ABSTRACT

Title of Dissertation:	ACCESSIBLE ON-BODY INTERACTION FOR PEOPLE WITH VISUAL IMPAIRMENTS
	Uran Oh, Doctor of Philosophy, 2016
Dissertation directed by:	Assistant Professor Leah Findlater College of Information Studies

While mobile devices offer new opportunities to gain independence in everyday activities for people with disabilities, modern touchscreen-based interfaces can present accessibility challenges for low vision and blind users. Even with state-of-the-art screenreaders, it can be difficult or time-consuming to select specific items without visual feedback. The smooth surface of the touchscreen provides little tactile feedback compared to physical button-based phones. Furthermore, in a mobile context, hand-held devices present additional accessibility issues when both of the users' hands are not available for interaction (e.g., on hand may be holding a cane or a dog leash).

To improve mobile accessibility for people with visual impairments, I investigate on-body interaction, which employs the user's own skin surface as the input space. On-body interaction may offer an *alternative* or *complementary* means of mobile interaction for people with visual impairments by enabling non-visual interaction with

extra tactile and proprioceptive feedback compared to a touchscreen. In addition, on-body input may free users' hands and offer efficient interaction as it can eliminate the need to pull out or hold the device.

Despite this potential, little work has investigated the accessibility of on-body interaction for people with visual impairments. Thus, I begin by identifying needs and preferences of accessible on-body interaction. From there, I evaluate user performance in target acquisition and shape drawing tasks on the hand compared to on a touchscreen. Building on these studies, I focus on the design, implementation, and evaluation of an accessible on-body interaction system for visually impaired users.

The contributions of this dissertation are: (1) identification of perceived advantages and limitations of on-body input compared to a touchscreen phone, (2) empirical evidence of the performance benefits of on-body input over touchscreen input in terms of speed and accuracy, (3) implementation and evaluation of an on-body gesture recognizer using finger- and wrist-mounted sensors, and (4) design implications for accessible non-visual on-body interaction for people with visual impairments.

ACCESSIBLE ON-BODY INTERACTION FOR PEOPLE WITH VISUAL IMPAIRMENTS

by

Uran Oh

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Assistant Professor Leah Findlater, Chair
Professor Ben Bederson
Professor Gregg Vanderheiden
Associate Professor Niklas Elmqvist
Professor David Jacobs, Dean's Representative

Statement of Co-Authorship

All work in this dissertation was conducted under the supervision of Dr. Leah Findlater, and I am the primary contributor to all aspects of this research, with the exception of the work in Chapter 6. This work is collaborated work with Lee Stearns. Among the work that are presented in this chapter, he led the implementation of hardware prototype (Section 6.2), and image processing for localization (Sections 6.3.2.1 and 6.3.3), while I led data segmentation (Section 6.3.1), gesture classification (Section 6.3.4) and experimentation (Section 6.5). The rest, we worked on equally including extracting motion features where I focused on non-optical sensors while Lee focused on camera in Section 6.3.2.2.

Most of the research in this dissertation from Chapters 4 and 5 are updated versions of published papers or submitted manuscripts. Specific chapters and sections that directly derived from each publication are listed below:

- **Chapter 4**: Uran Oh and Leah Findlater. (2014). Design of and subjective response to on-body input for people with visual impairments. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility* (ASSETS '14). ACM, New York, NY, USA, 115-122.

- **Chapter 5**: Uran Oh and Leah Findlater. (2015). A Performance Comparison of On-Hand versus On-Phone Nonvisual Input by Blind and Sighted Users. *ACM Trans. Access. Computing.* 7, 4, Article 14 (November 2015), 20 pages.

Dedication

To my parents.

Acknowledgements

First of all, I would like to thank my advisor Dr. Leah Findlater for her endless support and encouragement. I've learned a lot from you, more than I could imagine. I truly thank you for guiding me towards the goal throughout this long path. I do not know how to thank you enough even now, but I am sure I will start thanking you more and more once I become a mom. I also thank Dr. Jon Froehlich for his sharp and practical feedback and enthusiasm which was always inspiring.

I'd also like to thank the members of my dissertation committee: Ben Bederson, Gregg Vanderheiden, Niklas Elmqvist, and David Jacobs. Thank you for providing valuable advice that made my work considerably stronger.

I would like to thank my lab mates and all the precious people who would always be by my side: Lee, Tak, Kotaro, Meethu, Kristin, Maia, Jonggi, Alisha, Paul, Kyoung-Eun, Ginnah, Angela, Eunhui, Heera, Narae, Minsun, Boram and Bok-Hee Ssam.

Lastly, I'd like to thank my parents for all the encourage and faith. I thank all my grandparents, Uncle Jeany, Aunt Sunny and my sister for their love and care. ☺

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1. Motivation

Mobile phones play an important role in our current society. These devices not only enable basic communication (e.g., phone calls, text messages), but also provide productivity- and entertainment-related functionalities. Furthermore, for people with disabilities, mobile devices may offer new opportunities to gain greater independence in everyday activities [1,72]. However, while screen reader software (e.g., Apple VoiceOver[1], Google Talkback[2]) has contributed to the wide adoption of touchscreen devices for users with visual impairments (VI users) [137], basic tasks such as locating specific items on modern touchscreen-based smartphones can still be inaccessible or time-consuming with limited tactile feedback from the input controls [19,44,53,75].

Compared to using the touchscreen, on-body interaction, which employs the user's own body as an always-available input surface (e.g., [27,39,43,69]), may provide an *alternative* or *complementary* means of mobile interaction for people with visual impairments. On-body interaction could be particularly compelling for VI users who use screenreaders and thus do not need pull out the device to see the visual output of the screen—an action that in itself takes 4.5 seconds on average [11]. Moreover, because the input is performed on the user's own body, there is no need to hold an additional device, which may be particularly beneficial in a mobile context if the user is holding a cane or dog leash. Moreover, research with sighted users has shown that

1 http://www.apple.com/accessibility/ios/voiceover

2 https://support.google.com/accessibility/android/answer/6283677?hl=en&ref_topic=3529932

the tactile feedback from one's own body can offer more efficient non-visual interaction for sighted participants compared to the smooth surface of a touchscreen [38]. This on-body tactile benefit is likely to be at least as useful for blind users, who have greater tactile acuity than sighted users [20,32,34,63,116].

While on-body interaction is potentially beneficial for VI users, little work has explored this prospect or how to design and implement such interaction to be accessible specifically for VI users. Instead, almost all studies of on-body interaction have focused on supporting sighted users, either with visual output (e.g., projecting visual interface elements such as buttons on the skin) [40,42,43,76,117] or for non-visual use [27,39,69]. While Gustafson et al. [38] identified the possible performance benefit of on-body interaction for VI users, this potential remains largely unexplored, as they collected data from only one blind participant.

1.2. Dissertation Objectives and the Thesis Statement

To support accessible mobile computing for VI users, I have investigated on-body interaction. The objectives of the dissertation include: (1) identification of perceived advantages and limitations of on-body input compared to a touchscreen device, (2) assessment of performance benefits of on-body input over touchscreen input in terms of accuracy and speed, (3) implementation and evaluation of an on-body gesture recognizer using finger- and wrist-mounted sensors, and (4) design implications for accessible non-visual on-body interaction for people with visual impairments. My work toward these objectives shall demonstrate the overarching thesis of this dissertation:

On-body interaction can provide an alternative or complementary means of accessible mobile computing for visually impaired users, with improved speed and accuracy as compared to touchscreen interaction.

1.3. Approach and Overview

Towards the objectives of the dissertation outlined above, we first explored on-body gestures created by 13 sighted and 11 VI participants to understand the characteristics of on-body gestures preferred by VI users and whether those characteristics differed from sighted users (Chapter 3). We asked participants in both the sighted and VI groups to create a set of gestures, and we categorized those gestures based on attributes such as the location at which the gesture was performed or whether the gesture was static or included motion. The findings, though preliminary, suggested that VI users might have a greater tendency than sighted users to create *location-specific* gestures (e.g., pointing to a specific finger for different tasks) and static gestures, which motivated the need to design on-body interaction specifically for VI users.

Building on this preliminary result, we conducted a more in-depth user study with 12 low vision and blind participants [84] to investigate the design of and subjective response to non-visual on-body interaction (Chapter 4). Here, the focus was on understanding participants' preferences for different on-body input locations (e.g., palm, forearm, neck and face) and to compare on-body input to touchscreen phone input both with one-handed and two-handed interaction. We asked participants to create gestures for different locations on their body, and then to complete basic mobile tasks (e.g., item navigation/selection) with either one or two hands on a touchscreen phone or on the participant's hand. These basic mobile tasks employed *location-*

independent input, that is, taps or swipes that could be theoretically performed anywhere, rather than location-specific input. The results revealed that the face/neck area was the least preferred location for on-body gestures mainly because touching at that location may be seen as socially inacceptable, while locations on the hands were considered to be the most discreet and natural. Further, for the location-independent input tested here, on-body input was considered to be especially useful for contexts where one hand is busy (e.g., holding a cane or dog leash).

After identifying formative user needs and preferences for on-body interaction, we designed a controlled lab study [85] to examine the impact of on-body interaction on VI users' input performances (Chapter 5). Eleven blind and 12 sighted participants completed non-visual target pointing and gesture drawing tasks (e.g., circle, triangle) on both a touchscreen phone and on their own palm. The results showed that users were able to point to targets on their hand faster and more accurately than on the touchscreen, and that shapes drawn on the hand had higher recognition rates than those drawn on the phone (with the implication being that the shapes were more consistent on the hand). The findings confirmed the performance benefit of on-body input over touchscreen input for VI participants, extending previous pointing input results with only sighted participants [38] both to this new user group and to the more complex shape-based gestures.

Following the design and performance studies, we investigated a novel approach to recognizing on-body gestures using wearable sensors (Chapter 6). Camera-based sensing techniques have been widely explored as images can provide rich contextual information (e.g., [39,40,76]). However, for people with visual impairments, the use of a camera often leads to issues such as out-of-frame (e.g., gesturing finger is beyond the camera) or occlusion (e.g., gesturing finger is not visible/hidden by other objects) (e.g., [3,129]). Thus, we employed a finger-mounted camera, along with other sensors (e.g., accelerometer, gyroscope) to collocate touch, sensing and feedback (*i.e.*, tactile feedback from their own skin). Furthermore, this approach can extend the input space to anywhere the finger can reach and increase the input vocabulary as it can classify not only the gesture, but its performed location, which we call location-specific gestures (e.g., tapping on wrist versus tapping on thigh). As a collaborative work, my main contribution was to implement a location-specific on-body gesture recognizer with the non-optical sensors, while my colleague focused on the camera-based sensing. The recognizer was trained with both temporal and descriptive features from inertial motion unit (IMU) and infrared (IR) sensor values using a frame-based support vector machine (SVM). To evaluate our approach, we collected on-body gesture examples from 24 people using the full set of sensors, and conducted an offline performance evaluation on the recognition accuracy with different sensor combinations. The results demonstrated that our finger-mounted input sensing system is feasible for supporting 24 location-specific gestures with the average accuracies of 86.2% with non-optical sensors, and 86.3% with the camera alone, which increased to 94.9%, when all sensors were used. The algorithms developed here also form the basis of the real-time sensing system (not a contribution of this dissertation) used for the next study.

As for the final work, we focused on microinteractions. Microinteractions are single-purpose interactions that can be completed within only a few seconds with minimal effort [9], such as dialing a number or adjusting volume. Supporting efficient interaction is desirable; however, even basic interactions such as typing may take a longer time for people with visual impairments particularly for touchscreen devices (e.g., [15,19]). As such, the inefficiency can lead many to forgo security features (e.g., passcode) [14], or may contribute to concerns about safety [2,135]. Thus, we conducted an online survey with 134 respondents (78 who had visual impairments) and a semi-structured interview to understand the barriers for microinteractions, and how wearable devices are perceived, and whether they may be able to reduce the overall interaction time by eliminating the need for device retrieval [11] for visually impaired users compared to sighted users. Moreover, we conducted a design probe study with 12 blind participants using a real-time on-body interaction system (Chapter 7) to evaluate different types of interaction designs for supporting microinteractions. The results revealed nine microinteractions (e.g., activating voice input, responding to a text message) that would be most valuable to be supported for users with visual impairments compared to sighted people. The findings also showed that the perceived benefits of smartwatches, specifically, for microinteraction may not outweigh the limitation of the small, watch-sized touchscreen. Furthermore, although having to learn the location for each of the specific microinteractions was considered to be a drawback, we found that all participants liked the idea of location-specific on-body gestures for its efficiency.

1.4. Organization of the Dissertation

The rest of this dissertation is organized as eight chapters. In Chapter 2, we discuss a literature review related to my thesis. Chapter 3 explores usable on-body gestures for people who have visual impairments and how these gestures differ from those of sighted participants. Chapters 4 and 5 present controlled lab studies of non-visual, on-body interaction for people with visual impairments in terms of preference and performance. Chapter 6 addresses the implementation of an on-body input sensing system using finger-mounted sensors. Next, Chapter 7 focuses on the accessibility of microinteractions on mainstream mobile and wearable devices for people with visual impairments, and compares three different real-time on-body interfaces to support accessible microinteractions. Finally, Chapter 8 describes possible future research projects that can extend the current scope of this dissertation.

Chapter 2: Related Work

In this chapter, I provide an overview of mobile and wearable technologies in terms of accessibility for people with visual impairments, as well as detailed discussion of the subjective and technical challenges that are most relevant to my dissertation contributions. This chapter begins with a general background on mobile accessibility for users with visual impairments (Section 2.1). Section 2.2 describes various studies related to on-body interaction both in terms of sensing on-body input and design implications. Section 2.3 reviews projects on gestural interfaces using wearable sensors, which may offer many of the desirable features for VI users such as hands-free and quick access to physical devices. Finally, Section 2.4 summarizes this chapter.

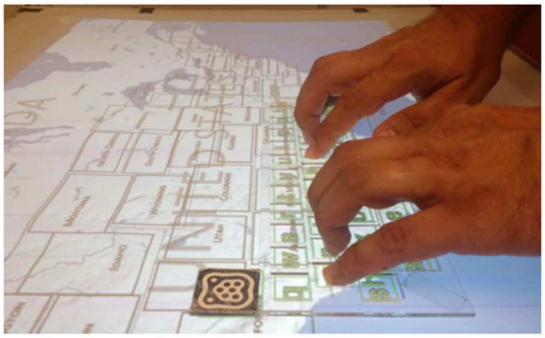
2.1. Accessible Mobile Computing for People with Visual Impairments

While mobile phones had traditionally been used as a communication device, they have evolved to serve users with a wide range of functionalities (e.g., music player, camera, route navigation), and are thus often compared to a "Swiss Army Knife" [35,107]. Especially for people with visual impairments, mobile technology can play an important role for increasing independence and safety [51,102]. However, modern mobile devices with touchscreens may have accessibility issues [66,75,102,135], weakening the advantages of possessing mobile devices due to a high dependency on visual cues and the smooth surface that may introduce additional challenges compared to physical button-based phones.

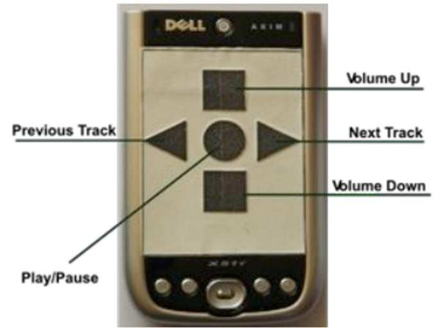
To improve touchscreen accessibility, most widely adopted techniques synthesize speech output upon a user's touch. The *Talking Fingertip Technique* [120], for instance, provided auditory description of touched items on a kiosk touchscreen for users who need nonvisual access while scanning and finding specific items with their fingers. Adding on to speech feedback, gestural interfaces were proposed for blind users to allow location-independent input, allowing the user to perform a gesture (e.g., swipe, tap) anywhere on the screen instead of having to touch their finger to a specific region [36,50]. *Slide Rule* [50], for example, supported multi-touch gestures that can be performed anywhere on the screen, and found that these gestures were more efficient than tapping spatially designated regions (e.g., top right corner), similar also to the approaches demonstrated by Sánchez and Aguayo [105] and Mobile Speak Pocket[3]. *NavTouch* [36] adopted directional gestures with a single finger for blind users to navigate the target character on a touchscreen for text entry. Commercial products have also employed location-independent gestures accompanied with screenreading software, such as iOS's *VoiceOver*[4] and Android's *Talkback*[5]. For instance, a *left-to-right* flick gesture will move a cursor to the next item, regardless of the spatial location of the input. While these gesture-based interfaces are widely disseminated, touchscreen gestures may not be usable [53], or are at the very least challenging to learn for visually impaired users [86].

As opposed to gestural interfaces, researchers have also investigated tactile enhancements for the spatial layout of the graphical user interface of the touchscreen touchscreen (e.g., [52,60,75]). *Touchplates* [52], for example, allows blind users to

[3] http://www.humanware.com
[4] https://www.apple.com/accessibility/ios/voiceover/
[5] https://support.google.com/accessibility/android/#topic=3529932



A physical overlay with a QWERTY keyboard layout on a large touchscreen.



Raised paper control panel overlaid on touchscreen-based MP3 player.

Figure 2.1. Examples of physical overlays to increase tactile feedback of touchscreen devices.

interact with large touchscreens by offering tactile feedback with a physical guide overlaid on the smooth screen for various layouts such as keyboard and maps (see Figure 2.1a). With physical overlays that provide tactile guidance for users to easily distinguish one control to another, spatial-layout based interfaces may enable more efficient interaction than location-independent gestural interfaces on a touchscreen device, while location-independent gestures were found to be faster than location-specific gestures on a touchscreen when no extra tactile feedback was given in Kane et al. [50]. For instance, McGookin et al. [75] compared both gesture-based and overlaid location-based interaction (as in Figure 2.1b) for a touchscreen MP3 player, and showed that location-independent gestures were less accurate and slower to complete basic digital music player operations (e.g., adjusting volume) compared to location-specific input on a button-shaped overlaid control panel. Of course, faster interaction may be possible for spatially stable items (e.g., keypads on QWERTY keyboard, marking menu items [57]) once users develop spatial memory of the interface layouts, as less time will be required for searching.

In this dissertation, I demonstrate the potential of on-body interaction for supporting non-visual access to mobile computing, compared to a touchscreen device in terms of subjective responses (Chapters 4 and 5), and task performance (Chapter 5). Moreover, we conducted an online survey and a semi-structured interview (Chapter 7) to understand how on-body interaction can be better designed by reflecting on participants' responses on their touchscreen devices including both smartphones or smartwatches. Then we designed, implemented and evaluated three different on-body interfaces focusing on the impact of both gestural and spatial interactions with a real-time system to understand design implications specific for VI users.

2.2. On-Body Interaction: Appropriating Skin as an Input Surface

On-body interaction has the same advantages as most wearable devices; it enables hands-free interaction and discreet use, and it can also allow quick information access by eliminating the device retrieval time. Moreover, with extra tactile and proprioceptive feedback from users' own body, it can allow accurate and fast performance under eyes-free conditions [27,38,39,69]. In this regard, various studies have been conducted to understand design implications focusing on the range of on-body input vocabulary, and to investigate input sensing techniques. However, except for one study by Gustafson et al. [38], the work discussed in this section has focused on sighted users rather than visually impaired users, who may have different preferences and needs.

In terms of studying input vocabulary for on-body interaction, Weigel et al. [128] investigated characteristics of different skin input modalities and preferred on-body input locations for different applications. They found that users prefer to use touchscreen gestures over other modalities such as scratching or squeezing the skin. They also showed that users may prefer different locations for different input types (e.g., handwriting on palm, keyboard on forearm) while their hand is the most preferred location in general. Other studies focused only on touch input, investigating to what extent users are capable of consistently pointing to a specific region at various locations on the body [27,39,40,69,119,126,136], where each region can served as a distinct input. Lin et al. [69], for example, examined the forearm as an input surface, and showed that sighted participants without visual cues were able to segment the area into six regions for pointing without overlapping. *PalmRC* [27] investigated nine regions on the palm and fingers, and found that participants were able to point to five landmarks on the palm with the average accuracy of 94-98% in a eyes-free manner. Furthermore, Gustafson et al. [38] revealed that the extra tactile feedback from users' palms enables faster target pointing performance for blind-folded sighted participants compared to a smooth surface such as a touchscreen. While a number of researchers have investigated the eyes-free on-body interaction for sighted users, the findings from these studies may not directly apply to VI users who have different needs and preference (e.g., as is already known for touchscreen gestures [53]). The benefits of tactile feedback from the skin may be even greater for blind users for whom tactile acuity has been found to be higher than it is for sighted users [20,32,34,63,116]. However, little work has studied the performance advantages of appropriating skin as an input surface for people with visual impairments, except for Gustafson et al. [38], which included one blind participant.

