challenging. Both types of gaps were indicated by white space, but participants were frequently unable to determine whether the white space occurred between two paragraphs or between two columns. We intended for paragraph gaps to be distinct from column gaps by the direction in which the finger is moving—vertically for paragraphs or horizontally for columns. However, without sight, many participants tended to move their finger more diagonally, drifting accidentally between paragraphs and columns. This challenge could be addressed by designing cues to identify the horizontal and vertical edges of a block of text.

Finally, we did not evaluate the potential utility of layout information for blind readers. And, arguably, for half the documents we used (the plain text documents), spatial information offered little benefit. The finger-based reading approach may be more beneficial for other types of documents, particularly those with inherent spatial characteristics such as maps or graphs.

**5.2.2. Cognitive Load and Physical Effort.** Our studies indicate that line-by-line reading incurs high mental and physical effort. The reader must simultaneously attend to directional guidance, document events (*e.g.*, start and end of line), and the synthesized speech content. Study II, in particular, highlighted the increased concentration and physical dexterity required to use HandSight compared to KNFB Reader iOS. This issue of physical effort confirms previous findings from a much smaller study (3 participants) [Shilkrot et al. 2015]. With more practice, HandSight should not require as much effort to use, and, if the technology provides enough

benefit, the need for this practice is not necessarily a barrier to adoption—braille and the Optacon [Millar 2003; Hislop et al. 1983], for example, require extensive practice. However, a multi-session study would be needed to assess just how much practice is needed and how efficiently experienced users can read with a finger-based approach.

**5.2.3. Camera Placement.** Whether they use crowdsourcing or automated OCR, both mobile document scanning approaches (e.g., KNFB Reader iOS) and body-mounted solutions (e.g., OrCam) require a global image of the document, properly aligned and in focus within the camera’s field of view. All participants in Study II reported at least some difficulty with this type of image capture using KNFB Reader iOS, but we had not introduced them to KNFB’s document-framing feature. That feature, along with findings from blind photography research (e.g. [Vázquez and Steinfeld 2012; Jayant et al. 2011; Cutter and Manduchi 2015]), should reduce the issue. The global image captured by KNFB also allowed for more fluent text-to-speech than with HandSight, which participants valued. At times, however, our own use of KNFB Reader iOS and observations of participants showed that this fluency can provide a false sense of confidence. That is, it is not always clear from the speech output if a part of the document is missing or the application has parsed and played text blocks in the wrong order.

HandSight’s finger-mounted camera and direct control over text scanning and speech playback may overcome these issues to some extent, but Study II showed it also introduces new camera placement challenges. For example, participants frequently encountered difficulties tracing lines near the upper and lower margins because the system could not provide directional guidance when no text was visible.

In future work, we plan to explore hybrid methods that may combine a body- or head-mounted camera with a finger-mounted one, potentially overcoming the weaknesses of each and supporting a wider range of reading situations. A body-mounted camera could capture complete documents and allow for efficient, fluent reading using a screen-reader interface and relative exploration of content (e.g., swipes). At the same time, the finger-camera interface could provide knowledge about the document layout, acting as a cursor to quickly search through the content or provide contextual information. It would be useful to compare how well having both types of interaction works compared to only the global, relative interaction or the finger-based interaction.

**5.2.4. Physical Design and Social Acceptability.** Physical design and social acceptability influence the adoption of wearable technology [Rico and Brewster 2009; Rico and Brewster 2010; Profita et al. 2013]. While our early HandSight prototype is bulky, future versions could be substantially reduced in size since the underlying technology (i.e., the endoscopic camera) is extremely small. Still, whether blind users are interested in wearing a finger-mounted device for accessibility is an open question. Social acceptability could also change how users feel about the haptic line guidance compared to the audio guidance in practice. The majority of users preferred haptic guidance in Study I, but even in future iterations of the physical design, the haptic vibration motors would likely add bulk compared to audio alone. These issues are not unique to HandSight, and the question of where users will feel most comfortable having a camera mounted on their body (if at all) should be explored in future work.

**5.2.5. Target Users.** While our prototype was designed to support totally blind users, the question of who may benefit most from a finger-based reading approach remains open. Low vision users, for example, may benefit from the direct access and physical gestures that a finger-based reading approach provides, without finding the line tracing as time consuming as for a totally blind reader. We recruited one low vision pilot participant, who experienced no difficulties with describing the layout of the document or with line finding and line tracing. The device could then act as a more portable alternative to closed-circuit television (CCTV) magnifiers, automatically processing the words and providing additional information about the text upon request (*e.g.*, font, spelling). Further investigation, however, is needed to explore this possibility and how it is received compared to commonly used magnifiers.

### 5.3. Design Iteration

In addition to the future work mentioned above—such as investigating the utility of spatial layout information, conducting a longer-term study, and evaluating HandSight with low-vision users—our findings lead to several design revisions that may improve blind users’ experience with HandSight.