In terms of sensing on-body input, a wide range of approaches have been explored, including cameras ([23,27,39,40,126]), IR [61,80–82], ultrasonic rangefinders [68,69], bio-acoustics [43], magnetic fields [24], electromyography (EMG)

[74], and capacitive sensors [68,74,106,127]. While these approaches are promising and have inspired our own work, each is limited in some way by its sensor types and placement. First, the interaction space is often constrained by the sensing range. Although widely used, cameras mounted on the upper body [27,39,40,117], for example, restrict the interaction space to a pre-defined region within the camera's field of view, which would be also problematic for people with visual impairments as aiming the camera can be challenging without visual feedback (e.g., [3,129]). Moreover, optical [23,126] and non-optical [43,61,68,81,127] sensors mounted on one arm or hand to detect gestures performed by the other similarly limit the on-body interaction space to a relatively small area, and cannot easily be scaled to other body locations without requiring additional sensors. Furthermore, prior work focuses either on input localization or detecting two-dimensional (2D) input position within a specific region and, therefore, cannot support location-specific gestures, which requires both. For example, *Skinput* [43] and *PUB* [69] can localize touch input on various locations on body using either bio-acoustics or ultrasonic range-finding. However, these systems cannot recognize gestures in 2D space such as directional swipes. In contrast, systems such as *FingerPad* [24], *PalmGesture* [126], and *SenSkin* [81] can estimate x and y coordinates of on-body input, enabling more complex gestures like shapes. However, these methods require sensors affixed on or near the interaction surface in order to achieve such precision, and they therefore cannot easily be extended to multiple locations.

We investigated characteristics of on-body input preferred by VI users in general (Chapters 3 and 4) based on subjective responses, and studied performance-based design implications (Chapter 5). To reliably support location-specific gestures and to implement accessible on-body interaction for people with visual impairments, we explored a finger-, and wrist-mounted sensors rather than sensing using hardware mounted on the target input surface or on some central body location (Chapter 6). Finally, for Chapter 7, we directly compared a gesture-based interaction with spatial interactions with two different levels of resolutions–dense locations on palm, and locations spread across different body parts.

2.3. Finger-mounted Devices for Gestural Input

A number of wearable devices have been studied for supporting gestural input, varying the mounting locations and sensor types (e.g., [23,41,55,70,76,112,117]). Some researchers studied body-mounted cameras for recognizing whole hand gestures (e.g., [23,70,112]), or tracking the fingertip [39,40,76]. *Gesture Pendant* [112], for example, uses a wearable camera around the chest to control a home automation system such as adjusting the room lighting or volume with hand gestures. *PinchWatch* [70] also deployed a camera on the chest (or ear or belt), supporting one-handed interaction to allow users to quickly switch their visual attention from their primary task for a brief moment, using different thumb-based pinching gestures.

Others explored wrist- or finger-mounted sensors [10,25,41,49,134]. *Abracadabra* [41], for example, uses magnetometers to support wireless, unpowered input for tracking the two-dimensional spatial location of the hovering finger to interact with small mobile devices such as smart watches. Similarly, *Nenya* [10] employed a magnetic ring worn on a finger, together with a 3-axis magnetometer for eyes-free discreet interaction. An accelerometer is also examined in *Magic Ring* [49] which can recognize six different gestures including rotating direction, and up/down direction and distance of each finger movement for controlling appliances.

Often, finger-mounted cameras were also examined for supporting efficient interactions with a user's surroundings (e.g., [23,78,98,134]). For example, *EyeRing* [78] used a finger-mounted camera to leverage the pointing gesture to interact with physical objects in the environment such as money identification. *Magic Finger* [134] also uses a micro RGB camera and an optical flow sensor instrumented on the user's finger. It can detect not only the x and y movements and contact but it can also recognize different materials by discriminating textures from the scanned images on contact for supporting tasks such as checking appointments or muting a call. While a finger-mounted camera can be used for identifying different parts of the user's body, little has been explored except for two input locations on the body (e.g., thumb and hand skin) in *Magic Finger* [134].

While wearable devices can be perceived positively by VI users [135], and observers [96], most work on wearable cameras for people with visual impairments has focused on aiding visual tasks such as way-finding (see surveys [26,103,124]) or object/character recognition for reading assistance ([110,113]). For instance, *OrCam* [89] has a small camera attached to a pair of glasses, and it allows users to activate speech by performing a pointing gesture on texts or an object.

In this dissertation, I implemented and evaluated finger-mounted sensors for recognizing different on-body gestures (Chapter 6), and investigated how current mobile devices provide accessible interaction for people with visual impairments, and how wearable devices, including on on-body interaction, can be designed to support non-visual interaction for mobile computing (Chapter 7). Furthermore, when investigating the accessibility of wearable devices for VI users, I focused on "microinteractions [10]", the interactions that are designed to be completed within a short period of time with minimal effort (e.g., dialing a number or adjusting the volume), which wearable devices are often designed for (e.g., [11,70,78,134]).

2.4. Summary

While on-body interaction is potentially beneficial for VI users as a means of accessible mobile computing complementing touchscreen-based interaction, little work has investigated the prospect. Toward this dissertation, I investigated design implications for supporting non-visual on-body interaction for VI users based on the assessment of the subjective preferences (Chapters 3 and 4), and performance (Chapter 5). Then I designed and implemented an accessible on-body interaction system for VI users with a finger-, and wrist-mounted sensors as a replacement of body-mounted or hand-held devices (Chapter 6). Finally, I investigated how to employ this finger-mounted on-body sensing system to support blind users in completing microinteractions—that is, brief interactions that are typically seen as a strength of wearable computing devices (e.g., [11,70,112]) in Chapter 7.

Chapter 3: Usable On-Body Input for People With Visual Impairments

3.1. Motivation and Introduction

For early prototyping of new gesture-based interfaces, understanding the unbiased interaction behavior of users without having to be concerned about reliable gesture recognition or technical limitations can be useful. As such, Wobbrock et al. [131] proposed a user-defined gesture protocol to elicit users' natural input—specifically, to understand what gestures users would *expect* to be able to use for tabletop interaction—and the research team asked non-technical users to create gestures using their own intuitive gestures. This user-defined gesture protocol has since been applied in variety of contexts such as three-dimensional motion gestures [104], flexible display [59], and multi-touch marking menus [65]. Moreover, this protocol may be used to capture the differences in gesture preference or usability between different user groups. For instance, Kane et al. [104] explored user-defined gestures created on a touchscreen by both sighted and blind users, and found that each user group has different gesture preferences (e.g., blind people created significantly more gestures using the edge of the screen than did sighted people).

As a preliminary study for understanding the challenges and opportunities of applying on-body interaction for people with visual impairments, we conducted a single-session lab study with 24 participants (13 sighted, 11 blind and low vision) where they were asked to create gestures for ten mobile actions, following a user-defined gesture protocol. The goal was to explore and compare preference for different types of on-hand gestures between the two groups of participants. The two main research questions were:

- *What gesture characteristics do visually impaired users prefer for on-body interaction?*
- *How do on-body interaction preferences of visually impaired users differ from those of sighted users?*

3.2. Experimental Methodology

To understand the accessibility challenges and opportunities of on-body interaction, focusing on the differences in gesture preference and characteristics between users with or without visual impairments, we conducted a single session study both with sighted and visually impaired users. The study was designed to capture emergent preferences for on-hand gestures by asking participants to create their own gestures for ten mobile actions—similar to a user-defined gesture protocol [39] but with a more exploratory goal.

3.2.1. Participants

We recruited 24 participants: 13 who reported having no visual impairments (**sighted**; 6 female), and 11 with visual impairments (**VI**; 6 female) where eight VI participants were blind since birth, while the remaining three had low vision. Sighted participants were on average 28.5 years old (*SD* = 5.9) versus 39.9 (*SD* = 11.8) for VI participants. In terms of smartphone experience, all sighted participants own touchscreen phones, while three VI participants were not touchscreen phone owners. All participants reported using their phone at least once every few hours, with the exception of one VI

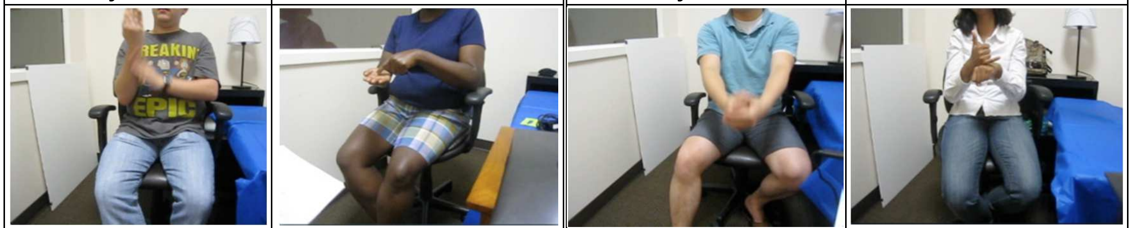
Category	Sub-Category	Action Name
Action	System/Phone	Answer A Phone Call
		Dial A Phone Number
		Check Current Time
		Volume Up
		Volume Down
	Application	Open Selected Item
Navigation	System/Phone	Return To Home Screen
		Open An E-mail App
	Application	Move To Next Item
		Move To Previous Item

Table 3.1. Mobile actions for user-define gesture task, grouped by category.

participant who used it once a day. We recruited via campus e-mail lists and local organizations that serve people with visual impairments. Participants were compensated for their time.

3.2.2. Procedure

Participants were instructed to imagine a device worn on their chest that can sense gestures when their two hands are in contact. They were asked to do so for ten mobile actions (e.g., answering a phone call, returning to home screen) that were selected to cover a range of mobile actions from Ruiz et al. [53]. For each action, the name and a brief description were read aloud by the experimenter. Then, each participant was asked to create a gesture for the described action. During the task we did not provide feedback on the quality of the gestures or on whether the hypothetical system would be able to recognize them (following [131]). For each action shown in Table 3.1, the name and a brief description was read aloud by the experimenter in a random order. Then, we asked participants to evaluate each gesture. Think-aloud protocol was used throughout the task. Since our goal was exploratory, we selected the mobile actions in Table 3.1 not



On-Hand		Mid-Air	
Dynamic	Static	Dynamic	Static
(a) One hand sliding down the side of the other hand.	(b) Pressing center of the palm to press 5 as if there is a number pad.	(c) Hands moving downward as one hand grabs the other in a fist.	(d) One hand in a phone shape, placed on top of the other fist.

Figure 3.1. Examples of dynamic and static gestures that are performed on hand versus mid-air. Corresponding actions are (a, c) volume down, (b) dial a phone number, and (d) place a phone call.

necessarily to be comprehensive, but to cover a range of motion gesture categories as identified Ruiz et al. [104]. A brief discussion period concluded the session.

3.2.3. Data and Analysis

All sessions were video recorded and later analyzed to identify characteristics of each gesture; participants' comments and quotes were also transcribed. We collected 240 gestures (10 gestures × 24 participants), and two subjective 7-point Likert ratings for each gesture. Each gesture was categorized based on its characteristics for the analysis. We applied non-parametric test such as Mann-Whitney U test.

3.3. Results

For self-evaluation of gestures, we used metrics from Wobbrock et al. [131]. For identifying types of gestures created by participants, we adapted existing gesture taxonomies for surface and three-dimensional (3D) gestures [5,131] to accommodate on-body gestures, and categorized into two dimensions: relationship between the hands (*on-hand vs. mid-air* gestures), and movement (*static vs. dynamic*), as shown in Figure 3.1.

Action	Easiness		Good Match	
	Mean	SD	Mean	SD
Answer A Call	6.63	0.77	6.00	1.18
Dial A Phone Number	5.42	1.41	5.67	1.55
Check Current Time	6.58	1.06	5.96	1.46
Volume Up	6.17	1.40	5.83	1.46
Volume Down	6.42	1.21	5.83	1.49
Open Selected Item	6.54	0.88	5.79	1.28
Return To Home Screen	6.42	1.14	5.71	1.49
Open an E-mail App	6.08	1.06	5.50	1.22
Move to Next Item	6.5	0.93	5.58	1.38
Move to Previous Item	6.25	1.03	5.54	1.18

Table 3.2. The mean and standard deviations (*SD*) for ratings on Easiness and Good Match in 7-point Likert scales (1 = *strongly disagree*, 7 = *strongly agree*) for the gesture created for each of the 10 actions (*N* = 24).

3.3.1. Quality of Self-Defined Gestures

After the creation of each gesture, participants were asked to provide ratings for two 7-point Likert-scale questions: (1) *Easiness*: "The gesture I picked is easy to perform.", and (2) *Good Match*: "The gesture I picked is a good match for its intended purpose." Both user groups gave high ratings for both metrics with the average ratings of 6.3 (*SD* = 1.1), and 5.74 (*SD* = 1.4) for *Easiness* and *Good Match*, respectively. Since our main focus is to explore differences between two participant groups, no in-depth comparisons were conducted for the ten actions; descriptive statistics are shown in Table 3.2.

3.3.2. On-Hand Versus Mid-Air Gestures

One-handed gestures that uses the other hand as a reference (e.g., swiping a finger on the opposite palm as on a touchscreen), namely *on-hand* gestures, was more popular than *mid-air* gestures, which involve both hands equally (e.g., two stacked fists) for both user groups (see Figure 3.1 for examples). Roughly two thirds of the gestures were *on-hand* gestures for both user groups: 72.3% of total gestures on average (*SD* = 34.9%) for VI participants, 66.2% (*SD* = 30.0%) for sighted participants. The difference was not statistically significant by a Mann Whitney U test.

3.3.3. Static Versus Dynamic Gestures

Gestures were also categorized into *dynamic* or *static* depending on whether the gestures involved movement or not. While roughly half of the gestures created by VI participants were static, with on average 46.4% (*SD* = 20.1), only 28.5% (*SD* = 18.6) of gestures were static on average for sighted participants. A Mann Whitney U test showed that the tendency to create static gestures was significantly greater for VI participants than sighted participants ($U_{(22)} = 29.00$, $Z = -2.51$, $p = .013$).

3.3.4. Location-Specific Versus Non-Specific Gestures

We further categorized gestures into *location-specific* (e.g., pointing specific fingers) and *location-independent* (e.g., a left-to-right swipe anywhere on the palm). While more than half (M =52.7%, SD=22.8) of the gestures created by VI participants were location-specific, only the average portion was 16.2% (SD=18.5) for sighted participants. Mann Whitney U showed that the difference was statistically significant ($U_{(22)} = 131.0$, $Z = 3.49$, $p < 0.001$). Note that all location-specific gestures were performed on hand, where the majority of these gestures were static (74.4%).

3.3.5. Subjective Feedback for On-Hand Interaction

Both participant groups reported high interest in on-hand interaction in general. Almost all of the sighted participants (11 out of 13) were positive, although they would prefer having visual feedback like on their phone. Two participants with low vision (with 20/200, 20/200) felt similarly. The remaining participants (including two participants with 20/800 and 20/3000 vision) from the VI group, however, highly valued the potential of on-hand interaction. For example, P24 commented that, *"If I have a choice of both [on-hand interaction and touchscreen phone interaction], I would choose this [on-hand]. It's more natural for me because I see with my hand"*.

3.4. Discussion

Here we discuss our findings, although preliminary, and address the limitations of the study.

3.4.1. Using Hand as a Touchscreen of the Phone

For both sighed and VI user groups, when they were asked to create gestures while two hands in contact, on-hand gestures were dominant where they perform a gesture on the other hand (especially on the palm) as if on a touchscreen of their phone. The findings confirmed the recommendation from Kane et al. [53] and Rico et al.'s [101] to reproducing layouts (either virtual or physical) or interfaces similar to existing interfaces they use, which were also demonstrated in other studies on on-body interaction (e.g., [27,39,40]). Moreover, as shown in Gustafson et al. [10], enabling the user's hand to behave like a smartphone can relieve the learning curve as users can transfer their mental model of the standard phone interface (e.g., app layouts) to their hands. Thus, supporting the palm as one of the input locations for allowing touchscreen-like gestures that users are already familiar with is recommended when designing on-body interaction.

3.4.2. Supporting Various Input Locations for VI Users

Our findings indicate that VI users may favor gestures performed at specific locations and landmarks (e.g., tapping a specific finger or knuckles) on their hand more than sighted users do, although further investigation is needed as there was a restriction on the gestures that both hands be touching as was required in this study. Besides the preference, gestures at a specific location may provide performance benefits over gestures that are insensitive to input locations (e.g., a directional swipe anywhere on the palm) with extra tactile feedback as shown in McGookin et al. [75]. Furthermore, beyond the tactile cues for users to locate users' fingers to specific regions, supporting locations with distinctive physical landmarks on the body may be desirable for VI users, similar to how edge-based gestures were preferred in Kane et al. [53].

3.5. Limitations

While the goal is to directly compare the differences in the preferences between sighted and VI user groups, the participants we recruited may not be representative due to random recruitment with a small sample size. Moreover, the two user groups were not matched in terms of their age; the participants in blind user group are about 10 years older on average than those in sighted user group. This age gap may have contributed to differences in terms of the types of gestures each user group created since older adults' usage of and familiarity with technologies differ from that of younger people [58,109]. Furthermore, we required that participants touch their hands together while creating gestures, but without this constraint, the gestures would likely have been different—particularly, for example, one may expect to see more mid-air gestures. Furthermore, although our focus was on understanding the design implications of on-body interaction for supporting mobile computing, the characteristics of gestures we collected would be different if the participants were asked to create gestures for actions in different applications (e.g., manipulating objects in virtual or augmented reality, documenting on a desktop).

3.6 Conclusion

In this work, we conducted an exploratory study to understand usable on-body gestures for people with visual impairments. The findings from this preliminary study suggest that VI participants may have different preferences for on-body gestures than sighted participants as well as lead to the following design implications for VI users, namely, (1) supporting on-hand gestures similar to gestures performed on touchscreen mobile phones, and (2) supporting various locations as input surfaces for mapping different tasks.

Findings from this study suggested the need for designing on-body interactions specifically for VI users, whose preferences may be different from those of sighted participants. This motivates our next study, which focuses on VI users only, in terms of subjective feedback on different on-body input locations (e.g., forearm, neck and face) and their perception of on-body interaction as opposed to touchscreen.

Chapter 4: Design of and Subjective Responses to On-Body Input for People With Visual Impairments

4.1. Introduction and Motivation

The study in the previous chapter showed that on-hand gestures were preferred over mid-air gestures, and that there could be potential needs for designing on-body interaction particularly for people with visual impairments (e.g., location-specific gestures). However, the study is limited as we did not explore on-body input locations beyond hand, and on-body interaction was not compared with touchscreen-based interaction. Thus, we conducted a more in-depth study.

While most of the other researchers have explored the focused area (e.g., hand or forearm) as an input space (e.g., [27,39,40]), here, we explored on-body input preferences across five different locations on the user's body (e.g., palm, neck and face). In addition, we collected subjective responses comparing touchscreen and on-hand interaction given different hand availability (one versus two hands), which simulates mobile contexts (e.g., one hand holding a cane). Moreover, as most wearable devices are designed for both private and public use, we examined whether the interaction is perceived to be acceptable by others in social contexts. The primary goal of the study was to gain more comprehensive insights into the design of accessible on-body input for people with visual impairments, reflecting users' needs, concerns and preferences. We had two main research questions:

- *How visually impaired users respond to on-body interaction?*
- *How should on-body interaction be designed for visually impaired users?*

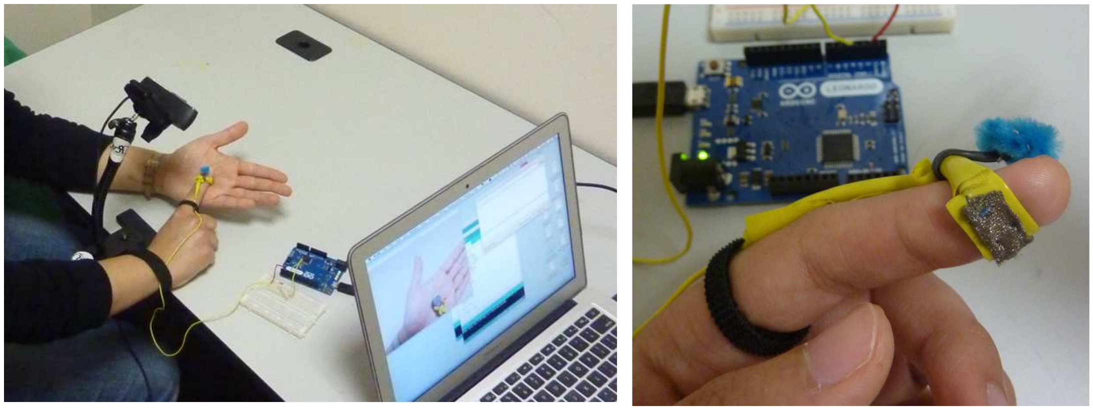
To answer these questions, we conducted a lab study with 12 participants with visual impairments. This study included two tasks: (1) user-defined on-body gesture creation to investigate what locations and gestures are preferred (e.g., touchscreen-style, location-specific), and what factors affect these preferences, and (2) a comparison between touchscreen and on-hand interaction with different hand counts (one versus two hands) to comparing subjective responses to on-body input versus mobile touchscreen input.

4.2. Experimental Methodology

To investigate the design of and subjective response to on-body interaction for people with visual impairments, we designed a single-session lab study consisted of two tasks. For the first task, we adapted a method employed by Weigel et al. [128] to examine gesture preferences at different on-body locations, and to evaluate these locations on factors such as social acceptability, comfort, and ease of use. The second task was designed to compare subjective responses to on-body input versus mobile touchscreen input, we implemented an on-hand sensing system that controls the VoiceOver software on an Apple iOS device and asked participants to complete basic mobile tasks with both on-hand input and a touchscreen smartphone.

4.2.1. Participants

Twelve people with visual impairments (6 male, 6 female) participated in the study with the average age of 44.3 (*SD* = 12.9, range 23–62). Nine participants were totally blind; six were born blind while the rest became blind later in life (years post onset: *M* = 22.8, *SD* = 14.4, range 3–42). Three participants had low vision. While nine



(a) Experiment setup (b) A finger-worn touch sensor

Figure 4.1. On-hand input sensing system used in Task 2. Participants wore a lightweight ring that included a color marker (tracked by a camera) and capacitive touch sensor. This ring could be placed so as not to cover the fingertip.

participants used touchscreen phones on a regular basis, the remaining three had feature phones. Participants reported using their phone at least once every few hours, with the exception of one participant who used it once a day.

4.2.2. Apparatus

For the second task, we made a lightweight ring including a color marker on the top and a capacitive touch sensor on the bottom as shown in Figure 4.1. In addition, we built a custom sensing system consisting of a tracking and a touch-detection module running on a laptop with an Intel Core i5 processor. A Logitech Webcam C930e was attached to a desk for tracking the x, y location of the color marker on the participant's gesturing finger. For touch detection, we ran another software on an Arduino Leonardo board using the *SoftwareSerial* and *CapSense* libraries. The laptop communicated timestamped finger locations to the Arduino software, which combined them with the touch state to classify the users' gestures. Finally, the Arduino converted the sensed gestures to VoiceOver keyboard shortcuts and sent them via Bluetooth to an iPhone 4S.

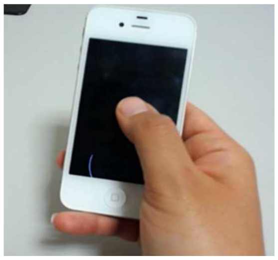
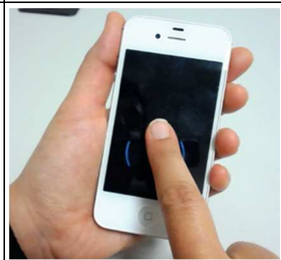
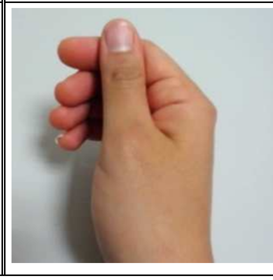
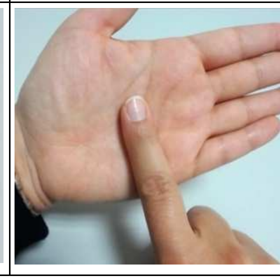
Phone		Hand	
One-handed	Two-handed	One-handed	Two-handed
			

Figure 4.2. The four input methods varying device (phone *vs.* hand), and hand count (one *vs.* two) for Task 2. Participants were instructed to use their thumb for one-handed condition, and the index finger for two-handed condition.

4.2.3. Procedure

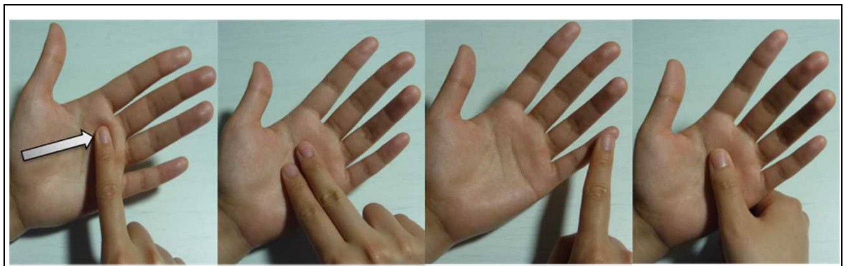
The study began with the first task with the goal to identify types of gestures (e.g., tapping, pinching) and locations that users would prefer for on-body interaction. Following a user-defined gesture protocol [128,131], we asked participants to create on-body gestures for five mobile actions (e.g., opening a selected app) at each of five different on-body locations: *same hand* (the same hand as the gesturing hand), *other hand-palm* (the palm of the other hand), *other hand-back* (the back of the other hand), *forearm*, and *neck and face* area. The mobile actions were chosen to cover a range of common mobile tasks and the locations were selected from Weigel et al. [128] but we focused on those where the skin would likely be exposed. The second task was designed to compare touchscreen and on-hand input varying number of available hands. The task was set up as a 2×2 within-subjects design, with factors of *device* (touchscreen phone *vs.* hand) and *hand count* (one *vs.* two) as illustrated in Figure 4.2. It began with brief training on basic iOS VoiceOver gestures including: horizontal flick to navigate left and right, double tap to open/activate selection, and pressing the home button (or, for on-hand input, a long press) to return to the home screen. For each condition, participants performed the same set of basic mobile actions (e.g., navigating through apps, opening a selected item). At the end of each condition, participants' subjective responses were collected in terms of ease of use, physical comfort, social acceptance and openness to perform inputs in different contexts. Participants were also asked to vote for their most and least preferred conditions at the end of each task.

4.2.4. Data Analysis

Because the work is exploratory, we did not have specific hypotheses. For subjective ratings, we specify which statistical tests we used throughout the Results section. In general, however, because the normality assumption of parametric tests may not hold for the 5-point rating scale data that we collected, we used non-parametric tests: Friedman tests, repeated measures ANOVAs with Aligned Rank Transform (ART) [130], and, for pairwise comparisons, Wilcoxon signed ranks tests. Bonferroni adjustments were used to protect against Type I error for all posthoc pairwise comparisons. For qualitative data, observation notes on gesture characteristics were recorded during the sessions and later categorized. We also conducted a qualitative analysis of the think-aloud data and other participant comments.

4.3. Results

We examined gesture creation strategies, location preferences for on-body input, and perceived trade-offs between touchscreen phone and on-body input depending the number of available hands.



	Basic gesture	Number of fingers	Specific fingers	Specific landmarks
Same hand	10	7	4	10
Other hand-palm	12	7	3	7
Other hand-back	12	9	3	7
Forearm	12	9	1	4
Neck and Face	12	6	4	11

Table 4.1. Four strategies for creating distinct gestures by varying (left to right): the basic gesture itself (e.g., taps or swipes), number of fingers, landmarked used, or the gesturing finger (e.g., thumb or index), and the number of participants who used a given strategy at each on-body location ($N = 12$). Multiple strategies could be used for each location.

4.3.1. Gesture Creation Strategies

During the gesture creation process for the first task, we collected 300 gestures (5 gestures × 5 locations × 12 participants). For all participants across all body locations, *directional swipe* was the most commonly used gesture for navigating to a *previous* or *next* item. For the other mobile actions, however, participants created a wider variety of gestures. Participants used the four following strategies to create their on-body gestures (Table 4.1): varying a common touchscreen gesture (e.g., swipe, single tap, double tap), varying the number of fingers (e.g., one *vs.* two), using specific body landmarks (e.g., pointing to a fingertip *vs.* palm), and varying which fingers were used (e.g., index *vs.* middle). Table 4.1 shows the number of participants who used each strategy while creating a set of gestures at each on-body location. All participants employed more than one common touchscreen gesture at each location (*i.e.*, varying the "*basic gesture*" in Table 4.1), except for with the *same hand* location, where two

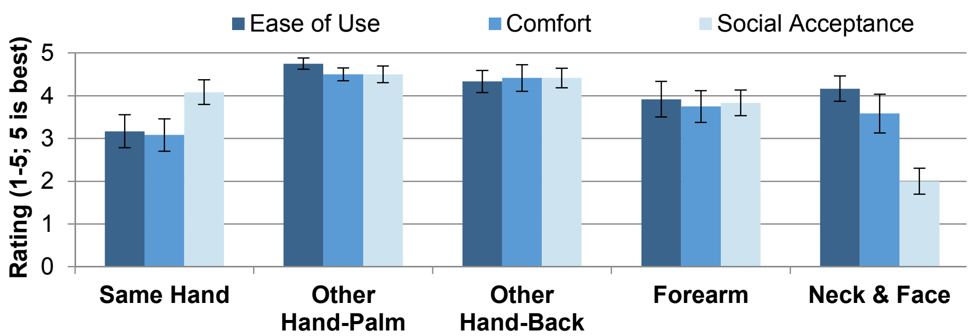


Figure 4.3. Average ratings for ease of use, comfort, and social acceptance for the on-body locations; 5-point scale (1-*least*, 5-*most*). Error bars are standard error. Locations on the other hand (both palm and back) consistently fared well ($N = 12$).

participants used only variations of a single tap gesture. The least frequent strategy was to vary which fingers were used; only six participants created distinct gestures by switching their fingers.

As found in the previous study in Chapter 3, touchscreen-like gestures and specific landmarks were most frequently observed. While gestures varying number of fingers was also recurrent, varying specific finger were less common. We did not find clear differences between different on-body locations.

4.3.2. On-body Location Preference

The location of a gesture on the body had an impact on the reported ease of use, physical comfort, and social acceptability of that gesture. Overall, of the five locations, the majority of participants favored *other hand-palm* (8 responses). *Same hand* and *other hand-back* also received two votes each. The *neck and face* was the least preferred location by 10 participants, while *same hand* received two votes. This overwhelming selection of *neck and face* as least preferred is particularly interesting given that its raw scores on ease of use and physical comfort were higher than *same hand* (Figure 4.3). This result suggests that the social unacceptability of *face and neck* overrode ease and comfort concerns. participants most preferred *other hand-palm*. We also examined on predicted use under physical constraints and for specific tasks.

4.3.2.1. Ease of Use

Figure 4.3 shows the rating scale results for ease of use, comfort, and social acceptability. The *other hand-palm* location received the highest average rating at 4.8, while the *same hand* received the lowest rating at 3.2. A Friedman test showed there was a significant effect of on-body location on ease of use ($\chi^2_{(4,N=12)} = 10.46$, $p = .033$). After a Bonferroni adjustment, no posthoc pairwise comparisons with Wilcoxon signed-rank tests were significant. When participants were asked to choose both the easiest and most difficult locations to use, *other hand-palm* was selected as easiest by 5 participants, followed by *other hand-back* (3 participants). The most common reasons for choosing locations on the other hand were that it was natural, offered a relatively wide input space compared to *same hand*, and that it was similar to using a mobile phone. For example: *"I can use my palm as a touchpad. It has enough space to perform any gesture."* (P11). In contrast, eight participants chose the *same hand* as the most difficult location to use, mostly because they found the interaction unfamiliar, for example: *"I'm not used to it. It's different."* (P4).

4.3.2.2. Physical Comfort

In terms of physical comfort, participants were again positive about *other hand-palm* and *other hand-back*, giving them mean ratings of 4.5 and 4.4 out of 5, respectively; see Figure 4.3. As with ease of use, the *same hand* was perceived to be least comfortable, at 3.2. A Friedman test showed that the impact of location on physical comfort was statistically significant ($\chi^2_{(4,N=12)} = 10.24$, $p = .037$). After a Bonferroni adjustment, no posthoc pairwise comparisons were significant.

In terms of the single most comfortable location, *other hand-palm* received the highest number of votes (7/12). For least comfortable, six participants chose *face and neck* and five chose *same hand*. The most common reasons for finding the face and neck to be uncomfortable were that it is relatively far from where hands natural rest, and that it is curved; three participants preferred flat surfaces for performing gestures. For *same hand*, participants were concerned they would be limited in the number of gestures they could comfortably perform, for example: *"You don't have the freedom to move around. The movements and gestures would be limited"* (P12).

4.3.2.3. Social Acceptance

For social acceptance, *other hand-palm* and *other hand-back* again fared well, receiving the highest ratings at 4.6 and 4.4 out of 5, respectively; *face and neck* was considered to be unacceptable (Figure 4.3). A Friedman test showed that there was a statistically significant impact of location on social acceptability ($\chi^2_{(4,N=12)} = 30.31$, $p < .001$); again, due to the Bonferroni adjustment, no posthoc pairwise comparisons were significant.

Other hand-palm was selected as the single most socially acceptable location by 8 participants, who appreciated that it allows for discreet use thanks to its similarity with everyday activities, for example: *"It doesn't draw a lot of attention"* (P2). Not surprisingly, all participants considered the *face and neck* to be the least socially acceptable, most commonly because it attracts too much attention, or interferes with

	Private	Crowded public	Non-crowded public	Workplace	Acceptance Rate (%) *M (SD)*
Same hand	12	9	11	9	85.5 (12.5)
Other hand-palm	12	10	12	12	95.8 (8.3)
Other hand-back	11	8	12	10	85.4 (14.2)
Forearm	11	6	9	8	70.8 (17.3)
Neck and face	9	1	4	3	35.4 (28.4)

Table 4.2. Number of participants who would perform on-body gestures in different contexts, with average acceptance rates across places (*N* = 12).

	Alone	Partner	Friends	Family	Colleagues	Strangers	Acceptance rate (%) *M* (*SD*)
Same hand	12	11	12	12	10	8	90.3 (13.4)
Other hand-palm	12	12	12	12	12	10	97.2 (6.8)
Other hand-back	12	12	12	11	11	9	93.1 (9.74)
Forearm	11	10	9	9	7	5	70.8 (18.1)
Neck and face	9	8	6	8	3	1	48.6 (26.6)

Table 4.3. Number of participants who would perform on-body gestures in front of different audiences, with average acceptance rates across audiences (*N* = 12).

other activities. For example, P1 said: *"You would be considered as rude or have bad manners [if gesturing on face] during conversations."*

We further investigated social acceptability in terms of place of use and audience (Tables 4.2 and 4.3). All three hand input locations had high acceptance rates regardless of place—private, crowded public, non-crowded public, and workplace; across the four places their average acceptance rates ranged 85–96%. The *forearm* had a somewhat lower acceptance rate (M = 71%), while the *face and neck* was generally unacceptable except in private. In terms of audience—alone, partner, friends, family, colleagues, and strangers—again, the *neck and face* had a much lower acceptance rate than other on-body locations. It was also interesting to note that *forearm* was again considered to be not as acceptable as the hand locations. One participant even said, while scrubbing his forearm: *"They may say I might have fleas"* (P12).

	Two-hands free			One-hand busy			Acceptance rate (%)
	Seated	Standing	Walking	Seated	Standing	Walking	M (SD)
Same hand	11	10	9	9	9	9	79.2 (7.0)
Other hand-palm	12	12	8	2	3	3	55.6 (38.6)
Other hand-back	11	12	8	3	5	5	61.1 (30.1)
Forearm	10	9	7	1	3	3	45.8 (30.6)
Neck and face	10	8	6	6	4	4	52.8 (19.5)

Table 4.4. Number of participants who would perform on-body gestures under different physical constraints, with average acceptance rates across constraints (N = 12).

	Hand-writing	Keyboard	Number pad	Sketching	Touchpad	Acceptance rate (%) M(SD)
Same hand	2	0	5	2	7	26.7 (23.1)
Other hand-palm	9	6	10	7	11	71.7 (17.3)
Other hand-back	7	4	8	5	11	58.3 (22.8)
Forearm	6	5	7	6	9	55.0 (12.6)
Neck and face	0	2	3	1	7	21.7 (22.5)

Table 4.5. Number of participants who would perform different types of input at each on-body location, with average acceptance rates across types of input (N = 12).

4.3.2.4. Physical Constraints

We expected that pose (seated, standing, or walking) and whether one hand is holding a cane or dog leash would be critical factors affecting on-body input for people with visual impairments. Table 4.4 shows the number of participants who were willing to perform gestures at each on-body location under these physical constraints. With two hands free, the majority of participants were willing to perform gestures at all on-body locations, whether seated, standing, or walking. For one hand holding a cane or leash, however, only the *same hand* location was popular, with nine participants willing to use *same hand* whether they were seated, standing, or walking. For two hands free, the responses suggest that participants may be less likely to want to make on-body gestures while walking than seated or standing, although further work is needed to confirm this possibility.

4.3.2.5. Input Type

To understand how participants would want to use each on-body location, we asked about five input types previously evaluated with sighted users by Weigel et al. [128]; see Table 4.5. *Other hand-palm* was seen as particularly flexible for supporting a range of input types. It was the most popular for handwriting, keyboard, number pad, and sketching, and was tied with *other hand-back* for touchpad-style input (e.g., taps, swipes). *Same hand* and *face and neck* were the least likely to be used; for example, no one was willing to use the *same hand* as a keyboard.

4.3.3. Touchscreen Versus On-hand Input with Different Number of Hands

For Task 2, we examined the subjective preference for phone versus on-hand input for different number of hands. Overall, the preferences differ depending on whether both hands are available or not. Again, we also collected responses on predicted use under physical constraints.

4.3.3.1. Overall Preference and Perceived Trade-Offs

The majority of the participants preferred the two-handed on-phone condition (8 responses); three preferred one-handed on-hand input, while one preferred two-handed on hand input. The least preferred condition was one-hand with the phone (8 responses). Below, we summarize participants' perceived advantages and disadvantages for touchscreen versus on-hand input.

While conventional touchscreen input (two-handed) was overall the most preferred, perhaps due to its familiarity, participants valued the advantages of on-hand interaction as well. A primary reason was that it allowed the phone to be safely stowed away, for example: *"You don't have to take out iPhone, you don't have to worry about* *getting it wet"* (P8). Some participants also commented that eliminating the need for the screen positively impacted ease of access and efficiency. For example, P9 said: *"You can just go right to your hand, you don't have to take the phone out. It could eliminate the screen."* Related, P7 commented on the aesthetic feel of the on-hand interaction, saying that it feels better not to have to interact with a piece of metal or glass.

All participants appreciated one-handed input for mobile computing because it can be important to have a hand free. For one-handed use, the on-hand input won out over the phone. One hand made it difficult to hold the phone and control it at the same time. P6, for example, commented that there was increased risk of dropping the phone, while P3 said: *"[It's] very uncomfortable, certain gestures can be mistaken for other gestures."* In comparison, two-handed use was considered easier, more accurate, and more stable. Four participants also noted that two hands allow for a greater variety of gestures. For example: *"It's easier because I have a free hand to maneuver the phone… do whatever you want to gesture"* (P11).

In general, all 12 participants felt it would be difficult to use their phone with both hands at times, particularly when they are walking. For example, P11 said: *"I have to stop walking and do the gestures and continue walking, or I have to wait until I get to the place where I can use it."* However, only six participants expressed the same concern with two-handed on-hand input. At the same time, even on-hand input with one hand was not always considered to be good, with five participants commenting on physical limitations of using the thumb for input with the one-handed use case. Two

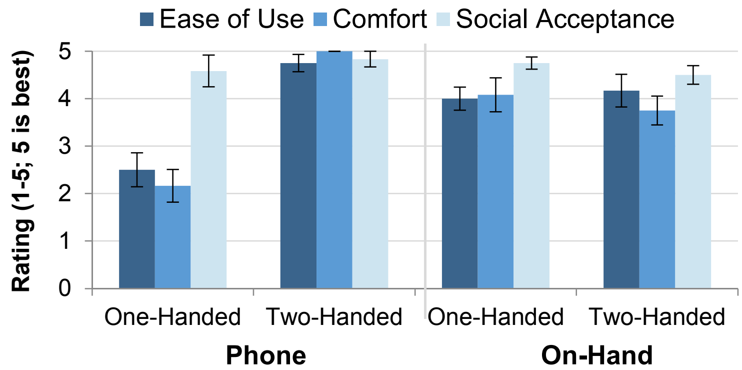


Figure 4.4. Average ratings for ease of use, comfort, and social acceptance for Task 2; 5-point scale (1-*least*, 5-*most*). Error bars are standard error (*N* = 12).

participants did not wish to interact with their phone at all when walking because of safety.