We designed the speech interface to adapt to the user’s finger speed so that they could easily control the rate of feedback. Some participants liked this feature, but others found it to be choppy when compared to the continuous speech feedback of screen reader software, noting that it was difficult to identify the end of a sentence. More fluid speech feedback and additional audio cues to mark punctuation could ease the reading experience.

An important observation from Study II is that a finger-based reading device should provide an easy way of determining when text is no longer contained within the camera’s field of view. Participants occasionally confused situations where the system could not provide guidance (not enough text in the frame) with being correctly centered over the current line and not receiving directional feedback. To address this issue and provide users with more information while reading, document exploration and reading modes could be integrated. In doing so, however, we must take care not to further increase cognitive load and distract from the content of the text.

To reduce the image capture issues seen in Study II, another possibility is to redesign the physical prototype to either move the sensor farther away from the text (as with FingerReader [Shilkrot et al. 2014; Shilkrot et al. 2015], which is on the upper part of the finger) or to use a wide-angle lens. These options could expand the camera’s field of view, for example, allowing users to drift farther away from a line before the text is lost.

Finally, our prototypes only allowed users to continue reading forward and did not support backtracking, rereading, or jumping to an arbitrary location in the text. Study I focused on sequential line-tracing guidance, but it would be interesting to implement and evaluate these additional reading actions.

### 5.4. Limitations

Using an iPad rather than a physical prototype to compare haptic and audio line guidance in Study I was a conscious study design choice, allowing us to bypass technical challenges in implementing a real-time prototype and to focus on the user experience and collect precise line traces. A limitation of this choice, however, is that the experience of reading with a physical prototype and paper may be different. As

well, the font size and document layout for Study I were constrained to two specific formats, which may not fully represent the variety of real-world documents that users may encounter. In Study II, an important limitation of the proof-of-concept prototype is that we assumed that the content of the page was known prior to beginning reading, and constrained the system to allow participants to read text sequentially from left to right and top to bottom. These choices simplified how the system provided finger guidance—it only needed to estimate the finger location on the page and provide upward or downward guidance to return to the last known line. However, these artificial limitations also disregarded some of the potential advantages of a finger-based reading approach, such as re-reading or jumping to arbitrary locations. We also asked participants to hold their hand in a specific position for Study II, constraining their natural behavior when using a device such as ours. Study II was not meant to offer a controlled comparison of KNFB Reader iOS and HandSight, but limitations even for gathering exploratory feedback include that we did not evaluate the document framing feedback of KNFB Reader iOS, and that only one participant had previous experience with KNFB Reader iOS (all 4 had used finger-based reading). While a more controlled comparison is thus needed, it is important to note that participants still identified many strengths of KNFB Reader iOS. Finally, while we focused on blind readers, it would be interesting to expand the evaluation of finger-based reading to users with a wider range of visual abilities.

## 6. CONCLUSION

We conducted an in-depth study with 19 blind participants comparing audio and haptic cues for directional guidance to support finger-based reading. Our findings showed similar performance and user preference between the two types of guidance, although audio resulted in significantly more accurate line tracing for some tasks. Subjective feedback was split, but suggests that haptic guidance may be slightly preferred. In addition, our findings highlighted general strengths and weaknesses of a finger-based reading approach, such as improved understanding of a document's layout and the difficulty encountered by blind users in accurately tracing a line of text with a finger. In follow-up sessions where 4 of the participants used a proof-of-concept finger-reading prototype as well as KNFB Reader iOS, the mobile scanner was seen as offering a more fluent reading experience. Ultimately, a finger-based reading approach may be best suited to material that is inherently spatial, such as maps or graphs, whereas existing applications that capture a global image of the document for text-to-speech (e.g., KNFB Reader iOS) may be preferred for text-heavy material. Future work should investigate this possibility, as well as assess the potential of finger-based reading for low-vision users, for whom precise directional finger guidance may not be necessary.

## ACKNOWLEDGMENTS

We thank our participants and the Maryland State Library for the Blind and Physically Handicapped. This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award W81XWH-14-1-0617.

## REFERENCES

- Chieko Asakawa, Hironobu Takagi, Shuichi Ino, and Tohru Ifukube. 2003. Maximum listening speeds for the blind. In *Proceedings of the 2003 International Conference on Auditory Display*. 276–279.
- Jeffrey P. Bigham et al. 2010. VizWiz: nearly real-time answers to visual questions. In *Proceedings of the*

- 23rd annual ACM symposium on User interface software and technology - UIST '10*. New York, New York, USA: ACM Press, 333–342. DOI:<http://dx.doi.org/10.1145/1866029.1866080>
- Jeffrey P. Bigham, Richard E. Ladner, and Yevgen Borodin. 2011. The design of human-powered access technology. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '11*. New York, New York, USA: ACM Press, 3. DOI:<http://dx.doi.org/10.1145/2049536.2049540>
- Nicoletta Di Blas, Paulo Paolini, and Marco Speroni. 2004. Useable accessibility to the web for blind users. In *Proceedings of the 8th ERCIM Workshop: User Interfaces for All*.
- J.C. Bliss. 1969. A relatively high-resolution reading aid for the blind. *Man-Machine Syst. IEEE Trans.* 10, 1 (March 1969), 1–9. DOI:<http://dx.doi.org/10.1109/TMMS.1969.299874>
- Yevgen Borodin, Jeffrey P. Bigham, Glenn Dausch, and I. V. Ramakrishnan. 2010. More than meets the eye: a survey of screen-reader browsing strategies. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A) - W4A '10*. New York, New York, USA: ACM Press, 1. DOI:<http://dx.doi.org/10.1145/1805986.1806005>
- Erin Brady, Meredith Ringel Morris, Yu Zhong, Samuel White, and Jeffrey P. Bigham. 2013. Visual challenges in the everyday lives of blind people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. New York, New York, USA: ACM Press, 2117–2126.
- Karl Burklen. 1932. *Touch reading of the blind*, New York: American Foundation for the Blind.
- M. Capp and P. Picton. 2000. The optophone: an electronic blind aid. *Eng. Sci. Educ. J.* 9, 3 (2000), 137–143. DOI:<http://dx.doi.org/10.1049/esej:20000306>
- Xiangrong Chen and A.L. Yuille. 2004. Detecting and reading text in natural scenes. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. II–366–II–373 Vol.2. DOI:<http://dx.doi.org/10.1109/CVPR.2004.1315187>
- Franklin S. Cooper, Jane H. Gaitenby, and Patrick W. Nye. 1983. *Evolution of reading machines for the blind: Haskins Laboratories' research as a case history*, Haskins Laboratories.
- Andrew Crossan and Stephen Brewster. 2008. Multimodal trajectory playback for teaching shape information and trajectories to visually impaired computer users. *ACM Trans. Access. Comput.* 1, 2 (October 2008), 12:1–12:34. DOI:<http://dx.doi.org/10.1145/1408760.1408766>
- Michael Cutter and Roberto Manduchi. 2015. Towards Mobile OCR: How To Take a Good Picture of a Document Without Sight. *Proc. ACM Symp. Doc. Eng. ACM Symp. Doc. Eng.* 2015 (January 2015), 75–84.
- EE Fournier D'Albe. 1914. On a type-reading optophone. *Proc. R. Soc. London. Ser. A* 90, 619 (1914), 373–375.
- Sarah L. Dowhower. 1989. Repeated reading: research into practice. *Read. Teach.* 42, 7 (1989), pp. 502–507.
- Leah Findlater et al. 2015. Supporting Everyday Activities for Persons with Visual Impairments Through Computer Vision-Augmented Touch. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. 383–384.
- Vincent Gaudissart, Silvio Ferreira, Céline Thillou, and Bernard Gosselin. 2004. SYPOLE: mobile reading assistant for blind people. In *9th Conference Speech and Computer (SPECOM)*. St. Petersburg, Russia.
- Nicholas A. Giudice, Hari Prasath Palani, Eric Brenner, and Kevin M. Kramer. 2012. Learning non-visual graphical information using a touch-based vibro-audio interface. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '12*. New York, New York, USA: ACM Press, 103.
- Louis H. Goldish and Harry E. Taylor. 1973. The Optacon: a valuable device for blind persons. *New Outlook Blind* (November 1973).
- Susumu Harada, Hironobu Takagi, and Chieko Asakawa. 2011. On the Audio Representation of Radial Direction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. New York, NY, USA: ACM, 2779–2788. DOI:<http://dx.doi.org/10.1145/1978942.1979354>
- Richard Hartley and Andrew Zisserman. 2003. *Multiple View Geometry in Computer Vision*, Cambridge University Press.
- Sharon A. Haymes, Alan W. Johnston, and Anthony D. Heyes. 2002. Relationship between vision impairment and ability to perform activities of daily living. *Ophthalmic Physiol. Opt.* 22, 2 (March 2002), 79–91. DOI:<http://dx.doi.org/10.1046/j.1475-1313.2002.00016.x>