4.3.3.2. Ease of Use, Comfort, and Social Acceptability

Ratings on ease of use, comfort, and social acceptance are shown in Figure 4.4. To assess the effect of *device* and *hand count* on ease of use ratings, we ran a two-way repeated measures ANOVA with ART. The interaction effect between input methods (on-phone, on-hand) and number of hands was statistically significant ($F_{1,11} = 37.66$, $p < .001$, $\eta^2 = .77$). Posthoc pairwise comparisons using Wilcoxon signed rank tests showed that the phone was easier than the hand for two-handed use ($Z = -2.31$, $p = .021$). The opposite was true for one-handed use, with on-hand input being easier than the phone ($Z = -2.36$, $p = .018$). There was also a significant main effect of hand count ($F_{1,11} = 16.01$, $p = .002$, $\eta^2 = .59$) but the main effect for device was not significant.

Physical comfort ratings mirrored the ease of use results. A two-way repeated measures ANOVA with ART revealed a statistically significant interaction effect between input location and hand count ($F_{1,11} = 63.15$, $p < .001$, $\eta^2 = .852$). Again, posthoc pairwise comparisons using Wilcoxon signed rank tests showed that participants felt more physically comfortable using two hands on the phone ($Z = -2.77$,

	Two-hands Free			One-hand Busy			Acceptance rate (%)
	Seated	Standing	Walking	Seated	Standing	Walking	*M (SD)*
Phone One	6	6	6	6	6	6	50.0 (0.0)
Phone Two	12	12	8	1	2	2	51.4 (43.0)
Hand One	11	11	10	10	11	11	88.9 (4.3)
Hand Two	10	9	5	0	0	0	33.3 (39.1)

Table 4.6. Number of participants who would use on-body input under different physical constraints, with mean acceptance rates across constraints (*N* = 12).

p = .006), but that on-hand interaction was more comfortable for one-handed use (Z = -3.07, *p* = .002). There was no statistically significant main effect of input method, however, there was a main effect of number of hands ($F_{1,11} = 49.65$, $p < .001$, $\eta^2 = .82$).

Social acceptance ratings were high across all four conditions, ranging on average from 4.5 to 4.8 out of 5. A two-way repeated measures ANOVA with ART revealed no significant main or interaction effects of device and hand count on these ratings. While we asked about use of the input methods in different places and in front of different audiences as with Task 1, no clear trends emerged based on either contextual factor.

Table 4.6 shows the popularity of the four input conditions under different physical constraints—that is, seated, standing, or walking, and two hands free versus one holding a cane or dog leash. These results again reflect the trade-offs between on-hand and phone input when only one hand is free. While the two-handed input conditions were both popular when two hands are free and the participant is seated or standing, these numbers quickly drop off for walking, and drop even further when only one hand is available. One-handed on-hand input, however, was perceived to be the most versatile, with 10 or 11 participants out of 12 willing to use it regardless of the physical constraints.

4.4. Discussion

Reflecting on our findings from the study, we provide design implications for accessible on-body interaction for people with visual impairments.

4.4.1. Creation of Gesture Sets

The user-defined gestures created in Task 1 strengthened the suggestion from Chapter 3.4.1 to utilize the basic gestures that are now well-established with modern touchscreens (e.g., tap *vs.* swipe) when designing on-body interactions for mobile computing. Varying the number of fingers and using specific landmarks on the body can broaden the set of distinguishable gestures. Again, as mentioned in Chapter 3.4.2, landmarks, in particular, such as pointing to different parts of the hand, may be useful for rapid mode switching or to respond to a notification (as is done in *Imaginary Phone* [39]). Finally, participants rarely varied their gesturing finger (e.g., index *vs.* thumb), which mirrors touchscreen input preferences by sighted users [131].

4.4.2. Dominance of Hands for On-Body Input Locations

The results from Task 1 showed that the hand-based locations (e.g., *other hand-palm*) were better received by participants than the forearm or face and neck areas. This finding contrasts Weigel et al.'s [128] that the forearm to be the easiest and most comfortable location to use; note that their participants did not have any visual impairments, and they did not consider social acceptability when investigating on-body input location preferences. Supporting hand-based locations as the input surface for on-body interaction for people with visual impairments would allow more discreet and natural interaction compared to the relatively socially unacceptable forearm or face/neck locations.

4.4.3. Consideration of Physical Constraints

Relatedly, blind mobile phone users often have one hand busy with a cane or dog leash. Support for one-handed input is thus critical for supporting accessible information access on the go, a need that came out in participant comments. An issue that two participants commented on was that they hold their cane with their dominant hand, which was the same hand we had tested in our study. A system to support people with visual impairments would need to either allow for input on the same hand while also holding the cane or would need to easily support switching control to the non-dominant hand temporarily.

4.4.4. Social Acceptability

As found in our prior research on mainstream wearable devices [135], social acceptability plays an important role for on-body input for people with visual impairments. Our Task 1 findings suggest that participants prioritized social acceptability over ease of use and physical comfort by choosing the neck and face location as least preferred even while input on one hand was considered relatively hard to use and uncomfortable. Although a recent study showed that observers were more understanding of a device which is often socially unacceptable (namely, Google Glass) if it is used to support a person with disability [96], people with disabilities themselves may not feel comfortable using a device that would attract unwanted attention [111]. Thus, a careful consideration is needed to allowing discreet use.

4.4.5. Complementary Input Techniques for Always-Available Interaction

The preference for input locations depended on the number of hand availability in Task 2; the touchscreen input was preferred when two hands are free, while on-body input was preferred for one-handed interaction. This suggests that on-body input would be especially more desirable as a *complement* to the phone when both hands are available. A downside of one-handed on-body input as identified by some participants is that it does not offer the same kind of input flexibility as other locations. Even so, as a complementary form of input, it could be used to support a specific set of tasks more easily and quickly than pulling out and using a phone (e.g., controlling navigation instructions, notifications, and audio).

4.4.6. Importance of Hands-On Experience

Two-handed on-body input, namely *other hand-palm*, was more preferred than one-handed on-body input (*same hand*), in Task 1, when various on-body locations were compared amongst others. However, participants' preferences changed to being more positive about one-handed on-body input than two-handed on-body input, in Task 2, after they had hands-on experience with the actual prototype for on-body interaction, compared against smartphone interaction. This change highlights the importance of having users interact with working systems rather than assigning too much weight to subjective responses collected in the largely imaginary scenarios that the user-defined gestures method traditionally employs (e.g., [128,131]).

4.5. Limitations

In terms of limitations, the primary shortcoming is the disparity in how familiar participants were with touchscreens versus on-body input in Task 2, likely leading to a bias toward the touchscreen input; a multi-session study with more in-depth tasks with the interactive system could partly address this issue. Moreover, having to grasp the phone for one-handed on-phone conditions in Task 2 may also have affected results; for example, users who need to interact using only one hand may employ other solutions such as securing the phone with a hand strap or a belt clip. Additionally, while on-body input is meant to support mobile information access, we conducted the study in a controlled lab setting and participants were seated while using the system. Different contexts of use may impact participants' reactions to the input. Finally, while most past work on wearable and on-body input has focused on sighted users, we did not include sighted users in our study. A direct comparison would be useful to understand the differences between sighted and visually impaired users' needs.

4.6. Conclusion

The findings from this study provided a better understanding of how to design accessible on-body input for people with visual impairments compared to the previous exploratory study. The results confirmed the tendency (already seen in Chapter 3) to create gestures that are commonly used for touchscreen devices. Also, we found that the palm-side of the other hand is the most preferred location for on-body input because of its natural and discreet use, while the *neck and face* area is the least preferred location. The findings also suggest that users prioritize social acceptability over ease of use and physical comfort. Lastly, we observed that input performed on the hand was preferred to touchscreen input for one-handed use when performing location-independent gestures for basic mobile tasks (e.g., menu navigation), while it was the opposite when two hands were available.

The prior study in Chapter 3 and this study provide insight into the subjective responses to and the design of on-body interaction for VI users. However, the question of how on-body interaction impacts input performance compared to touchscreen interaction for VI users remains unanswered. We address this question in the next chapter with a controlled lab study.

Chapter 5: A Performance Comparison of On-Hand Versus On-Phone Nonvisual Input by Blind and Sighted Users

5.1. Motivation and Introduction

Focusing on subjective feedback, findings in Chapter 4 showed that blind and low-vision participants reacted positively to the idea of on-body input, in particular preferring the hand to a touchscreen phone for location-independent gestures when only one hand is free (e.g., when the other hand holds a cane or dog leash). The hand as an input surface was also considered to be more discreet and natural than other body locations.

While most studies of on-body interaction have included visual output [40,42,43,76,117], only a smaller number have investigated non-visual use with sighted users [38,39,69]. Gustafson et al. [38], for example, assessed non-visual pointing performance and found that sighted users could point to targets more precisely on their hand than on a touchscreen phone. These on-body tactile benefits may be even stronger for users with visual impairments, for whom tactile acuity has been found to be greater than it is for sighted users [20,32,34,63,116]. However, this performance question remains unexplored, as [38] only collected data from one blind participant. Thus, we conducted a study to answer the following main research questions:

- *Does on-hand input outperform touchscreen input?*
- *Is there any performance difference between blind and sighted users?*

To assess the performance of on-body input with blind users, we designed and conducted a study comparing non-visual input on the hand versus on the phone with 12 sighted and 11 blind participants. Compared to the simple location-independent gestures (e.g., press-and-hold, swipe right) that we previously studied in Chapter 4, here, we focused on absolute pointing and more complex gestures involving two hands. Absolute pointing, pointing to a specific location on the hand or phone, could result in highly efficient interaction if it is accurate. This study included two tasks, the first of which was a controlled *pointing task* with 20 target locations on the hand or phone, based on the Gustafson et al.'s study [38]. The second task moved beyond pointing, by comparing input of more complex shape gestures on the hand and phone (e.g., a circle and a '+' sign) in terms of performance and subjective feedback.

5.2. Experimental Methodology

To assess the performance of non-visual on-body input compared to input on a flat surface for blind and sighted users, we designed and conducted a single-session study with 23 participants (11 blind). Our study includes two non-visual tasks. The first task builds on Gustafson et al.'s [38] work with sighted users by (1) including blind participants, and (2) examining a more thorough set of target locations per participant (20 *vs*. 5)—this latter difference allows for an analysis of input performance based on location. In designing the study tasks, our goal was to make each condition as realistic as possible while still conducting a controlled performance experiment. This motivation underlies many of our study decisions, such as using the hand's natural landmarks to configure pointing targets rather than having targets of uniform size.

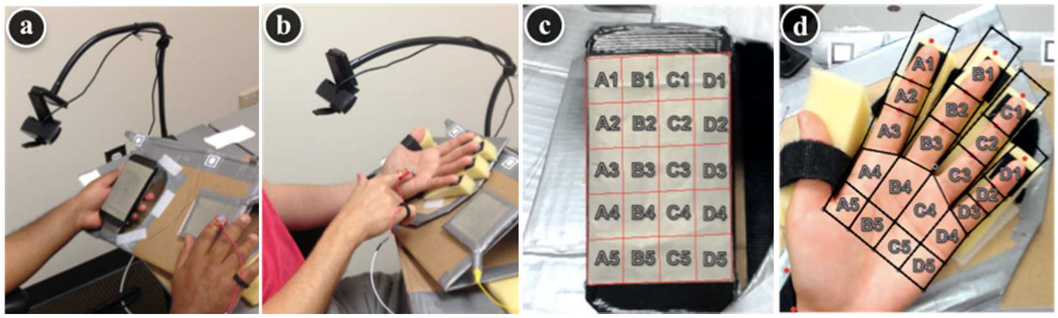


Figure 5.1. Experimental setup showing the camera and approach used to stabilize (a) the phone or (b) non-dominant hand during tasks, and corresponding target locations (c, d). At the start of each trial, participants tapped on the square touchpad on the right side of the stand. The entire setup was reversible for left-handed participants.

5.2.1. Participants

Participants were recruited via campus e-mail lists and local organizations that serve people with visual impairments. In recruiting participants, we addressed two limitations commonly seen in studies comparing visually impaired and sighted users: variation in vision levels of the visually impaired group and an age discrepancy between the two groups. Twelve sighted (7 female, 11 right-handed) and 11 totally blind (5 female, 8 right-handed) individuals participated in this study.[6] Sighted participants were on average 51.8 years old (*SD* = 11.9, range 26-67) versus 52.4 years old (*SD* = 10.8, range 33-67) for blind participants. Six of the blind participants had become blind later in life (years post onset: *M* = 31.0, *SD* = 16.7), while five were born blind. All participants had touchscreen phone experience, and they were compensated for their time.

5.2.2. Apparatus

As shown in Figure 5.1, the custom experimental system consisted of a Logitech 1080p HD webcam C930e, touch sensors connected to an Arduino Leonardo board, and

[6] Four more participants (1 sighted) were initially recruited but excluded from analysis because they were unable to learn the target-naming scheme in the pointing task even after training.

tracking software running on a laptop with an Intel Core i5 processor and OSX 10.9.4. To ensure data consistency across the hand and phone conditions, we used this tracking setup for both interfaces rather than using the native touchscreen sensing on the phone. The main software was written in C++ and used the OpenCV library for image processing. The system: (1) tracked the pointing finger, (2) detected touch on the phone or hand, (3) automatically generated pointing targets, (4) provided audio feedback upon touch, and (5) included a conductive touchpad for participants to tap at the start of each trial.

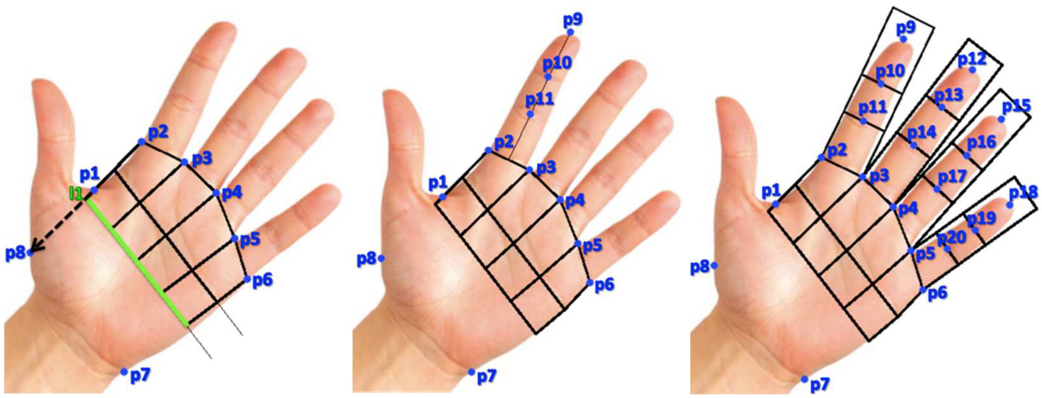
5.2.2.1. Stabilizing the Phone and Hand

We affixed the phone or non-dominant hand to a stand to ensure that it did not move and the camera angle remained steady during study tasks. The stand itself was angled for comfort, based on feedback from pilot participants. Cutouts in the stand allowed the participant to hold the phone, which was attached to the stand by Velcro (Figure 5.1a). For the hand, Velcro straps fastened the wrist and thumb to the stand. To prevent curling of the fingers—which would impact the size and shape of targets—the fingers rested on slightly angled foam wedges and the fingernails were secured to the foam using small pieces of Velcro tape (Figure 5.1b). This allowed the upper (input) side of the hand to be completely bare, so as not to hinder tactile feedback. The rotation of the phone or hand could be adjusted to be comfortable for each participant during an initialization step. Finally, the entire setup was reversible, to support both left- and right-handed participants.

(a) Top view.

(b) Bottom view.

Figure 5.2. Finger-worn touch sensor.



(a) Targets on the palm. (b) Targets on the fingers. (c) After configuration.

Figure 5.3. Configuring targets for the hand condition for a right-handed user. See main text for detail.

5.2.2.2. Finger Tracking and Touch Detection

Unlike previous studies that used depth information alone to track the fingertip and detect touch (e.g., [27,39,42]), we found such an approach to be insufficient for precise measurements. Instead, we combined a color marker for *x, y* tracking with a separate lightweight touch sensor on the pointing finger (Figure 5.2). The color marker was placed 5 mm down from the tip of the participant's finger. The *x*, *y* coordinates of the touched point were determined by image moments of the camera frame after filtering colors and removing noise. The touch sensor consisted of conductive thread that was shielded from the user's skin with 3*mm*-wide non-conductive tape. It connected to an Arduino Leonardo board running software that used the *CapSense* library to detect changes in capacitance from touching the pointing finger to the user's non-dominant hand. For the phone condition, conductive fabric covered the screen, allowing for the same touch detection approach.

For the phone, a researcher clicks the corners of the input area, which is then split evenly into columns and rows. For the hand, we laid out 20 targets where the five rows start at the fingertips and do not cover the bottom half of the palm. The choice to include 20 targets (5 rows × 4 columns) on both the phone and the hand follows Gustafson et al. [38], and is based on smartphone home screens that typically lay out icons in 5 or 6 rows × 4 columns. A set of reference points is used to maximize the use of natural landmarks by aligning targets with fingertips, phalanges of the fingers, thumb versus palm, webs between fingers, and border with the wrist. As shown in Figure 5.3, points from $p1$ to $p6$ segment the palm from the fingers and thumb; $p7$ is the outer join between palm and wrist. For the thumb, this segmentation also requires extending $\overleftrightarrow{p2p1}$ and demarcating its intersection with the edge of the palm ($p8$). To automatically generate palm targets, $\overline{p2p8}$ and $\overline{p6p7}$ are each divided into four equal segments, the top two of which delineate the two rows of the palm. These rows are in turn divided into columns based on the finger webs, but with equal widths at the bottom; the outside two cells are also adjusted to reach the edge of the hand (compare Figure 5.3a to Figure 5.3b).

Finally, to generate finger targets, the researcher selects the fingertips (e.g., $p9$) and, on the line segment from the midpoint of the finger base (e.g., $\overline{p2p3}$) to the tip, selects the natural divisions between phalanges (e.g., $p10$ and $p11$). The targets are then automatically generated as shown in Figure 5.3c. In pilot studies, participants regularly pointed to the very top of the finger, which meant that the pointing finger would be touching, yet *above* the non-dominant hand's finger. To support this common interaction, we extended the height of the top row targets on the hand by 50% of the width of the target.

5.2.3. Procedure

Participants completed both the *pointing* and *shape-drawing tasks* for one interface condition (hand or touchscreen), followed by the other interface condition. The order of presentation for hand versus phone was fully counterbalanced within each participant group; the pointing task always preceded the shape-drawing task. To ensure non-visual performance, sighted participants were blindfolded during the tasks. Each session lasted two hours. The touch sensor was placed on the index finger of the participant's dominant hand, and the hand and phone were affixed to the stand as described above (see Figures 5.1a and 5.1b).

5.2.3.1. Pointing Task

Participants first explored the names and locations of targets by running their pointing finger over the hand or phone. Based on touch-and-explore interfaces, the system read each target aloud as it was touched. The selection of a target occurred on finger lift-up. For each trial, participants tapped on the starting touchpad to reposition their hand, and to initiate timing and cause the instruction to be read aloud. Upon a correct selection, a chime sound played from the laptop; no audio feedback was provided for incorrect selection. For example, if a participant makes a contact, the name of the touched location will be read aloud. Then participants can either lift up their finger to confirm the location or keep browsing for the correct target location depending on the current target. As brief practice, participants performed one random practice trial before beginning the main task.

Participants then performed three blocks of 20 trials as a *learning* phase and another three blocks as a *trained* performance phase (120 trials in total). Each block included the 20 locations in Figures 5.1c and 5.1d, presented in random order. Participants were asked to find the target location as quickly and accurately as possible. They were allowed to retry the trial if they realized they had misunderstood where the location was after hearing the target name, for example, pointing to D1 instead of A1, or if they needed to have the target name to be repeated. After the task, participants provided subjective feedback on ease, accuracy and speed.

5.2.3.2. Shape Drawing Task

Participants drew five shape gestures: circle, equilateral triangle, square, plus sign (+), and equal sign (=). This set was chosen to cover a variety of characteristics such as curviness/straightness of lines, length of lines, shape closure, and angle between two strokes (e.g., parallel, perpendicular). Beforehand starting the task, a brief verbal description of each gesture was given and participants were allowed to explore a raised physical guide for each shape (geometric shapes or mathematical symbols may not be familiar to all blind participants [6,48]). Participants then completed six blocks of trials including one practice and five test blocks, where each block consisted of the five shapes presented in random order (25 trials in total). As with the pointing task, participants tapped on the starting touchpad to begin each trial. Participants were asked to take as much time as they needed to draw the shape accurately and consistently on their palm. For feedback, a brief, high-pitched sound played at every *touch-down* or *touch-up* event.

5.2.3. Experimental Design and Hypotheses

For the pointing task, the study used a 2 × 2 × 2 mixed factorial design, with user *Group* as a between-subjects factor (levels: sighted *vs.* blind), *Interface* as a within-subjects factor (levels: hand *vs.* phone), and *Phase* as a within-subjects factor (levels: learning *vs.* trained). For the shape drawing task, the study used a 2 × 2 mixed factorial design, with *Group* as a between-subjects factor and *Interface* as a within-subjects factor.