- David W. Hislop, B.L. Zuber, and John L. Trimble. 1983. Characteristics of reading rate and manual scanning patterns of blind Optacon readers. *Hum. Factors J. Hum. Factors Ergon. Soc.* 25, 4 (August 1983), 379–389.
- David W. Hislop, B.L. Zuber, and John L. Trimble. 1985. Text-scanning patterns of blind readers using Optacon and braille. *J. Rehabil. Res. Dev.* (1985), 54–65.
- Sture Holm. 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* (1979), 65–70.
- Bill Holton. 2014. KNFB Reader for iOS: does this app live up to all the hype? *AFB AccessWorld Mag.* 15, 11 (2014).
- Chandrika Jayant, Hanjie Ji, Samuel White, and Jeffrey P. Bigham. 2011. Supporting blind photography. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*. ASSETS '11. New York, NY, USA: ACM, 203–210. DOI:<http://dx.doi.org/10.1145/2049536.2049573>
- Shaun K. Kane, Brian Frey, and Jacob O. Wobbrock. 2013. Access Lens: a gesture-based screen reader for real-world documents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. New York, NY, USA: ACM, 347–350. DOI:<http://dx.doi.org/10.1145/2470654.2470704>
- R. Keefer, Yan Liu, and N. Bourbakis. 2013. The development and evaluation of an eyes-free interaction model for mobile reading devices. *Human-Machine Syst. IEEE Trans.* 43, 1 (2013), 76–91. DOI:<http://dx.doi.org/10.1109/TSMCA.2012.2210413>
- Deborah Kendrick. 2008. A picture becomes a thousand words: an evaluation of the knfbReader mobile. *Am. Found. Blind AccessWorld Mag.* 9, 6 (2008).
- Georgios Kouroupetroglou and Dimitrios Tsonos. 2008. Multimodal accessibility of documents. In *Advances in Human-Computer Interaction*. Vienna, Austria: I-Tech Education and Publishing, 451–470.
- Jonathan Lazar, Aaron Allen, Jason Kleinman, and Chris Malarkey. 2007. What frustrates screen reader users on the web: a study of 100 blind users. *Int. J. Hum. Comput. Interact.* 22, 3 (2007), 247–269. DOI:<http://dx.doi.org/10.1080/10447310709336964>
- Betty Ann Levy. 2001. Text processing: Memory representations mediate fluent reading. *Perspect. Hum. Mem. Cogn. aging Essays honour Fergus Craik* (2001), 83–98.
- Roberto Manduchi and James M. Coughlan. 2014. The last meter: blind visual guidance to a target. In *Proceedings of ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*. To appear.
- Susanna Millar. 2003. *Reading by touch*, Routledge.
- S. Mori, C.Y. Suen, and K. Yamamoto. 1992. Historical review of OCR research and development. *Proc. IEEE* 80, 7 (July 1992), 1029–1058. DOI:<http://dx.doi.org/10.1109/5.156468>
- Elizabeth D. Mynatt and Gerhard Weber. 1994. Nonvisual presentation of graphical user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI '94*. New York, New York, USA: ACM Press, 166–172.
- National Center for Health Statistics. 2012. *National Health Interview Survey 2012 Data Release*.
- National Federation of the Blind Jernigan Institute. 2009. The braille literacy crisis in america: facing the truth, reversing the trend, empowering the blind. (2009).
- Nicolas Noble and Benoit Martin. 2006. Shape discovering using tactile guidance. In *Proceeding of the 6th International Conference EuroHaptics*.
- Uran Oh, Shaun K. Kane, and Leah Findlater. 2013. Follow that sound: using sonification and corrective verbal feedback to teach touchscreen gestures. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. Article 13.
- Peter Parente and Gary Bishop. 2003. BATS: The Blind Audio Tactile Mapping System. *Proc. ACM South East. Conf.* (2003).
- Beryl Plimmer, Peter Reid, Rachel Blagojevic, Andrew Crossan, and Stephen Brewster. 2011. Signing on the tactile line: a multimodal system for teaching handwriting to blind children. *ACM Trans. Comput. Interact.* 18, 3 (August 2011), 17:1–17:29. DOI:<http://dx.doi.org/10.1145/1993060.1993067>
- Benjamin Poppinga, Charlotte Magnusson, Martin Pielot, and Kirsten Rasmus-Gröhn. 2011. TouchOver map. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '11*. New York, New York, USA: ACM Press, 545.
- The Princeton Review. 2000. *ITBS preparation and practice workbook, Glencoe language arts, grade 8*.

- New York, NY: McGraw-Hill/Glencoe.
- Halley P. Profita et al. 2013. Don't mind me touching my wrist. In *Proceedings of the 17th annual international symposium on International symposium on wearable computers - ISWC '13*. New York, New York, USA: ACM Press, 89.
- Julie Rico and Stephen Brewster. 2009. Gestures all around us: user differences in social acceptability perceptions of gesture based interfaces. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '09*. New York, New York, USA: ACM Press, 64. DOI:<http://dx.doi.org/10.1145/1613858.1613936>
- Julie Rico and Stephen Brewster. 2010. Usable gestures for mobile interfaces. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. New York, New York, USA: ACM Press, 887. DOI:<http://dx.doi.org/10.1145/1753326.1753458>
- Ruby Ryles. 1996. The impact of braille reading skills on employment, income, education, and reading habits. *J. Vis. Impair. Blind.* 90 (1996), 219–226.
- R. Safabakhsh and S. Khadivi. 2000. Document skew detection using minimum-area bounding rectangle. In *Proceedings International Conference on Information Technology: Coding and Computing (Cat. No.PR00540)*. IEEE Comput. Soc, 253–258. DOI:<http://dx.doi.org/10.1109/ITCC.2000.844226>
- Huiying Shen and JamesM. Coughlan. 2012. Towards a real-time system for finding and reading signs for visually impaired users. In Klaus Miesenberger, Arthur Karshmer, Petr Penaz, & Wolfgang Zagler, eds. *Computers Helping People with Special Needs SE - 7*. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 41–47. DOI:[http://dx.doi.org/10.1007/978-3-642-31534-3\\_7](http://dx.doi.org/10.1007/978-3-642-31534-3_7)
- Roy Shilkrot, Jochen Huber, Connie Liu, Pattie Maes, and Nanayakkara Suranga Chandima. 2014. FingerReader: a wearable device to support text reading on the go. *CHI '14 Ext. Abstr. Hum. Factors Comput. Syst.*, Vi (2014), 2359–2364. DOI:<http://dx.doi.org/10.1145/2559206.2581220>
- Roy Shilkrot, Jochen Huber, Meng Ee Wong, Pattie Maes, and Suranga Nanayakkara. 2015. FingerReader: a wearable device to explore printed text on the go. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*. 2363–2372.
- Lee Stearns et al. 2014. The design and preliminary evaluation of a finger-mounted camera and feedback system to enable reading of printed text for the blind. In *Workshop on Assistive Computer Vision and Robotics (ACVR'14) in Conjunction with the European Conference on Computer Vision (ECCV'14)*. Zurich, Switzerland: Springer, 615–631.
- Jing Su, Alyssa Rosenzweig, Ashvin Goel, Eyal de Lara, and Khai N. Truong. 2010. Timbremap: enabling the visually-impaired to use maps on touch-enabled devices. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*. 17–26. DOI:<http://dx.doi.org/10.1145/1851600.1851606>
- J.L. Toennies, J. Burgner, T.J. Withrow, and R.J. Webster. 2011. Toward haptic/aural touchscreen display of graphical mathematics for the education of blind students. In *2011 IEEE World Haptics Conference*. IEEE, 373–378.
- Marynel Vázquez and Aaron Steinfeld. 2012. Helping visually impaired users properly aim a camera. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '12*. New York, New York, USA: ACM Press, 95. DOI:<http://dx.doi.org/10.1145/2384916.2384934>
- Kai Wang, B. Babenko, and S. Belongie. 2011. End-to-end scene text recognition. In *Computer Vision (ICCV), 2011 IEEE International Conference on*. 1457–1464. DOI:<http://dx.doi.org/10.1109/ICCV.2011.6126402>
- Xing-Dong Yang, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2012. Magic finger: always-available input through finger instrumentation. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*. UIST '12. New York, NY, USA: ACM, 147–156. DOI:<http://dx.doi.org/10.1145/2380116.2380137>
- Koji Yatani, Nikola Banovic, and Khai Truong. 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*. CHI '12. New York, NY, USA: ACM, 415–424. DOI:<http://dx.doi.org/10.1145/2207676.2207734>
- Koji Yatani and Khai Nhut Truong. 2009. SemFeel: a user interface with semantic tactile feedback for mobile touch-screen devices. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*. UIST '09. New York, NY, USA: ACM, 111–120. DOI:<http://dx.doi.org/10.1145/1622176.1622198>

## APPENDIX A

In this appendix we list the text of the documents used in Studies I and II, along with associated comprehension questions. We adapted six test documents from a Grade 8 Iowa Test of Basic Skills practice book [Princeton Review 2000]. The original text was modified slightly for length and to ensure clarity with our speech synthesis engine (*e.g.*, removing unnecessary proper nouns). We created three additional training documents of a similar length and reading level, as well as a two-column magazine document for testing KNFB Reader iOS, using documents adapted from articles in USA Today.

### Study I, Training Document (plain, both conditions):

*Scientists counting Antarctica's emperor penguins from space have found twice as many of them as expected. Using high-resolution satellite images to study each of 44 colonies around the coastline of Antarctica, experts said Friday that they put the total emperor penguin population at 595 thousand, or roughly double previous estimates of 270 thousand to 350 thousand. Seven of the colonies had never been seen before.*

*Satellite technology was a boon for researchers; visiting dozens of remote colonies in temperatures as low as minus 58 degrees is expensive, dangerous and time-consuming. With their distinctive black and white plumage, emperor penguins stand out against the snow, making them clearly visible on satellite images.*

### Study I, Test Document 1 (plain, first condition):

*People have used coins as a means of exchange for thousands of years. Valued for their craftsmanship and purchasing power, coins have been collected in great numbers throughout history and buried for safekeeping. Because stores of coins gathered and hidden in this manner lie untouched for many years, they can reveal a great deal about a given culture.*

*Coins are useful in revealing many aspects of a culture. They can provide clues about when a given civilization was wealthy and when it was experiencing a depression. Wealthy nations tend to produce a greater number of coins made from richer materials. The distribution of coins can also reflect the boundaries of an empire and the trade relationships within it. Roman imperial gold coins found in India, for example, indicate the Romans purchased goods from the East.*