We derived the following hypotheses based on the findings from Gustafson et al. [38] that sighted users were faster at pointing on their hand than on a phone when visual cues are absent, and a number of studies that show blind individuals have higher tactile acuity than sighted individuals [20,32,34,63,116]:

- H1: The hand is faster for pointing than the phone.
- H2: The hand is more accurate for pointing than the phone.
- H3: The pointing performance benefits (speed and accuracy) of the hand are greater for blind participants than for blindfolded sighted participants.
- H4: The hand results in more accurate shape gestures than the phone (compared to an ideal reference shape).
- H5: The hand results in more consistent shape gestures than the phone when gestures are redrawn repeatedly.
- H6: The shape-drawing benefits (accuracy and consistency) of the hand are greater for blind participants than for blindfolded sighted participants.

5.2.4. Data and Analysis

For both tasks, the system continuously logged timestamped x, y coordinates and touch status for the index finger. For the target pointing task, we collected data from six blocks of 20 trials from 23 participants for both phone and hand conditions. To reduce the influence of outlier trials, we removed 58 trials that were three standard deviations above and below the mean per participant. There were 16 miss-recorded trials, leaving a total of: 6×20×23×2-58-16 = 5446 trials. For the shape drawing task, five blocks of 5 trials for two interfaces were collected from 23 participants, for a total of: 5×5×2×23 = 1150 trials.

We specify which statistical tests were used throughout the results. In general, we apply paired *t*-tests and repeated measures ANOVAs for pointing accuracy and speed. Wilcoxon signed-rank tests and repeated measures ANOVAs with Aligned Rank Transform (ART) [130] were used for subjective ratings, and for shape accuracy and consistency, which violated the normality assumption of the parametric tests. For posthoc pairwise comparisons (*t*-tests or Wilcoxon signed-rank tests), Holm's sequential Bonferroni adjustments were used to protect against Type I error [45]. Finally, the audio recordings were analyzed to thematically group participants' comments and open-ended responses.

5.3. Results

We assessed accuracy, speed and subjective measures, as well as conducting a secondary analysis of performance across target locations on the phone and the hand for the *pointing task* and preliminary recognition rates that could be achieved with a gesture recognizer for the *shape-drawing task*.

5.3.1. Pointing Task

We report on the primary performance measure of selection time, which encompasses both speed and accuracy, and a secondary measure of the accuracy of only the first point of contact. We also examine performance based on target size and location.

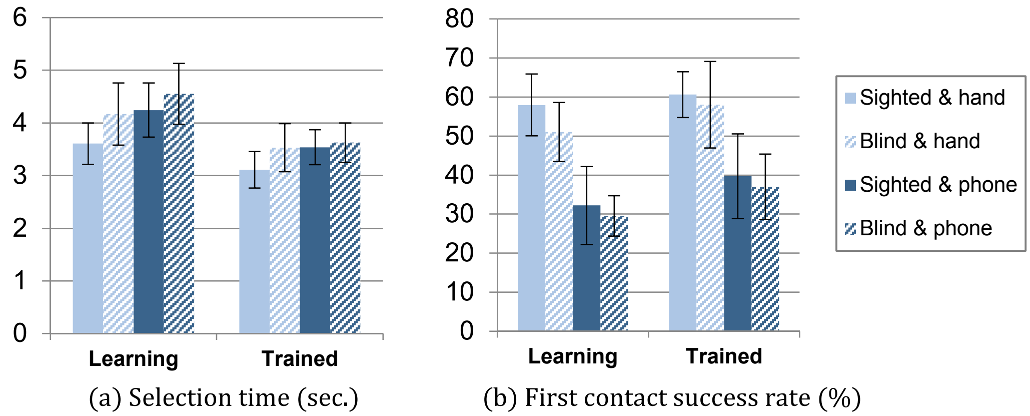


Figure 5.4. Average selection time (a), and average first contact success rate (b) for the learning and trained phases of the pointing task ($N_{sighted} = 12$; $N_{blind} = 11$). Error bars indicate 95% confidence intervals. Sighted particpants were blindfolded during the task.

5.3.1.1. Target Selection Time

Selection time per trial was defined as a comprehensive performance measure that includes an implicit time penalty for errors. It was calculated from the starting signal (when the participant tapped the start touchpad) until a finger up event occurred on the *correct* target location (Figure 5.4a). This measure is impacted both by proprioception, as the participant moves their pointing finger through the air, and by tactile feedback, after the finger touches down and moves on the surface of the opposite hand. Supporting H1, the average selection time on the hand was 3.59s ($SD = 0.77$), which was faster than the 3.98s for the phone ($SD = 0.72$). A 2 × 2 × 2 (*Group*, *Interface*, *Phase*) repeated measures ANOVA found that the difference was statistically significant, by main effect of *Interface* ($F_{1,21} = 7.17$, $p = .014$, $\eta^2 = .25$). Participants also improved significantly between the learning and trained phases (main effect of *Phase*: $F_{1,21} = 102.18$, $p < .001$, $\eta^2 = .83$). As shown in Figure 5.4a, the average time in the learning phase was 4.13s per target ($SD = 0.75$), compared to only 3.45s per target in the trained phase ($SD = 0.60$). No other main or interaction effects were significant.

Although we had expected to see greater performance advantages on the hand for blind participants than for blindfolded sighted participants, no support was found for H3.

5.3.1.2. First Contact Success Rate

While selection time, above, encompassed both speed and overall pointing accuracy, we also isolated *first contact success rate* as the percentage of trials where the participant's first touch point landed within the bounds of the target (see Figure 5.4b). This secondary measure of accuracy relies solely on proprioception. Because selection occurs on lift up, these rates are not comparable to standard error measures, but do provide insight into one aspect of performance efficiency. Supporting H2, the hand was more accurate than the phone, with an average accuracy of 57.0% (SD = 12.8) across groups, compared to 34.8% for the phone (SD = 14.4). A 2 × 2 × 2 repeated measures ANOVA (*Group* × *Interface* × *Phase*) revealed that this difference was significant, by a main effect of *Interface* ($F_{1,21} = 57.6$, $p < .001$, $\eta^2 = .73$). Accuracy also improved significantly from the learning phase to the trained phase, jumping from 42.8% (SD = 11.9) to 48.9% (SD = 12.9) by main effect of *Phase* ($F_{1,21}$=12.38, $p = .002$, $\eta^2 = .37$). Finally, although we had hypothesized that the performance advantages of the hand would be greater for blind participants than for blindfolded sighted participants (H3), no other main or interaction effects were significant.

5.3.1.3. Impact of Target Location and Size

As a secondary analysis, we computed selection time and first contact success rate for each target on the hand and phone. To reflect more experienced use, this analysis includes only the trained phase data. As well, because of the lack of performance

Figure 5.5. Heat maps for average selection time (left; sec.), and first contact point accuracy (right; %) per target in the pointing task, averaged across participants (with *SD* in parentheses) (*N* = 23). The fingertips resulted in particularly strong performance results.

differences between the blind and blindfolded sighted user groups above, we combined data from the two.

As Figure 5.5 shows, while performance was generally better on the hand than on the phone, there was also a greater range in results across targets. The fingertips, for example, appear to be particularly fast and accurate. To broadly compare the impact of different target locations on selection time and first contact success rate, we grouped the targets by row and by column, and conducted one-way repeated measures ANOVAs with the following single factors for each device (hand and phone): *Rows* (5 levels: from fingertip to palm) and *Columns* (4 levels: left to right / index to baby finger). We report only posthoc pairwise comparisons that were significant at $p < .05$ after a Holm-Bonferroni adjustment.

Compared to the phone, the performance of each target on the hand could be affected more by its location because some location might have more distinctive landmarks than other locations (e.g., fingertip *vs.* palm). For the hand, rows and columns both significantly impacted speed (*Rows:* $F_{4,88} = 7.396$, $p < .001$, $\eta^2 = .252$; *Columns*: $F_{3,66} = 4.359$, $p = .007$, $\eta^2 = .165$). For rows, the fingertips and the top row on the palm (fourth row overall) offered a speed advantage. Pairwise comparisons showed that participants were significantly faster pointing to the fingertips than to the second, third and fifth rows, and were faster with the fourth row than the fifth row. For columns, the third column was slower than the rightmost column. Similarly, rows and columns both significantly impacted first contact success rate (*Rows*: $F_{4,88} = 21.13$, $p < .001$, $\eta^2 = .490$; *Columns*: $F_{3,66} = 4.037$, $p = .011$, $\eta^2 = .155$). Similar to the selection time results, pairwise comparisons showed that the fingertips (top row) were more accurate than all other rows, and the fifth row was less accurate than all other rows. For columns, the third column was less accurate than the first two.

For the phone, different rows did not significantly impact speed, but columns did ($F_{3,66} = 5.431$, $p = .002$, $\eta^2 = .198$). Posthoc pairwise comparisons showed that the outer edges were the fastest—the leftmost column was significantly faster than the middle two columns. No significant effects were found for first contact success rate. The more marked performance differences across locations for the hand could be due at least partly to variation in target size across location, a decision that we had purposely made to ensure that the hand condition was realistic and made use of physical landmarks. Hands also varied from one participant to the next in size and shape. To investigate the relationship between target size on the hand and the measures of selection time and first contact success rate, we computed Pearson's correlation coefficients for each measure. Although statistically significant due to the large sample size (23 participants × 20 targets), the correlation between target size and speed was negligible in magnitude ($r = -.026$, $n = 460$, $p < .001$). However, a moderate positive correlation was found between target size and first contact success rate ($r = .424$, $n =$

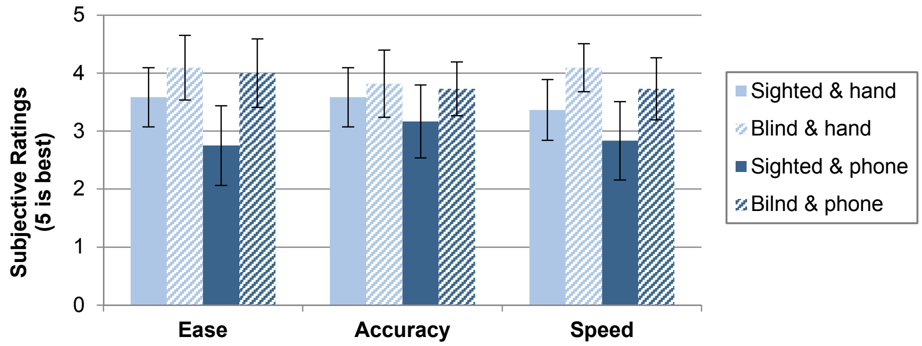


Figure 5.6. Average subjective ratings for ease, accuracy, and speed for the pointing task ($N_{sighted} = 12$; $N_{blind} = 11$). Error bars indicate 95% confidence intervals.

460, $p < .001$). While our study design does not allow us to isolate the impact of target size on the performance measures, these results suggest that size may play some role.

5.3.1.4. Subjective Feedback

In terms of overall preference, 9 out of 12 sighted participants preferred the hand to the phone, while blind participants were more evenly split, with 6 votes for hand and 5 for phone. This trend could be due to blind participants' familiarity with non-visual interaction with a touchscreen phone, since all of them had a smartphone. For example, one blind participant, B7, said: *"I'm more familiar with the phone, I've been an iPhone user for almost two years"*. Participants also rated the hand and the phone in terms of subjective ease, accuracy and speed using 5-point scales (5 is best); see Figure 5.6. For each measure, we ran a 2 × 2 repeated-measures ANOVA with ART. Blind participants reported generally higher ease and speed ratings compared to sighted participants, perhaps due to their comfort level with non-visual interaction. Significant main effects of *Group* on ease ($F_{1, 21} = 8.57$, $p = .008$, $\eta^2 = .29$) and speed ($F_{1, 21} = 5.80$, $p = .025$, $\eta^2 = .22$) were observed. No other main or interaction effects were significant.

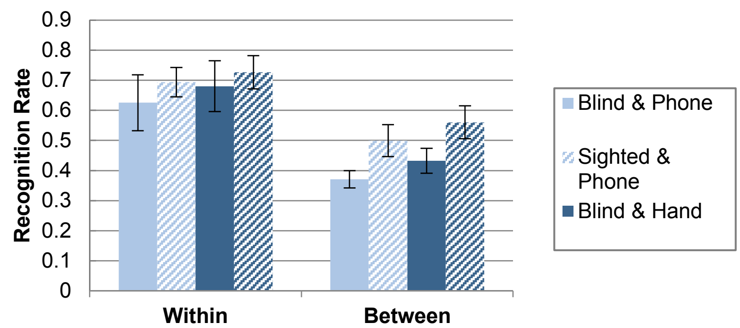


Figure 5.7. Average recognition rate for the shape drawing task for both within and between participants ($N_{sighted}$ = 12; N_{blind} = 11). Error bars indicate 95% confidence intervals. Sighted participants were blindfolded during the task.

5.3.1.5. Summary

The speed results support H1 and confirm Gustafson et al.'s [38] conclusion that the hand allows faster target pointing than the phone for non-visual use. Furthermore, we extended this result to show that it applies to blind users who already have experience with non-visual interaction. We also found support for H2, which provides new insight on first touch location accuracy, showing that it is higher on the hand than the phone (likely due to proprioceptive differences). No support was found for H3 that the performance benefit is greater for the blind group than the sighted group. Per-target analysis revealed that target location impacted pointing performance, particularly on the hand (e.g., fingertips *vs.* palm). Finally, though not conclusive, sighted users may have a stronger preference than blind users for the hand compared to the phone for non-visual pointing input.

5.3.2. Shape-Drawing Task

To explore the feasibility of supporting non-visual gestural input, we computed gesture recognition rates based on the drawn shape gestures. We also assessed geometry-based accuracy and consistency measures of the gestures.

5.3.2.1. Recognition Rate

Consistency and accuracy are important for shape gestures because they will ultimately impact recognition rates for a gesture recognizer. To assess the practicality of shape gestures on the hand versus touchscreen, we applied the $N multi-stroke recognizer, which is meant for fast prototyping of shape-based gestures and does not require many training examples [8]. Recognition rates were calculated twice: (1) 5-fold cross-validation *within* a single participant by training on four gesture examples and testing on the remaining one, and (2) *across* participants in the same user group by testing on each participant after training on the rest. The results are summarized in Figure 5.7.

For the recognition rates we ran a 2 × 2 × 2 (*Group*, *Interface*, *Training Set*). Overall, the hand resulted in significantly higher recognition rates than the phone ($F_{1,21} = 13.61$, $p = .001$, $\eta^2 = .39$), and gestures created by blindfolded sighted participants were more accurately recognized than blind participants' gestures ($F_{1,21} = 7.09$, $p = .015$, $\eta^2 = .25$). As one would expect, for *Training Set*, rates were significantly higher if training was personalized within each participant as compared to the user group as a whole ($F_{1,21} = 150.00$, $p < .001$, $\eta^2 = .88$). No other main or interaction effects were significant.

5.3.2.2. Geometry-based Accuracy and Consistency Measures

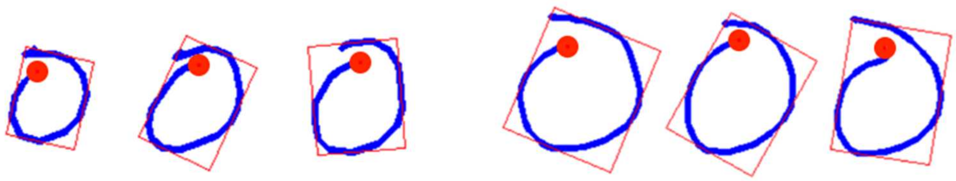
While the recognition rate analysis indirectly requires that gestures be accurate and consistent, we also explicitly examined geometry-based accuracy and consistency of the shapes collected. Accuracy was defined as the *absolute difference* between the drawn shape and an ideal shape (e.g., a perfect square). For the circle, equilateral triangle, and square, we calculated this difference for *aspect ratio*, the ratio of width to

		Blind		Sighted	
		Phone: *M* (*SD*)	**Hand:** *M* (*SD*)	**Phone:** *M* (*SD*)	**Hand:** *M* (*SD*)
Accuracy	Aspect ratio[a]	0.36 (0.22)	0.39 (0.33)	0.28 (0.11)	0.30 (0.08)
	Closure[a] (px)	30.17 (15.59)	34.84 (18.60)	25.95 (12.75)	28.96 (14.60)
	Angle[b] (°)	7.77 (3.37)	11.29 (5.66)	8.11 (4.90)	10.81 (5.57)
	Length Ratio[b]	0.20 (0.11)	0.25 (0.17)	0.14 (0.08)	0.16 (0.04)
Consistency	Aspect Ratio[a]	0.17 (0.07)	0.14 (0.05)	0.24 (0.20)	0.27 (0.25)
	Closure[a] (px)	11.26 (12.14)	7.02 (8.12)	6.58 (3.92)	8.43 (6.74)
	Angle[b] (°)	4.59 (3.49)	4.18 (3.28)	2.81 (2.01)	6.70 (6.43)
	Length Ratio[b]	0.10 (0.07)	0.12 (0.13)	0.07 (0.06)	0.08 (0.04)
	Size[ab] (px^2)	3751(7838)	3407 (5536)	1000 (1178)	2157 (1216)

Table 5.1. Accuracy and consistency measures for Task 2. Accuracy is computed as the absolute difference between the raw measure and an ideal shape, and consistency is the standard deviation across each participant's five test trials per shape. Smaller numbers are better. The metrics with 'a' were used for circle, triangle, and square. The metrics with mark 'b' were used for the rest. ($N_{sighted} = 12$; $N_{blind} = 11$).

height (ideally 1), and *closure* [53], the Euclidian distance between the start and end points of the gesture (ideally 0). For the plus and equal signs, we calculated the *angle* between the two strokes (ideally 90° for '+' and 0° for '=') and the *length ratio* of the shortest stroke to the longest stroke (ideally 1, which represents equal length). Consistency for each of these measures was defined as the standard deviation across the five test trials each participant drew per shape. Finally, we also looked at the area of the minimum bounding box and consistency of that size.

The results are inconclusive regarding these accuracy and consistency measures. Examining the raw means for accuracy measures shown in Table 5.1, all measures are closer to the ideal shape on the phone than on the hand if compared within the blind group or the sighted group. This trend is contrary to H4 and H6, although 2 × 2 (*Group* × *Interface*) repeated measures ANOVAs with ART for each measure revealed no significant main or interaction effects. For the consistency measures, the same analyses revealed a few significant effects, although no clear picture emerged. The significant effects were: main effect of *Interface* on consistency of size ($F_{1,21} = 9.89$, $p = .005$, $\eta^2 = .32$), interaction effect between *Group* and *Interface* on consistency of size



(a) Phone (b) Hand

Figure 5.8. Examples of circles on the phone and on the hand by blind participant B5. The rectangles show the minimum bounding box for each shape, and the starting point of each stroke is marked with a dot. The sizes of the shapes are larger on the hand.

($F_{1,21}$ = 5.85, p = .025, η^2 = .22), and interaction between *Group* and *Interface* on shape closure ($F_{1,21}$ = 5.56, p = .028, η^2 = .21). It is unclear whether the results would change with a larger sample size.

Finally, participants created bigger gestures on their hand than on the phone (e.g., Figure 5.8). A 2 × 2 repeated measures ANOVAs with ART revealed a significant main effect of *Interface* on size ($F_{1,21}$ = 19.15, p < .001, η^2 = .48). No other main or interaction effects were significant.

5.3.2.3. Subjective Feedback

In terms of overall preference for shape gestures, 7 out of 12 sighted participants preferred the hand to the phone, compared with only 4 out of 11 blind participants. These trends are similar to the first task. Participants who preferred the hand valued its tactile feedback. For example, B7 said: *"You can feel where you started and where you ended. You may have more control to draw the shape."* In contrast, the flatness of the phone was the most popular reason for favoring it over the hand, mentioned by 4 sighted and 7 blind participants. For example, B2 said: *"phone is sort of drawing on a paper, because it's flat. [Because of the] valleys and peaks, I never know if the square on the hand was really a square."*

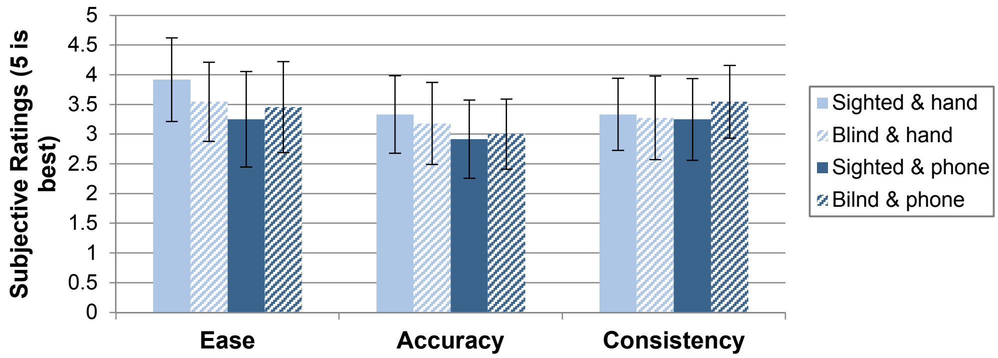


Figure 5.9. Subjective ratings for ease, accuracy, and consistency for shape drawing task ($N_{sighted}$ = 12; N_{blind} = 11). Error bars indicate 95% confidence intervals.