*The way the coins themselves are decorated sometimes provides key information about a culture. Many coins are stamped with a wealth of useful historical evidence, including portraits of political leaders, important buildings and sculptures, mythological and religious figures, and useful dates. Some coins, such as many from ancient Greece, can be considered works of art themselves and reflect the artistic achievement of the civilization as a whole.*

*Information gathered from old coins by historians is most useful when placed alongside other historical documents, such as written accounts or data from archeological digs. Combined with these other pieces of information, coins can help historians reconstruct the details of lost civilizations.*

### Comprehension Questions:

1. Which of the following do coins reveal about a civilization?
  - a. The average cost of clothing
  - b. Information about its economy
  - c. Its farming techniques
2. What is the main idea of the passage?
  - a. How difficult it is to find old coins
  - b. How coins reflect the artistic achievements of a culture
  - c. How coins can tell us about ancient civilizations

**Study I, Test Document 2 (magazine, first condition):**

*Despite the stubborn, widespread opinion that animals don't feel emotions in the same way that humans do, many animals have been observed to demonstrate a capacity for joy. People have often seen animals evincing behavior that can only be taken to mean they are pleased with what life has brought them in that particular moment.*

*A chimpanzee named Nim was raised by a human family for the first year and a half of his life. After that time, Nim was separated from them for two and a half years. On the day that Nim was reunited with his human family, he smiled, shrieked, pounded the ground, and looked from one member of the family to the next. Still smiling and shrieking, Nim went around hugging each member of the family. He played with and groomed each member of the family for almost an hour before the family had to leave. People who were familiar with Nim's behavior said they had never seen him smile for such a long period of time.*

**Comprehension Questions:**

1. What is the purpose of the second paragraph?
  - a. To criticize Nim's human family for abandoning him
  - b. To show how well Nim's human family treated him
  - c. To demonstrate that animals have the ability to feel joy
2. Why did Nim shriek and pound the ground?
  - a. He was overjoyed to see the family again.
  - b. He was hungry and wanted to be fed.
  - c. He was angry with the family for leaving him.

**Study I, Test Document 3 (plain, second condition):**

*Born in Spanish Harlem in the late 1950s, Raphael Sanchez learned at an early age to listen to the many voices of the city. It was as a boy in Harlem that he developed the powers of observation that would later make his writing truly great. In the 1970s, Raphael went to Columbia University, where he was exposed to a literary tradition. While his university education gave his writing new depth, the raw energy of the streets has always served as the primary fuel for his writing. This is what gives his works passion and power.*

*Raphael once told me that in order to escape from life he turns to books, and in order to escape from books he turns to life. It is this balance of the sights, sounds, and smells of the street with the perspective gained from his formal education that has made Raphael popular with both critics and regular readers alike.*

*For those of us who have read and admired his work, it seems natural that Raphael has won so many awards. He deserves them, and his humility in accepting them has been refreshing. When he received the Writer's Quill Award two weeks ago, for example, he told the audience, "This award is not really mine. It belongs to all the million things that have inspired me.*

*"That is the kind of man I am introducing to you this evening. He is a man who has been inspired by a million things, and he is a man who has provided inspiration to a million people. Ladies and gentlemen, it is my great honor to present to you, Raphael Sanchez.*

**Comprehension Questions:**

1. Which of these best describes why Raphael Sanchez's writing is so popular with critics and regular readers?
  - a. It has won the Writer's Quill Award.
  - b. It reflects both scholarship and city experience.
  - c. It is based on his experiences at Columbia University.
2. What does Raphael Sanchez mean when he says, "This award is not really mine"?
  - a. He owes everything to the people and things that inspired him.
  - b. He does not believe in the value of awards.
  - c. He feels Columbia University should be given an award too.

**Study I, Test Document 4 (magazine, second condition):**

*In the 1800s, most geologists thought the sea floor was a lifeless expanse of mud, sediment, and the decaying remains of dead organisms. They thought that, with the exception of some volcanic islands, the bottom of the sea had no major geographic features, such as peaks or valleys.*

*In the mid-nineteenth century, ships depth-sounding the ocean floor with sonar for a transatlantic telegraph cable made some interesting discoveries. To geologists' surprise, the ocean floor was found to be made up of long mountain ranges and deep valleys and troughs. Another surprise finding in the Atlantic was the existence of basalt, a volcanic rock thought only to exist in the Pacific Ocean. The presence of basalt in the Atlantic was a clue that volcanic activity occurs at the bottom of the sea. This and other discoveries, many of them accidental in the beginning, were signals to geologists that their knowledge of the sea floor was very limited.*

**Comprehension Questions:**

1. The discovery of basalt in the Atlantic Ocean suggested that
  - a. Iron, zinc, and gold would be found nearby.
  - b. Geologists still had much to learn about the ocean floor.
  - c. The Atlantic was deeper than previously believed.
2. How did ships in the mid-nineteenth century measure the ocean's depth?
  - a. By sending down scuba divers
  - b. By bouncing sound waves off the sea bottom
  - c. By photographing the sea floor with special cameras