Participants also rated the two interfaces on ease, accuracy and consistency using 5-point scales (5 is best), as shown in Figure 5.9. For each rating, we ran a 2 × 2 repeated-measures ANOVA with ART but no main or interaction effects were significant.

5.3.3. Summary

The speed results support H1 and replicate Gustafson et al.'s [38] conclusion that the hand allows faster target pointing than the phone for non-visual use. Furthermore, we extended this result to show that it also applies to blind users who already have more experience with non-visual interaction. We also found support for H2, which provides new insight on first touch location accuracy, showing that it is higher on the hand than the phone, and that it varies based on location of the target—targets on the fingers were more accurate than on the palm. No support was found for H3.

While the geometric analyses were inconclusive and did not provide support for H4-H6, the gesture recognition rate results provided indirect evidence that the hand results in more consistent and/or accurate gestures than the phone. These recognition rate findings thus provide some support for H4 and H5. Subjective preference trends were similar to the pointing task, with more sighted than blind participants preferring the hand.

5.4. Discussion

Our findings both replicate and broaden Gustafson et al.'s [38] study of non-visual pointing performance by sighted users on the hand versus a phone. Most importantly, we extended their results to blind users, showing that on-body input offers an alternative to the touchscreen phone as a means of accessible mobile interaction. Our results also show that the location of the first touchdown on the hand is more likely to be within the intended target's bounds than it is on the phone. This finding suggests that proprioception even before the hands touch is partly responsible for the performance advantage of the hand.

5.4.1. Difference Between Sighted Versus Blind Users

Because blind individuals have higher tactile acuity than sighted individuals [20,34,63,116], we had expected that the performance benefits of the hand would be particularly noticeable for the blind participant group. However, there was no difference for the pointing task, possibly due to the difference between tactile acuity is simply too small to matter (e.g., the tactile grating detection and 2-point gap discrimination studies [34,116] showed differences of 0.33mm and 0.37mm, respectively). Moreover, for shape drawing task, blind users had significantly lower recognition rates than sighted users, which may be due to differences in spatial cognition ability (e.g., [93,118]) or even simply due to lower familiarity with the shapes that were used [6,48].

Another somewhat contrary trend, though not statistically significant, suggests that sighted participants were more likely than blind participants to prefer the hand to the phone. This could be due to blind participants having more experience and familiarity with non-visual interaction on touchscreen devices, since all blind participants owned a smartphone. Further work is needed, however, to confirm whether these preferences would remain unchanged with more realistic or longer-term use of on-body interaction.

5.4.2. Pointing on Fingers and Shape Drawing on Palm

Despite perhaps not being more beneficial for blind users than sighted users, our findings suggest that pointing to targets on the hand is a viable and efficient input technique for accessible mobile computing. In designing future on-hand interfaces, the fingertips, which are known for high acuity (e.g., [25,115]), would be good locations for frequently needed shortcut commands as participants were faster and more likely to touch down immediately within a target's bounds in these regions than in other areas. In addition, pointing performance differs depending on which finger the target is located (e.g., index or ring finger), and should be taken into account when placing targets. As with the advantage of the fingertips, these differences across fingers may be at least partly due to known acuity differences (e.g., [63,123]). In terms of drawing shapes, palm may offer better gesture recognition accuracy than on the phone. As such, future work should explore the potential for accessible finger-specific or shape-based gestural shortcuts.

5.5. Limitations

When designing the target layout for the pointing task, our focus was to maximize the role of natural landmarks of the hand. We thus adapted the target layout to each participant's hand to provide a more realistic, ecologically valid assessment of on-hand input performance than could be achieved by replicating the rectangular shape of a phone on the hand. As a result, and as must have also been the case in [38], the average target size on the hand was bigger than the phone: $6.31cm^2$ (SD = 2.19) on average compared to $4.03cm^2$. Although we attempted to mitigate this difference by using a relatively large phone, further work is needed to clearly separate the impacts of size and other layout factors from tactile and proprioceptive feedback.

Another limitation is that we may have found different results had we provided more training for each task, considering that on-hand input was new to participants. It may also be useful to examine potential differences between *early-blind* (who became blind at birth or a young age) and *late-blind* individuals, to control visual experience in assessing performance on spatial tasks [32,64]. Tactile acuity is also affected by factors such as age [34], so a matched pair design that takes age into account could offer more experimental power in future work.

To reduce input variability and facilitate accurate sensing, we stabilized the hand and phone during data collection. A more realistic scenario would allow both hands to move freely and could uncover additional issues such as the need to keep the hands within the camera's field of view.

Finally, due to random recruitment, our participants might not be representative of the population for each user group. Thus, further investigation would be needed with a representative sample to accurately reflect the entire population.

5.6. Conclusion

Although the benefit of on-body input for blind users might not be *greater* than for sighted users for non-visual interaction, the results from this study confirmed the findings from Gustafson et al. [38] that users can benefit from on-body interaction for a non-visual target pointing task, compared to smooth touchscreen. We also extended this finding to blind users. Furthermore, our findings show that the hand offers better first touch location accuracy and results in higher shape gesture recognition rates than the phone—a new contribution compared to [38]. Furthermore, the findings from this study allowed us to gain insights for designing non-visual on-body interaction.

We now turn from investigating interaction design and input performance questions with on-body interaction to how to accessibly sense such interaction for VI users.

Chapter 6: Sensing Location-Specific On-Body Gestures Using Finger- and Wrist-Worn Sensors

6.1. Motivation and Introduction

To enable on-body interaction, researchers have explored a variety of approaches such as using body-mounted cameras [23,27,39,40,117,126] and capacitive sensors [68,74,106,127]. However, the interaction space is often constrained by the placement and range of the sensor, thus not feasible for supporting location-specific gestures that are found to be preferred by VI participants for on-body interaction (Chapters 3 and 4). For example, *Skinput* [43] can detect simple touches at a variety of locations, but cannot recognize more complex gestures. In contrast, *FingerPad* [24] and *PalmGesture* [126] can sense shape gestures performed on the fingertip or palm, but cannot easily be extended to multiple locations. Moreover, while camera-based sensing techniques are common for on-body input because optical images can offer rich contextual information, capturing in-focus images without occlusion or framing issues is often challenging for people with visual impairments, who may require assistance (e.g., [3,129]).

To support complex gestures at multiple body locations and to provide an accessible on-body input sensing approach for VI users, we employed finger- and wrist-mounted sensors. Mounting sensors on the gesturing hand enables collocation of touch, sensing and feedback (*i.e.*., tactile feedback from the user's skin), and it may also prevent out-of-frame (e.g., gesturing finger is beyond the camera) or occlusion issues (e.g., gesturing finger is not visible/hidden by other objects) that can occur with hand-held or body-mounted cameras. Furthermore, this approach can extend the input space beyond limited locations such as the palm, and support contextual location-specific on-body input (e.g., tapping the wrist to check the current time, swiping on the ear to dismiss an incoming call), which will also increase input vocabulary by combining input locations and gestures compared to having only location-independent gestures. The main research question we had was:

- *To what extent can we reliably support on-body input sensing using finger-, and wrist mounted sensors?*

To answer this question, we investigated a finger-based, multi-sensor approach to detect location-specific on-body gestures. We developed a physical prototype included (i) a finger-worn multi-sensor package with two infrared (IR) reflectance sensors, an inertial measurement unit (IMU), and a small camera with an adjustable LED for illumination; and (ii) a wrist-worn IMU as would be found in a smartwatch. Then, we collected on-body input data from 24 participants to evaluate the effectiveness of the sensors individually and in combination. The dataset includes: (i) a set of eight basic gestures (e.g., taps, swipes) at three of the above locations that offer a large surface area (palm, wrist, thigh) for a total of 24 location-specific gestures; and (ii) the same eight gestures performed only on the palm, but at three different speeds.

6.2. Prototype Hardware

Our physical prototype consists of two wearable components: (i) a multi-sensor package worn on the finger with two IR sensors, an inertial measurement unit (IMU), and a small camera with an adjustable LED for illumination; and (ii) a wrist-worn microcontroller with its own IMU—which was intended to simulate the sensing that a

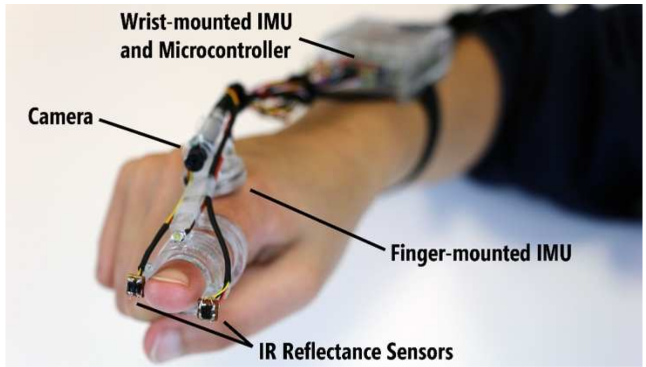


Figure 6.1. Our prototype hardware, consisting of a finger-mounted multi-sensor package (camera, inertial motion unit, and infrared reflectance sensors) and a wrist-mounted microcontroller and additional inertial motion unit.

smartwatch can provide. The finger-based sensors were mounted on three laser-cut rings and positioned to avoid impeding the user's sense of touch.

For the segmentation of the start and end of gestural input, the two IR sensors[7] (each 2.9mm × 3.6mm × 1.7mm) were used with the sensing range of ~2–10mm. As shown in Figure 6.1, these sensors were mounted on the sides of the front-most ring, approximately 5mm from the fingertip to avoid interfering with tactile sensitivity. For the gesture recognition, two IMUs[8] were used to investigate the role of sensor position (*finger vs. wrist*) on performance. These were mounted on the user's hand: one below the camera on the index finger and one on the wrist. The IMUs provide motion information at ~190 *Hz* with nine degrees of freedom—each contains a three-axis accelerometer, gyroscope, and magnetometer. A small (6*mm* diameter) CMOS camera[9] was also used to collect 640 × 640 px images at 90 fps. For consistent lighting, a bright LED (3mm diameter) was mounted below the camera lens. The IR and IMU sensors

[7] Fairchild Semiconductor QRE113GR
[8] Adafruit Flora LSM9DS0
[9] Awaiba NanEye GS Idule Demo Kit

were connected to a microcontroller[10] mounted on a Velcro wristband, and the camera and microcontroller were connected to a desktop computer via USB cables.

6.3. Input Recognition

To recognize localized on-body interaction, we developed a four-stage approach: Stage 1 – touch segmentation, Stage 2 – feature extraction, Stage – 3 location classification, and Stage 4 – gesture classification. For the touch segmentation stage, we used only IR sensor readings, while we explored features from all types of sensors for the rest of the stages.

6.3.1. Stage 1: Input Detection and Segmentation

We first segmented the input stream by detecting *touch-down* and *touch-up* events using IR sensor readings. The IR values represent distance from the touch surface (lower values are closer). While for real-world use, a segmentation approach would need to identify these touch events within a continuous stream of data, here the segmentation was done within a bounded trial that contains a single gesture. Based on experiments with pilot data, we developed a straightforward threshold-based approach; within a trial, a *touch-down* event is defined as an IR value lower than the threshold, while a *touch-up* event is the opposite. To be conservative, we segment the entire touch gesture from the first *touch-down* event in the trial to the last *touch-up* event. The threshold (θ) was set as below:

$$\theta = \begin{cases} 0.9 * IR_{max} \; if \; single \; IR \; is \; used \\ \min(0.9 * IR_{max}^{1}, 0.9 * IR_{max}^{2}) \; otherwise \end{cases}$$

[10] Sparkfun Arduino Pro Micro (5V/16MHz)

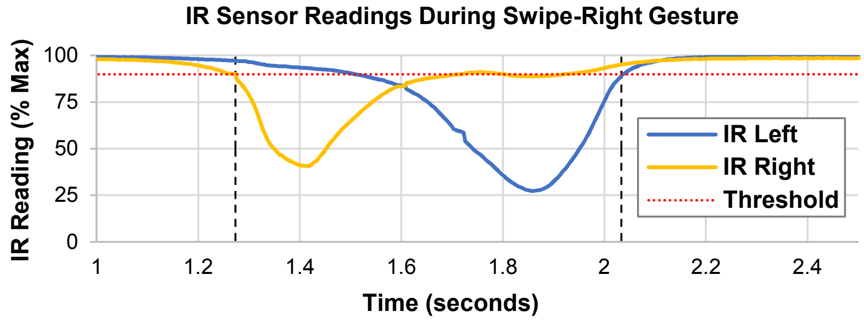


Figure 6.2. An example of left and right IR sensor readings during a swipe-right gesture performed on the thigh. Dashed lines indicate automatically segmented start and end times.

For a single IR sensor, for example, the threshold was set to 90% of the maximum IR values observed across the input stream.

Figure 6.2 shows the IR sensor readings for a *swipe-right* gesture on the thigh, illustrating why two sensors are needed for segmentation. With this left-to-right movement, the left sensor detects the touch first, but if used alone would segment the end of the gesture prematurely because of the direction of the gesture and shape of the surface. We crop each input stream to include only those sensor readings and video frames that lie between the *touch-down* and *touch-up* event timestamps.

6.3.2. Stage 2: Feature Extraction

We extracted static orientation and visual features for localization, and motion features for gesture classification.

6.3.2.1. Features for Input Localizations

To extract static features for localization, we first determined the video frame that has the maximum focus in the segmented sequence, where the focus is defined as the total number of pixels extracted using a Canny edge detector [22], tuned with a small aperture ($\sigma = 3$) and relatively low thresholds ($T_1 = 100, T_2 = 50$). We then used the timestamp of this video frame to extract static features from the IMU and IR sensors.

For the IMUs, we computed the orientation of users' finger and wrist, estimated by applying a Madgwick filter [71] on a sequence of raw accelerometer, magnetometer, and gyroscope readings. For the camera, we extracted texture features using local binary patterns (LBP), a common choice due to computational efficiency and robustness to changes in illumination (see [79,88]). Finally, features were extracted from IR sensor values for localization.

6.3.2.2. Motion Features for Gesture Classifications

For gesture classification, we extracted the following motion features from the sensor readings within the segmented timeframe.

IMU and IR. We used three preprocessing steps on the raw IMU and IR sensor readings, including smoothing, normalization, and resampling. We first smoothed the raw values using a Gaussian filter (σ =13, optimized based on pilot data) to reduce the effect of sensor noise, and then normalized the smoothed sequence by subtracting the mean and dividing by the standard deviation of the sequence. To obtain a fixed length sequence and improve robustness across different speeds, we resampled the sensor readings at 50 equally spaced discrete time steps, as in [133]. These values, however, were still sensitive to small variations in speed and orientation. Thus, as in [133], for each IMU and IR sensor we compute summary statistics for sets of 20 samples at 10-step increments (*i.e.,* four windows): mean, minimum, maximum, median, and absolute mean. Finally, for the 50 resampled accelerometer, magnetometer, and gyroscope readings, we computed *x*-*y*, *x*-*z*, and *y*-*z* correlations (also as in [133]). As a result, 79 features for each IMU and 70 for each IR sensor were extracted, which we used individually or concatenated together when classifying gestures.

Camera. We also extracted 2D motion feature vector from the full segmented sequence of camera video frames, estimated using a template-matching approach; matched against the successive two frames using a sliding window to compute the normalized cross-correlation [67]. Then, to reduce the impact of tracking noise, we smoothed the motion estimates by applying a moving average (window size = 10). We then re-sampled 50 points from this temporal sequence of motion vectors and computed statistics as with the IMU and IR sensor readings to obtain a fixed-length vector of 840 features for use in gesture classification.

6.3.3. Stage 3: Location Classification

For localization, we relied primarily on static visual features from the camera as well as IMU and IR reflectance sensor values. We classified the image using a support vector machine (SVM), commonly used in texture classification (e.g., [28,56,134]), that was trained on the computed texture features. We used a χ^2 kernel, which is known to perform well with LBP histograms [4].

During the classification, we resolved ambiguities using a sensor fusion approach. We combined predictions from the static visual features from a video frame with predictions from the IMU orientation and IR reflectance features with the same timestamp as that frame. Since the scales, lengths, and types of these feature vectors were all very different, we instead trained a separate SVM with a Gaussian kernel on the non-visual features rather than concatenating the features for use with a single classifier. To robustly combine the predictions from two disparate classifiers, we first tuned the SVMs to output normalized probability predictions for each class using Platt scaling, as is standard [95]. We concatenated these predictions into a single feature

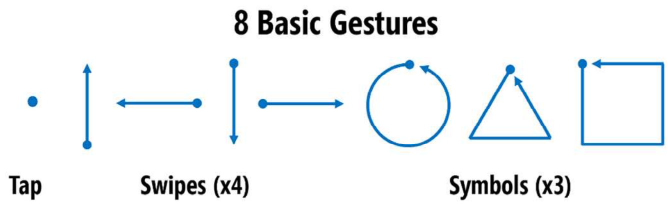


Figure 6.3. Participants were asked to perform these eight gestures on three different body locations (palm, wrist, and thigh) or on palm with three different speed levels (slow, medium, fast).

vector, which we then used to train a third sensor fusion classifier that automatically learned how to prioritize sensors based upon prediction confidence and location class. Inspired by [62], we used a feedforward neural network for this sensor fusion classifier. Our network had one fully connected hidden layer for flexibility of functional representation, and a softmax output layer to allow for multiclass output; it is trained using resilient backpropagation [73].

6.3.4. Stage 4: Gesture Classification

Individual support vector machine (SVM) gesture classifiers were trained for each of the locations where gestures are performed. Thus, once a location was predicted in Stage 4, the gesture was classified using the matching SVM. Each SVM was trained using all of the computed IMU, IR, and camera-based motion features concatenated into a single feature vector as in [133]. In our experiments, we combined features in various combinations (see Results). As with texture, SVM classifiers are commonly used for classifying gesture features because they are robust and efficient for problems with high dimensionality. We used a linear kernel with feature weights that were tuned to maximize performance across all participants. We trained three location-specific SVMs (*palm*, *wrist*, and *thigh*) to classify the gestures shown in Figure 6.3: *tap*, *swipe up*, *swipe down*, *swipe left*, *swipe right*, *circle*, *triangle*, and *square*.

6.4. Data Collection

To evaluate the accuracy of our finger- and wrist-mounted sensing system and the effectiveness of the sensors individually and in combination, we first collected input data from 24 participants including a set of eight basic gestures (e.g., taps, swipes) as shown in Figure 6.3 at three different locations that offer a large surface area (palm, wrist, thigh), and the same eight gestures performed only on the palm, but at three different speed levels (slow, medium, fast).

6.4.1. Participants

Twenty-four right-handed participants (16 female) were recruited via campus e-mail lists and word of mouth. Their average age was 28.9 (*SD*=7.95, range 19–51). Participants were compensated $25 for their time.

6.4.2. Apparatus

For data collection, participants wore the prototype described earlier in Chapter 6.2. We selected ring sizes to fit the participant's finger and adjusted positioning to ensure a consistent sensor range. A custom application in C# was written to display task prompts and a live feed from the finger-worn camera to assist with framing the target locations. Sensor readings from the IMU and IR sensors and video frames from the camera were logged along with timestamps and manual touch location and gesture labels to use as our ground truth. The raw data readings along with timestamps were saved to a log file and accompanying video file. Velcro straps were used on the upper arm to prevent the USB cables from interfering with the participants' range of motion. The focus of the camera was adjusted prior to the data collection.

6.4.3. Procedure

The procedure began with a brief demographic questionnaire, and lasted up to 90 minutes. After a brief demographic questionnaire and instrumentation of the hardware prototype, three types of on-body input were collected for each participant in the following order: (1) location-specific touches, (2) location-specific gestures, and (3) gestures with various speed levels.

6.4.3.1. Location-specific Touches

Participants made static touches (*i.e.*, touch and hold the finger down) at 15 locations including locations on the *palm*, *fingers*, *ear*, *shoulder* and *thigh*. Each trial began with a visual prompt of the target location shown on the monitor along with an audio alert. Participants were asked to hold their finger down and visually confirm on the monitor that they were pointing to the correct location. The experimenter confirmed that the location was correct and the image was in focus, then pressed a button to mark the current timestamp and trigger the start of the next trial. Participants completed 10 blocks of trials, where each block consisted of a different random permutation of the 15 locations (150 trials in total). In total, this dataset includes 3600 location-specific touches, across all participants.