**Study II, Training Document 1 (plain, HandSight):**

*When Mary Smallenburg opened a package from her mother to find cereal and ramen noodles, she burst into tears. Without it, she wouldn't be able to feed her four children. It got to the point where I opened my pantry and there was nothing. Nothing. What was I going to feed my kids? Smallenburg says, adjusting a bag of fresh groceries on her arm.*

*Smallenburg's family is one of 50 military families that regularly visit the Lorton Community Action Center food bank. Volunteers wave a familiar hello as she walks in the door. None of what we have been through has been expected, Smallenburg says. Three of her four children have special needs and her husband is deployed in Korea. The last few months actually, coming here has been a godsend.*

*Nationwide, 25 percent of military families, 620,000 households, need help putting food on the table, according to a study by Feeding America, a network of 200 food banks. The results are alarming, says Bob Aiken, chief executive officer of Feeding America. It means that people in America have to make trade-offs. They have to pick between buying food for their children or paying for utilities, rent and medicine.*

*One in seven Americans, 46 million people, rely on food pantries and meal service programs to feed themselves and their families, the study found.*

**Study II, Test Document 1 (plain, HandSight):**

*Henry Ford and his Model T automobile changed the face of America. His horseless carriage contributed to a movement from rural to urban and to the development of an industrial economy.*

*In 1903, Ford Motor Company was officially formed, and in 1908, Ford announced the birth of the Model T. He told the nation, I will build a car for the great multitude. This was a bold announcement, since most manufacturers planned to build only luxury cars for the very wealthy.*

*His idea worked. Ford's Model T was a hit with the American public, and demand grew with each passing year. In the course of nineteen years, around fifteen million Model T cars were sold in the United States, nearly one million in Canada, and another 250,000 in Great Britain. All told, these numbers equaled half the total number of automobiles manufactured in the world at that time.*

*The methods of production Ford used were revolutionary. Ford's assembly line could churn out the frame of a Model T in about six hours. This high rate of speed was made possible by the division of labor. Instead of one person controlling production from start to finish, the labor was divided into smaller and smaller tasks, with each person performing the same task all day long.*

*By 1927 the era of the Model T was coming to a close. General Motors, a major competitor, was producing better cars for only slightly more money. Customers with an eye for new styles just didn't see the appeal of the plain Model T.*

**Comprehension Questions:**

1. According to the passage, why was the Model T more popular than other cars available at the same time?
  - a. It looked like a buggy.
  - b. It was more spacious.
  - c. It was less expensive.
  - d. It was more stylish.
2. Which of the following best describes Ford according to the article?
  - a. A poor businessman
  - b. A visionary
  - c. A follower
  - d. A great metal worker
3. What led to the downfall of the Model T?
  - a. It was not very well made.
  - b. Its price went up.
  - c. Other competition emerged.
  - d. Many of Ford's workers quit their jobs.

**Study II, Training Document 2 (plain, KNFB Reader iOS):**

*Here's a tip. Don't stress over tipping.*

*Restaurant tips are more modest in Europe than in America. In most places, 10 percent is a big tip. If your bucks talk at home, muzzle them on your travels. As a matter of principle, if not economy, the local price should prevail. Please believe me, tipping 15 percent or 20 percent in Europe is unnecessary, if not culturally ignorant.*

*Virtually anywhere in Europe, you can do as the Europeans do and, if you're pleased with the service, add a euro or two for each person in your party. In very touristy areas, some servers have noticed the American obsession with overtipping, and might hope for a Yankee-size tip. But the good news is that European servers and diners are far more laid-back about all this than we are. The stakes are low, and it's no big deal if you choose the wrong amount. And note that tipping is an issue only at restaurants that have waiters and waitresses. If you order your food at a counter, don't tip.*

*At table-service restaurants, the tipping etiquette and procedure vary slightly from country to country. But in general, European servers are well paid, and tips are considered a small bonus, to reward great service or for simplicity in rounding the total bill to a convenient number. In many countries, 5 percent to 10 percent is sufficient.*

**Study II, Test Document 2 (plain, KNFB Reader iOS):**

*A clone is a life form engineered in a lab environment to be identical to another, through a process of asexual, or nonsexual, reproduction. This process of creating a new life form, called genetic engineering, can be useful in creating individuals of a given species that represent the best possible genetic traits of that species. People who work with plants have long used cloning techniques to create better strains of trees, fruits, and vegetables. The Macintosh apple, for example, was created by cloning techniques, and it supposedly represents the best qualities of all apple types.*

*In July of 1996, a group of Scottish scientists made a breakthrough by successfully cloning a sheep from the cells of another adult sheep. After scraping cells from the udder of one sheep, the scientists introduced the nucleus of one of these cells into the unfertilized eggs of a different sheep. Then, they placed the egg, which had begun to divide, into the uterus of a third sheep. The result was Dolly, a healthy sheep who was born in the natural way from the third sheep. Dolly was almost identical to the sheep from whose skin cells she had been formed.*