6.4.3.2. Location-specific Gestures

We defined a set of eight basic gestures: *tap*, *swipes* in the four cardinal directions, and shape-based *circle*, *square*, and *triangle* gestures (Figure 6.3). Participants performed these gestures at three body locations: the *palm*, *wrist*, and *thigh*. These locations were selected because they have a relatively large input area, thus allowing for more complex gestures, and, practically speaking, they are easy to access and unobtrusive [33,97,125].

As with the first task, participants completed 10 blocks of trials, where each block consisted of a different random permutation of the 24 gesture and location combinations (240 trials in total). In total, this dataset includes 5,760 location-specific gestures across all participants.

6.4.3.3. Gestures With Various Speed Levels

To assess robustness across different gesture articulation speeds, participants performed the set of eight gestures on their palm at three different speeds: *fast*, *medium*, and *slow*. Participants performed 10 blocks of trials per speed, with each block consisting of a different random permutation of the eight gestures (240 trials in total). At the beginning of each of the three speed levels, the tester demonstrated all eight gestures at that speed and guided the participant through a brief practice session to ensure consistency. The order of presentation for speed levels was fully counterbalanced across participants. In total, this dataset includes 5,760 gestures, varying speed across all participants.

6.5. Experiment and Results

We evaluated our on-body input sensing system with the collected data for location-specific gesture classification, along with a robustness analysis across gesture speeds, with a brief verification that both IRs are better for segmentation than either IR alone. We use classification accuracy (% correct) to measure performance. All experiments use leave-one-out 10-fold cross validation to train and evaluate our algorithms. For the location-specific gesture classification, we compare sensor combinations using paired t-tests and use a Holm-Bonferroni adjustment to protect against Type I error [45].

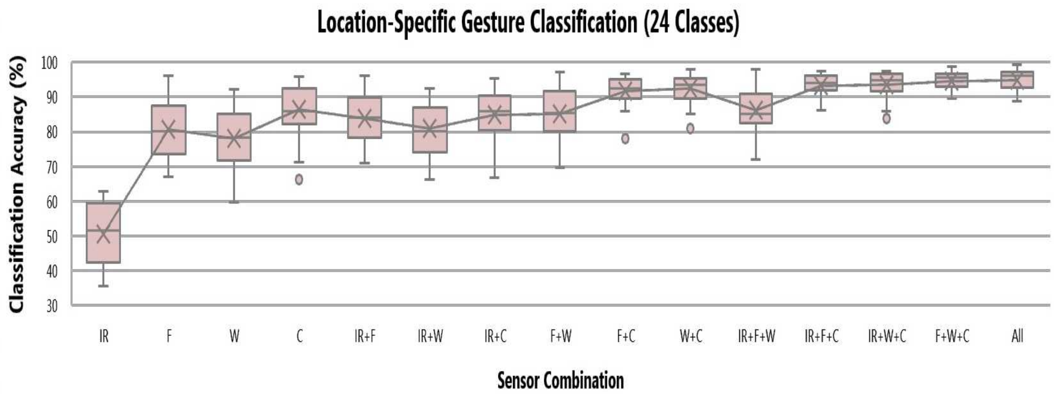


Figure 6.4. Classification accuracy of location-specific gestures for different sensor combinations averaged across 24 participants, showing that the camera (C) is necessary but not sufficient for gesture classification. Legend: IR = infrared reflectance, F = IMU_{finger}, W = IMU_{wrist}, C = camera). Means are marked by '×'.

6.5.1. Location-Specific Gesture Classification

To explore how well different sets of sensors can be used to recognized basic touch input performed at different locations, we conducted the same experiment with location-specific on-body gestures with 24 classes (3 locations × 8 gestures) to explore the performance of each sensor combination. The results are shown in Figure 6.4. This analysis includes all four stages of our Input Recognition approach, but does not require geometric verification in Stage III because the thigh, palm and wrist are coarse-grained locations.

The highest accuracy of 94.9% was achieved by combining all sensors (*SD* = 2.9). As with the touch localization, above, the camera is the best single sensor 86.3% (*SD* = 7.7), but in general adding more sensors improves accuracy. The camera alone is more accurate than the IR sensors or the wrist-based IMU (IMU_{wrist}) alone, although there is no statistically significant difference between the camera and finger-based IMU (IMU_{finger}), at 80.6% (*SD*=8.3). The accuracy comparisons among different sensor combinations that are statistical significant is shown in Table 6.1.

Comparison	t_{23}	*d*	*M (SD)*[1]	*M (SD)*[2]
Single sensors				
C *vs.* IR	21.01	4.26	86.3 (7.7)	50.4 (9.1)
F *vs.* IR	13.53	3.54	80.6 (8.3)	50.4 (9.1)
W *vs.* IR	13.49	3.28	78.0 (8.1)	50.4 (9.1)
C *vs.* W	4.71	1.05	86.3 (7.7)	78.0 (8.1)
Best single sensor (C) vs. two sensors				
C *vs.* C+W	-5.03	-1.03	86.3 (7.7)	92.5 (4.1)
C *vs.* C+F	-4.17	-0.90	86.3 (7.7)	91.7 (4.2)
C *vs.* W+IR*	3.56	0.74	86.3 (7.7)	80.8 (7.0)
Two-sensor combinations				
C+W *vs.* W+IR	9.15	1.95	92.5 (4.1)	80.8 (7.0)
C+F *vs.* W+IR	7.50	1.82	91.7 (4.2)	80.8 (7.0)
C+F *vs.* C+IR	7.02	1.23	91.7 (4.2)	84.9 (6.8)
C+F *vs.* F+IR	6.87	1.41	91.7 (4.2)	83.8 (7.0)
C+W *vs.* C+IR	6.84	1.38	92.5 (4.1)	84.9 (6.8)
C+W *vs.* F+IR	6.72	1.55	92.5 (4.1)	83.8 (7.0)
C+W *vs.* F+W	5.70	1.25	92.5 (4.1)	85.1 (7.5)
C+F *vs.* F+W	4.41	1.11	91.7 (4.2)	85.1 (7.5)
Best 2-sensor combination (C+W) vs. 3 sensors				
C+W *vs.* F+W+IR	5.28	1.19	92.5 (4.1)	86.2 (6.4)
C+W *vs.* C+F+W	-4.38	0.58	92.5 (4.1)	94.5 (3.1)
C+W *vs.* C+W+IR	-4.05	0.31	92.5 (4.1)	93.6 (3.7)
Three-sensor combinations				
C+W+IR *vs.* F+W+IR	5.95	1.45	93.6 (3.7)	86.2 (6.4)
C+F+IR *vs.* F+W+IR	5.85	1.37	93.3 (3.8)	86.2 (6.4)
C+F+W *vs.* F+W+IR	7.48	1.69	94.5 (3.1)	86.2 (6.4)

Table 6.1. Statistically significant comparisons between sensors combinations for location-specific gestures. Analysis included 37 total comparisons within the groups listed above (e.g.,** all pairs of single sensors). Holm-Bonferroni adjustment applied. All displayed results are significant at $p < .001$, except those marked '*' ($p = .002$). Legend: IR = infrared reflectance, F = IMU_{finger}, W = IMU_{wrist}, C = camera). The mean and standard deviation for each pairwise comprisons are also presented (marked '1' for the former, and '2' for the latter in the order that is presnted in first column).**

Overall, camera-based combinations are best for accurate location-specific gesture classification than combinations without the camera (see Table 6.2 for detail comparisons), though there is flexibility in what specific set of IR or IMU sensors are combined with it. With only two sensors involved, the camera plus either the finger- or wrist-worn IMU are the most accurate options, with accuracies rising significantly compared to the camera alone, to 91.7% (*SD*=4.2) and 92.5% (*SD*=4.1), respectively. These two camera-based combinations are also better than any other pairs of sensors. With three sensors involved, the camera-based options are all significantly more accurate than the combination of IR plus the two IMUs; no statistically significant

Comparisons	t_{23}	*d*	w/ camera	w/o camera
Singer non-optical sensor with or without the camera				
$\text{IMU}_{\text{finger}}$	7.36	1.69	91.7 (4.2)	80.6 (8.3)
$\text{IMU}_{\text{wrist}}$	12.0	2.26	92.5 (4.1)	78.0 (8.1)
IR	25.2	4.29	84.9 (6.8)	50.4 (9.1)
Two non-optical sensors with or without the camera				
$\text{IMU}_{\text{finger}}$ + $\text{IMU}_{\text{wrist}}$	7.00	1.64	94.5 (3.1)	85.1 (7.5)
$\text{IMU}_{\text{finger}}$ + IR	8.30	1.69	93.3 (3.8)	83.8 (7.0)
$\text{IMU}_{\text{wrist}}$ + IR	10.49	2.14	93.6 (3.7)	80.8 (7.6)
Three non-optical sensors with or without the camera				
$\text{IMU}_{\text{finger}}$ + $\text{IMU}_{\text{wrist}}$ + IR	7.85	1.23	94.9 (2.9)	86.2 (6.4)

Table 6.2. Average accuracies for 24 location-specific gesture classification of sensor different sensor combinations with or without the camera (the values within the parenthesis indicate the standard devications). The last two columns show the results of paired t-tests with Holm-Bonferroni adjustments. All displayed results are significant at $p < .001$ ($N = 24$).

differences are found between the camera-based combinations. Furthermore, the three-sensor combination with the highest accuracy, camera plus both IMUs (M = 94.5, SD = 3.1), is not significantly different from including all four sensor options.

6.5.2. Robustness to Variation in Gesture Speed

To evaluate the robustness of our approach to differences in gesture articulation speed, we compared gesture classification accuracy for individual sensors with the speed-varied dataset (8 gestures on the palm at 3 different speeds). The average duration of slow gestures is 2.8s (SD = 0.8), medium is 1.8s (SD = 0.5), fast is 1.2s (SD = 0.2). Our findings show that the accuracies from the IMUs are relatively robust to different speeds (> 93%); however, the camera and IR accuracies both drop as speed increases. Descriptive statistics are shown in Table 6.3. A 4 × 3 two-way repeated measures ANOVA with factors of *Sensor Type* (4 levels: *IR*, *finger IMU*, *wrist IMU*, and *camera*) and *Speed* (3 levels: *Slow*, *Med*, *Fast*) reveals a significant interaction effect between the two factors ($F_{6,276} = 17.75$, $p < .001$, $\eta^2 = .14$). This result confirms that the different

Speed	IR (%)	Finger IMU (%)	Wrist IMU (%)	Camera (%)
Slow	89.4 (9.1)	97.1 (3.0)	94.3 (6.5)	98.0 (2.6)
Med	85.9 (9.3)	97.9 (1.9)	93.6 (7.9)	96.3 (3.1)
Fast	74.7 (15.3)	94.9 (3.9)	93.2 (4.9)	82.8 (11.8)

Table 6.3. Mean (*SD*) of gesture classification accuracy (%) per sensor type and speed. Compared to the IMUs, the IR and camera were negatively affected by speed (*N* = 24).

sensors were impacted to differing degrees by the variation in speed. As suggested by Table 6.3, there were also significant main effects of *Sensor Type* ($F_{3,69} = 37.13$, $p < .001$, $\eta^2 = 0.42$) and *Speed* ($F_{2,46} = 25.32$, $p < .001$, $\eta^2 = 0.22$).

6.5.3. Touch Segmentation

To investigate the potential benefit of one versus two IR sensors on touch segmentation, we compare localized gesture classification performance (24 classes) under four conditions: using the full input stream for a given trial (*none*), and using touch segmentation with only the left, right, or both IR sensors; we trained and tested only on finger-mounted IMU sensor sequence to control other variables. Accuracy with both IRs is 80.6% (*SD*=8.3), much higher than none (*M*=59.5, *SD*=11.0), left IR (*M*=59.3, 10.9), and right IR (*M*=59.3, *SD*=11.0). These three comparisons are statistically significant, confirming our choice to include both IRs: none ($t_{23} = -10.99$, $p < .001$, $d = -2.21$), left ($t_{23} = -11.02$, $p < .001$, $d = -2.24$), and right ($t_{23} = -11.18$, $p < .001$, $d = -2.23$).

6.5.4. Summary

We were able to achieve 94.9% recognition accuracy across 24 classes of location-specific gestures (8 gestures × 3 locations) at best with sensor fusion. When only single-sensor was used, the accuracy with a finger-mounted IMU was not significantly different from the camera alone. However, with multi-sensor combinations, the inclusion of the camera is necessary for higher accuracy for input localization, although IMU sensors might outperform the camera when it comes to recognizing gestures performed at fast speed.

6.6. Discussion

Our experiments demonstrate the feasibility of sensing location-specific on-body gestures using a multi-sensor finger-worn approach. Here, we reflect on the implications of our findings.

6.6.1. A Broader On-Body Input Vocabulary

Our work increases the input vocabulary of on-body interaction over existing approaches to sense on-body input, which support either input at only one location or only simple input at multiple locations. Our experiments, for example, we show how sensor fusion can be used to accurately discriminate three body locations and eight basic gestures—resulting in 24 distinct gestural inputs. In terms of information entropy, a number n of successive location-specific gestures from a set of possible input m is $\log_2 m^n$. As such, a single location-specific gesture from a set of 24 input vocabulary (8 gestures × 3 locations) can convey approximately 4.58 bits ($\log_2 24^1$), while a location-independent gesture from one of the eight gestures can only convey about 3 bits ($\log_2 8^1$) of information. A large input vocabulary should allow an expert user to complete a wide range of efficient and non-visual interaction tasks. Having flexibility in the types and locations of gestures that are supported should also enable new intuitive mappings between input and functionality—for example, a user could gesture on their wrist to check the time or set an alarm, or tap their ear to answer or dismiss a phone call. As well, our experiments suggest that the IMUs are particularly robust to differences in gesture articulation speed and, indeed, are able to distinguish between three different speeds with high accuracy (at least on the palm). Thus, our approach could also support speed-based gestures such as distinguishing between flicks versus slower swipes.

The exact input vocabulary for location-specific on-body interaction will, of course, depend on what locations can be accurately recognized. While further work is required to fully understand what locations at each of these levels of granularity will be robustly detectable, our results already suggest that our finger-worn multi-sensor approach should be able to support interactions that require multiple input areas within a small space, such as text or numeric entry (e.g., by tapping different locations on palm or fingers).

6.6.2. Comparing the Impact of Sensors

A key contribution of this work is not only to demonstrate the feasibility of using a finger-worn multi-sensor approach for location-specific input but also to characterize the benefits of the individual sensors—camera, IMU and IR. Moreover, while finger-worn cameras have been explored to some extent for input tasks, very little work has examined the combination of camera and other sensors. *Magic Finger* [134] is an exception, but as noted earlier, uses a camera for texture detection and an optical mouse sensor to detect optical flow, but does not combine the two input streams.

While the finger- and wrist-worn IMUs appear to contribute to localization, with a more realistic task where the user's body pose changes, the advantage of the camera for localization would likely only become more prominent for all of our experiments. When detecting not just touch location but also classifying more complex gestures, the addition of at least one finger-worn or wrist-worn IMU is also needed, with the combination of camera plus one of these IMU achieving over 90% accuracy over 24 classes. While adding more sensors (other IMU, IR) further improves location-specific gesture recognition accuracy by 2–3%, each additional sensor increases size, weight, power requirements and overall design complexity—all critical factors in wearable systems. The camera plus one IMU offers a minimal combination that may be sufficient for some applications. Future work should explore other types of sensors (e.g., electromyography or bio-acoustic) and combinations as well as algorithmic improvements for achieving high and robust accuracy. For example, one possibility is to apply other sensor fusion approaches, such as the HMM employed by *Botential* [74], or to deploy state-of-art camera-based sensing technologies such as a part-based model approach (e.g., [100]).

6.7. Limitations

Our experiments have several limitations that should be addressed in future work. Our dataset consists of input examples under controlled conditions, so classification accuracy with more realistic use will need to be assessed, including changes in body pose and use during movement (e.g., walking, riding in a vehicle). Moreover, our simple threshold-based segmentation approach made use of holistic knowledge of each trial and was applied only to individual trials with artificial tuning, which cannot be directly applied to a real-time system with continuous input streams. In addition, although the purpose of the study is to assess the feasibility of our finger-mounted sensor-based on-body input sensing approach, because all the participants were sighted and recruited from a campus, the study lacks representation of our target user group, namely people with visual impairments. Participants were also shown a live feed of the video camera, which was useful for ensuring the correct location of touches during the localized touch-and-hold input task for the system evaluation purpose. Without this feedback, however, users or the system will need to learn to accommodate the offset between the location visible to the camera and the user's perceived touch location. Finally, while we proposed and focused on a finger-mounted sensor-based approach as a solution to resolve challenges that users with visual impairments might face when using a camera (e.g., out-of-focus, out-of-frame), employing state-of-art techniques in computer vision for recognizing hand gestures (see survey [99]) would also be useful to consider to address the open issue with more realistic constraints.

6.8. Conclusion

We introduced and investigated a new finger-based, multi-sensor approach to detecting location-specific on-body gestures. Our findings not only highlight the feasibility of our approach (95% accuracy at detecting 24 location-specific gestures), but also characterize the utility of the camera, IR, and IMU sensors. The importance of the camera sensor for localizing input is critical, but to also achieve high gesture recognition accuracy requires the addition of at least a finger- or wrist-worn IMU. The expanded vocabulary size of on-body input afforded by location-specific gestures has the potential to support efficient interaction for expert users, intuitive task-based interactions, and relatively fine-grained input for body areas that have distinctive visual features (e.g., fingertips and palm). Having demonstrated the feasibility of a finger- and wrist-worn approach to sensing location-specific on-body gestures, the next chapter now turns to the question of how to employ such interaction to support blind users in completing microinteractions—that is, brief interactions that are typically seen as a strength of wearable computing devices (e.g., [11,70,112]).

Chapter 7: Supporting Accessible Microinteraction Through On-Body Interaction

7.1. Motivation and Introduction

Microinteractions are brief interactions that are designed to be completed within a very short period of time with minimal effort (e.g., less than four seconds) [10], and, as such, are particularly useful during multi-tasking (e.g., using the device on the go) [90,91]. Examples include checking the time, answering a phone call, or reading a new text message. Support for microinteractions is seen as a primary strength of wearable computing devices (e.g., [11,70,112,119]) and as such microinteractions are a promising application area for accessible on-body interaction. For users with visual impairments, in particular, designing efficient wearable support for microinteractions could be particularly desirable because interaction time for VI users is generally longer than that of sighted users [15,51] and incurs a greater cost when using the device in mobile context [2,111,135]. However, no studies have investigated how to support microinteractions for VI users beyond money identification or barcode scanning [78].

To first identify the needs, barriers and strategies for supporting microinteractions for people with visual impairments compared to sighted people, we conducted an online survey with 56 sighted and 61 VI screenreader users. We focused on uncovering the usages of current hand-held smartphones, and soliciting perceived advantages and limitations of a smartwatch and of on-body input to learn design implications for supporting wearable, accessible microinteractions. Following this survey, we then conducted interview and design probe sessions with 12 VI screenreader users more specifically focused on supporting microinteractions through on-body input. The design probes included three different real-time on-body interface implementations that allowed us to explore user responses to location-independent gestures versus location-specific gestures. In Chapters 3 and 4, we had predicted that VI users would prefer location-specific gestures, but those chapters did not include use and comparison of location-specific and location-independent on-body interaction implementations. The design probes were implemented using the real-time version of the system we developed in Chapter 6. The main questions we seek to answer are:

- *What are the most useful tasks to be supported as microinteractions for users with visual impairments, and do these differ from microinteractions for sighted users?*
- *How should on-body interaction be designed to support microinteractions for people with visual impairments?*

7.2. Experimental Methodology

This section describes both the survey and in-person study session method.

7.2.1. Online Survey

We designed and conducted an online survey to collect responses from both sighted and people with visual impairments in terms of mobile and wearable device use, primarily focusing on existing use and perceptions of smartphones and smartwatches, but also soliciting open-ended feedback on the idea of on-body interaction. The main questions here include: (1) What are the most common current microinteractions used by VI and sighted users? (2) What tasks, if any, are not supported as microinteractions

Task Label	Task Description
Alarm	Set an alarm or timer
App Launch	Find and open a specific app
Calendar	Check your calendar for an overview of the day's schedule
Clock	Check the current time
Music	Pause a music player
Navigation	Set a destination to get navigation directions
Phone	Dial a phone number
Read Msg.	Read a text message that is two sentences long
Respond Msg.	Respond to a text message with a two-word reply
Weather	Check the weather

Table 7.1. A list of ten specific microinteractions examined in online survey. The tasks are alphabetically sorted by task label in asending order.

for VI users but would be valuable to support as such? (3) To what extent do VI users versus sighted users perceive wearable devices (specifically, a smartwatch) to enable microinteractions compared to the smartphone? (4) How do people react to the idea of on-body interaction and what are the foreseen potential use cases?