*In 1997 Dolly's story was widely publicized in the media, and her existence resparked a continuing debate about the use of cloning techniques on humans. Some people claim that genetic engineering should not be used on humans under any circumstances. Others urge slowness. They insist that if genetic engineering is to be used, there are many questions that need to be answered first.*

**Comprehension Questions:**

1. What event revived the debate about cloning?
  - a. The cloning of plants
  - b. The cloning of a sheep
  - c. The cloning of amphibians
  - d. The future plan to clone human organs
2. According to the passage, how many sheep did it take to produce Dolly?
  - a. One
  - b. Two
  - c. Three
  - d. Four
3. In paragraph 1, the author mentions the Macintosh apple as an example of
  - a. a case when cloning produced an improved organism.
  - b. a case when cloning failed.
  - c. a case when cloning created a controversy.
  - d. a case when cloning went too far.

**Study II, Test Document 3 (magazine, KNFB Reader iOS):**

*Let them sleep! That's the message from the nation's largest pediatrician group, which, in a new policy statement, says delaying the start of high school and middle school classes to 8:30 a.m. or later is an effective countermeasure to chronic sleep loss and the epidemic of delayed, insufficient, and erratic sleep patterns among the nation's teens.*

*Multiple factors, including biological changes in sleep associated with puberty, lifestyle choices, and academic demands, negatively impact teens' ability to get enough sleep, and pushing back school start times is key to helping them achieve optimal levels of sleep, 8 and a half to 9 and a half hours a night, says the American Academy of Pediatrics statement, released Monday and published online in Pediatrics.*

*Just 1 in 5 adolescents get nine hours of sleep on school nights, and 45 percent sleep less than eight hours, according to a 2006 poll by the National Sleep Foundation (NSF).*

*As adolescents go up in grade, they're less likely with each passing year to get anything resembling sufficient sleep, says Judith Owens, director of sleep medicine at Children's National Medical Center in Washington, D.C., and lead author of the AAP statement. By the time they're high school seniors, the NSF data showed they were getting less than seven hours of sleep on average.*

*Chronic sleep loss in children and adolescents can, without hyperbole, really be called a public health crisis, Owens says.*

**APPENDIX B**

In this appendix we list the text of the subjective questionnaires administered in Studies I and II. For ease of use questions, the choices were (1) Very difficult, (2) Somewhat difficult, (3) Neutral, (4) Somewhat easy, or (5) Very easy. For comparison questions, the choices were (1) Much worse, (2) Somewhat worse, (3) About the same, (4) Somewhat better, or (5) Much better.

**Study I, after each directional guidance condition:**

1. *How easy or difficult was it to follow a line of text with your finger? Why?*
2. *How easy or difficult was it to read and understand the text given this feedback? Why?*

3. *Do you feel like the feedback direction was correct, or did it feel backwards to you?*
4. *Do you have any other comments about what was good or bad about this type of feedback?*

**Study I, end of study:**

1. *Overall, how easy or difficult was it to find the beginning of the text?*
2. *How easy or difficult was it to find the beginning of each line?*
3. *How easy or difficult was it to notice the end of a line?*
4. *How easy or difficult was it to notice the end of a paragraph?*
5. *How easy or difficult was it to find the beginning of the next column?*
6. *You have tried two different types of feedback. Which did you prefer more? Why?*
7. *Overall, how was your experience of our system compared to how you would normally read braille? Why?*
8. *Overall, how was your experience of our system compared to how you would normally use a screen reader? Why?*
9. *Overall, how was your experience of our system compared to how you would normally read printed documents? Why?*
10. *Do you have any questions, suggestions for improvement, or other comments?*

**Study II, after HandSight tasks:**

1. *How easy or difficult was it to find the beginning of the text?*
2. *How easy or difficult was it to find the beginning of each line?*
3. *How easy or difficult was it to notice the end of a line?*
4. *How easy or difficult was it to notice the end of a paragraph?*
5. *How easy or difficult was it to follow a line of text with your finger?*
6. *How easy or difficult was it to understand the feedback?*
7. *Overall, how easy or difficult was it to read and understand the text given this feedback?*
8. *Overall, how was your experience with the app compared to how you would normally read braille?*
9. *Overall, how was your experience with the app compared to how you would normally read printed documents?*
10. *Do you have any other comments about what was good or bad about this type of feedback?*

**Study II, after KNFB Reader iOS tasks:**

1. *Overall, how easy or difficult was it to read and understand the text using the app?*
2. *What, if anything, did you like about using the app?*
3. *What, if anything, did you dislike about using the app?*
4. *Overall, how was your experience with the app compared to how you would normally read braille documents?*
5. *Overall, how was your experience with the app compared to how you would normally read printed documents?*
6. *Overall, how was your experience with the app compared to reading with HandSight?*
7. *Do you have any questions, suggestions for improvement, or other comments about the KNFB Reader app?*