7.2.1.1. Participant Recruitment and Survey Platform

We recruited smartphone owners through email lists, university bulletin boards, organizations working with people with visual impairments, social networking sites, and word of mouth. The survey was hosted on SurveyMonkey and was designed to take up to 20 to 25 minutes for screenreader users. Participants could opt into a draw for a $100 Amazon gift certificate after the completion of the survey.

7.2.1.2. Survey Outline

The survey consisted of 32-36 questions depending on the participants' level of vision and an additional 24 questions for smartwatch owners. Questions included general background (e.g., age, gender, level of vision), current mobile and wearable technology use if applicable (e.g., device type, frequency of use), perceived tradeoffs between smartphone and smartwatches, and estimated time needed for each of a set of microinteractions in Table 7.1 that are common with mobile or wearable devices [9,94].

At the end, we also asked participants' opinions of on-body interaction and potential use cases for such interaction. In case participants were not familiar with smartwatches or on-body interaction, we provided brief descriptions:

> Smartwatch description: A smartwatch offers many of the same features as a smartphone, but has a smaller screen the size of a large watch face. Smartwatches often include a camera, speaker, microphone, and sensors that can track information like the number of steps you've taken.
>
> On-body interaction description: Imagine a small wearable device such as a wristband or a ring that is able to sense when you do taps, swipes or other gestures on the surface of your body. This device would be paired with a small speaker for audio output or with a projected image on your arm or hand and would allow you to do many of the same actions as you can do on a smartphone or smartwatch. For example, you could do taps and swipes on your bare palm in the same way you usually use the touchscreen on a phone or smartwatch. You could also make tap or make other gestures at specific locations on your body, such as the wrist, a fingertip, or ear to do a specific action (e.g., check the time, answer a phone call, change a song's volume).

7.2.1.3. Data and Analysis

A total of 147 responses were collected during a one-month period from August 17th to September 17th, 2016. The dropout rate was 8.8%, leaving 134 completed surveys from 56 participants who had normal or corrected-to-normal vision and 78 participants who reported visual impairments (our analysis focuses on the 61 of these participants who used screenreaders). The median completion time was 10.5 minutes for sighted participants, and 22.0 minutes for participants with visual impairments. We followed an iterative coding process [47] for seven open-ended responses such as identifying the types of tasks that are frequently completed within a short period of time (namely ten seconds) and the perceived tradeoffs of smartphones versus smartwatches. Two researchers developed initial codebooks for each question. Three to four iterations were completed per question, where each researcher independently coded a randomly selected subset of 20-40 responses, Cohen's kappa was computed to assess interrater reliability for each code, and codes were refined. After the final iteration, the average kappa score across all codes was 0.87 (SD = 0.12, range 0.53 to 1.0). The worst performing code was "*Other/general*" for the general reactions to the on-body interaction scenario.

In our analysis below, we summarize data from sighted participants separately from that of participants with visual impairments. We further exclude 17 VI participants who reported not using a screenreader software on their phone for "*most of the time*" or *"always"*, as we are primarily interested in screen reader users. We ran Chi-square tests of association using contingency table analysis for close-ended questions to assess whether there is a significant relationship between user group and specific tasks on their smartphones.

7.2.1.4. Participants Demographics

As mentioned above, of the 134 completed responses, 78 participants reported having a visual impairment, while 56 did not (**sighted** participants). Of the VI participants, 61 (**$\mathbf{VI_{audio}}$** participants) reported using a screenreader on their phone "*most of the time*" or "*always*": one had vision from 20/70 to 20/200, 13 were legally blind at best 20/200, 17 were blind with some light perception, and 30 were totally blind. The remaining 17 VI participants (**$\mathbf{VI_{visual}}$** participants) relied more on visual feedback from the screen, for example, using a screen magnifier: eight had vision from 20/70 to 20/200 and nine were legally blind at best 20/200. While the median age of VI participants was 45-54 for both VI user groups, that of sighted participants was 25-34. The VI_{audio} user group included 38 female participants (22 male, and 1 *Other*), and 9 of the VI_{visual} participants

ID	Age	Gender	Visual impairment	Screen reader	Screen magnifier	Voice input
P1	50	Male	Blind	Always	N/A	Some of the time
P2	35	Female	Light perception only	Always	N/A	Some of the time
P3	41	Male	Low vison	Always	N/A	Rarely
P4	61	Female	Blind	Always	N/A	Always
P5	38	Male	Low vision one eye; none in other	Always	Some of the time	Some of the time
P6	56	Male	Light perception only	Always	N/A	Most of the time
P7	65	Male	Blind	Always	N/A	Most of the time
P8	41	Female	Light perception only	Always	N/A	Always
P9	56	Female	Blind	Always	N/A	Most of the time
P10	51	Female	Blind	Always	N/A	Rarely
P11	29	Female	Low vison	Most of the time	Some of the time	Most of the time
P12	31	Female	Light perception only one eye; none in other	Always	N/A	Most of the time

Table 7.2. Demographics of participants for the in-person study.

were female (7 male, and 1 *Other*). For the sighted user group, 30 were female (26 male).

7.2.2. Interview and Design Probe Study

We designed a two-hour single-session study consisting of semi-structured interview questions followed by use of three on-body interfaces as design probes. Compared to the survey, these sessions allowed us to collect in-depth qualitative data and responses based on use of a real-time system.

7.2.2.1. Participants

We recruited a total of 12 participants (7 female) through local organizations working with people with visual impairments. As shown in Table 7.2, all participants were blind or had low vision, and reported using a screenreader all the time, except for P11 who used a screenreader "*most of the time*". The average age was 46.2 (SD = 12.0, range 29 – 65) and all participants had owned a smartphone for more than a year. One participant (P10) reported having a fitness tracker. P5 had smartwatch for a year, but reported that

Main Menu	Submenu	Description
Clock	Time	Check the current time
	Alarm*	Check the next alarm
	Timer*	Check the time remaining
	Stopwatch	Check the time elapsed
Daily Summary	Date	Check today's date
	Calendar*	Check the next event
	Weather*	Check the current temperature
Notifications	Summary	Check the notification summary (e.g., "1 miss phone call and 2 new messages")
	Missed phone call*	Check missed phone call (e.g., "A missed phone call from Alice")
	New message #1	Check new message (e.g., "A new messge from Bob")
	New message #2*	Check new message (e.g.,"A new message from Charlie")
Health and Activities	Distance	Check the miles traveled
	Steps*	Check the number of steps taken
	Calories	Check the calories burnt
	Heart rate*	Check the heart rate
Voice Input*	None	Activate voice input

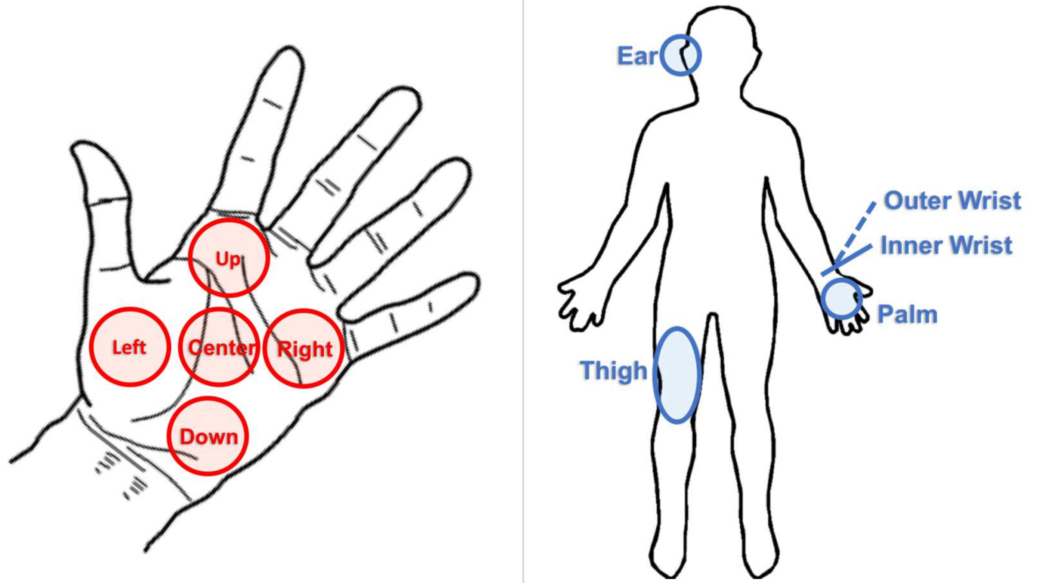
Table 7.3. We used two-level highrarchical menu for the study. There are five main menu items, and each of the main menu contained 3-4 submenu items except for "*Voice Input*", which has no submenu item. A total of 16 microinteractions were supported by our system. The mark '*' indicates that the submenu was used for our task.

he barely used it. The remaining participants reported not having any wearable device. Participants were compensated $30 for their time and additional $30 for the travel costs.

7.2.2.2. Design Probe Interfaces

To learn design implications for supporting microinteractions through on-body input, we designed three different on-body interfaces for menu navigation for people with visual impairments: location-independent gestures (LI_{any}), location-specific gestures on palm (LS_{palm}), and location-specific gestures at parts of the body (LS_{body}). In addition to sparking discussion on using on-body input to support accessible microinteractions, these three implementations allowed us to evaluate with a real-time system the prediction from Chapters 3 and 4 that users would prefer location-specific gestures to location-independent gestures for on-body input.

All three interfaces versions supported 16 microinteractions with two-level hierarchy menu as shown in Table 7.3. The main items in the top level were: "*Clock*",



(a) Five locations on the left palm for LS_{palm}: *up* – Notifications, *left* – Clock, *right* – Daily Summary, *down* – Health and Activities, and *center* – Voice Input.

(b) Five locations across body for LS_{body}: *palm* – Notifications, *outer wrist* – Clock, *inner wrist* – Daily Summary, *thigh* – Health and Activities, and *ear* – Voice Input.

Figure 7.1. The mapping between location and pre-defined microinteractions used in this study: (a) completely arbitrary locations on palm used for LS_{palm}, and (b) semi-contexual locations across different body parts for LS_{body} for right-handed participants. For left-handed participants, the other side of the body was used.

"*Daily Summary*", "*Notifications*", "*Health and Activities*", and "*Voice Input*". Each of the main menu contained submenu items except for *"Voice Input"*. The interfaces were designed as follows:

LI_{any} simulated the gestures used in VoiceOver or TalkBack, which allows users swipe left and right on any surface of body to navigate back and forth through a list of menu items, which would be time-consuming to navigate a long list of menu items [50].

LS_{palm} included five specific locations on the palm that mapped to the top-level items as shown in Figure 7.1a; the top-level menu items were arbitrarily mapped to the five locations. Users could either point directly to a location to select that item (similar to the interaction evaluated in Chapter 5), or could *"explore by touch"*, as supported by TalkBack and VoiceOver, where the user searches for an item by sliding their finger across the surface of their palm, hearing each item read aloud.

LS_{body} is also location-specific, but unlike *LS_{palm}*, it maps each menu item to a location across the full body and the mappings are more contextual rather than arbitrary; for example, as shown in Figure 7.1b, tapping the wrist checks the time.

For all three interfaces, double-tapping activated the currently selected item for (again following interaction in VoiceOver and TalkBack), and to access the submenu items, users would first need to find and select the top-level menu item. Once the top-level menu item is selected, navigating the submenu items were the same for all three interfaces.

7.2.2.3. Apparatus

We implemented these three interface designs on top of a real-time system based on the sensing hardware and algorithms described in Chapter 6. The wearable hardware here included a camera[11], an IMU[12], and a pair of IR reflectance sensors[13] mounted on the user's index finger via a 3D printed ring and Velcro strips. The IMU and IR sensors were connected to a wristband containing an Arduino microcontroller[14], which was in turn connected (along with the camera) to a desktop computer[15]. Speech feedback was provided through a pair of speakers using the Microsoft .Net speech synthesis libraries. This prototype was intended primarily to allow us to explore possible on-body

---

[11] Awaiba NanEye GS Idule Demo Kit

[12] Adafruit Flora LSM9DS0

[13] Fairchild Semiconductor QRE113GR

[14] Sparkfun Arduino Pro Micro (5V/16MHz)

[15] Dell Precision Workstation T7910 (CPU: Intel Xeon, 8-core, 2.1Ghz, GPU: NVIDIA GeForce GTX 750Ti)

interaction designs, and we envision a future system that is much smaller and self-contained.

7.2.2.4. Procedure

The procedure began with a background questionnaire, followed by a semi-structured interview (~20 minutes). The interview covered the same themes as the survey with additional questions related their use of devices with different input methods. After the interview and a short break, we transitioned to design probe, which consisted of using the three interfaces described above. First, we provided a short overview of the system as a whole, then started collecting images from nine different locations on their body (see Figure 7.1) to adaptively train the system. Following a short introduction and guided practice with each of the interface designs, we asked participants to complete a set of basic tasks (~5 minutes). The order of the three interfaces was fully counterbalanced. All interfaces had five applications in the main menu including: *Clock*, *Daily Summary*, *Notifications*, *Health and Activities*, and *Voice Input* that are commonly used on a smartwatch [94], and each of these had 3-4 sub-menu items except for Voice Input. Participants were asked to find and open one of the sub-menu items first by finding and opening the main menu item with each of the given interface, then by navigate between sub-menus by location-independent swiping.

For each interface, participants completed two tasks from each of the five main menu, for a total of 10 tasks; these were randomly ordered. Examples of tasks included finding and opening *Alarm* under *Clock*, *Weather* under *Daily Summary*, and *Steps* under *Health and Activities*. After a short description of each task, each task was completed after the text-to-speech signal saying “Begin”. After using each of the three interfaces participants rated the following (in 7-point scale) and provided open-ended comments on their experience:

- *Efficiency – How efficient do you think it is to use?*
- *Easiness – How easy do you think it is to use?*
- *Use in public – How comfortable would you feel using this version in public?*

Finally, the session concluded with a semi-structured interview (~15 minutes) on subjective feedback comparing the three conditions.

7.2.2.5. Data and Analysis

Open-ended responses were audio recorded and transcribed for analysis. For questions asking about specific tasks (e.g., "*What tasks do you think it would be useful to do on a smartwatch, if any?*"), we used the same codebook developed already for the online survey. Other questions were analyzed by one researcher based on themes of interest [21] (e.g., one-handed use of the phone, interaction on-the-go), while allowing for new, emergent themes. For the subjective ratings of the three on-body interfaces, we used Friedman tests because the normality assumption of parametric tests may not hold for the 7-point rating scale data we collected. Holm's sequential Bonferroni adjustments were used to protect against Type I error [45] for all posthoc pairwise comparisons.

7.3. Results

Here I present the findings from the online survey and the in-person sessions. As for the online survey, the responses are summarized to understand the differences in smartphone and smartwatch usage by the two user groups of primary interest: sighted users and VI_{audio} users (*i.e.*, screenreader users), together with the findings from the interview. For the design probe study, I summarize the perceived trade-offs between three interfaces we examined based on qualitative analysis.

7.3.1. Mobile and Wearable Devices: Ownership and Frequency of Use

As required, all participants had smartphones. While the mobile operating systems for their smartphone devices was evenly split between iOS and Android for sighted participants (29 iOS, 30 Android, 3 Windows, and 2 *Other*), there were more iPhone than Android phone owners in the VI_{audio} user group (59 iOS, 8 Android, and 1 *Other*) for online survey participants, and 11 iOS and one Android phone user (P5) for the interview participants, confirming previous studies including ours [13,135,137]; note that some participants reported having multiple devices. In general, participants tended to use their smartphone and smartwatch (if they owned one) frequently. A majority of the online survey participants overall reported using their phone at least once an hour, but broken down by group, this was true for 65.6% of the VI_{audio} group and for 75.0% of the sighted group. As for the participants from the interview, all but two participants (P4, P10) reported the same. A total of 13 online survey participants (9 sighted, 4 VI_{audio}) reported owning a smartwatch, and one interview participant (P5) reported that he once had a smartwatch for a year, but barely used it.

7.3.2. Identification of Useful Microinteractions

To identify mobile tasks that are most valuable to be supported for VI users as microinteractions compared to sighted users, we examined how the device use of VI participants differs from that of sighted users.

7.3.2.1. Frequently Used Microinteractions on Smartphones

We asked what tasks participants frequently do on their phone in about 10 seconds or less (open-ended)—these are tasks that would be particularly important to support as microinteractions. The most frequently reported tasks were related to immediate updates such as checking notifications, time or weather across both user groups: managing email (76.9% in general, where 82.2% of this response was specifically for checking for new emails), managing messages (72.6% in general, where its 62.4% was checking for new messages), checking weather (31.6%), time (29.9%) and social media updates (24.8%). But Chi-square tests revealed that the ratio of participants was significantly different between two user groups for managing phone calls such as answering or making a call ($\chi^2_{(1)} = 9.25$, $p = .002$, $\phi = .28$), and checking the weather ($\chi^2_{(1)} = 5.16$, $p = .023$, $\phi = .21$). While the percentage of VI_{audio} participants who reported the frequent use of phone calls is 34.4%, it was only 10.7% for sighted group. Also, a higher proportion of VI_{audio} participants (41.0%) check the weather more frequently than sighted participants (21.4%). The responses from the interview participants were similar to survey responses: managing phone calls (N=9 of 12), managing text messages (N=7), and checking the weather (N=4). But looking up information using voice search was also dominant (9 responses each). To sum up, tasks related to notifications, communications (namely email, text messages, and phone calls), time, weather, and voice input are the top five most frequent microinteractions that VI participants are currently using on their phone, which should be supported through on-body interaction as well.

7.3.2.2. Useful Tasks on Different Devices

To understand what tasks would be useful to be supported as microinteractions through wearable devices (e.g., on-body interaction), we examined what types of tasks that people would like to do on a smartphone versus a smartwatch. We asked participants (Q1) what tasks they think are especially useful on a smartwatch, and (Q2) what tasks they would prefer to do on their phone instead of on a smartwatch.

As found in [94], although their responses were collected from sighted participants only, the top five responses across both user groups for the first question were checking the time, health and fitness tracking, managing text messages (mostly for checking for new messages), GPS navigation, and notifications. As for the second question, the top five responses across both user groups also included managing emails, and text messages (but more for reading and responding rather than just checking new messages), phone calls, entertainment (e.g., watching Neflix, playing games), and text entry. No significant differences were found between the two user groups for either question.

However, a greater percentage of VI_{audio} participants (27.9%) showed a strong preference for using their phone for tasks in general than sighted participants (12.5%), responding that they would prefer to do everything on their phone rather than on a smartwatch. This relationship between strong preference for the phone and user group was found to be significant ($\chi^2_{(1)} = 4.23$, $p = .040$, $\phi = .19$). Confirming the sentiment from VI survey respondents, one third of interview participants (N=4 of 12) also reported that they would like to do everything on their phone as they consider a smartwatch as a redundant device; P6 specified, *"Honestly, I looked into it* *[smartwatch]. Honestly, I haven't seen anything where the smartwatch is going to help me justify purchasing one. Everything that the smartwatch does I can do with my iPhone, including counting steps. And I did that too by the way. It's another app. I use everything, I really do. Since the Smartwatch requires that you still have to have your iPhone? I see no advantage. If and when you don't have that requirement I could see the smartwatch becoming potentially helpful".* Unlike a prior study that showed positive attitude of VI users towards wearable devices in general [135], this finding suggests that VI users are less interested in using a smartwatch than sighted users. More specific perceptions of smartwatches compared to smartphones are examined in Section 7.3.5.

At the very end of the survey, we also asked what tasks would be useful to support through on-body interaction. The top five responses across two user groups were the tasks that do not necessarily require visual feedback, with the exception of managing text messages: phone calls (34.0%), checking the time (29.2%), device settings such as adjusting the volume (16.0%), controlling a media player (19.8%) and managing text messages (18.9%; mostly for reading). The differences in the response ratio between the two user groups were not significantly different for all coded responses with Chi-square tests, except for checking notifications (9.8%) and composing text messages (7.8%); no VI users listed these microinteractions. In terms of general comments, a greater percentage of VI_{audio} participants responded that they would like to do anything that they can do on their phone via on-body interaction (10.9%, N=55) whereas only 3.9% for sighted participants expressed a similar sentiment (N=51); although the differences in proportion was not found to be significant by the user group. Note that the question was optional that not every participant responded.

7.3.2.3. Use of Voice Input

Motivated by findings from previous studies that show voice input is more popular for VI_{audio} users than sighted users [13,135], we examined how speech input is being used for microinteractions by these two user groups. Confirming the prior findings, while 36.1% of the VI_{audio} participants reported that they use voice input (e.g., Apple Siri, Google Now) "*most of the time*" or "*always*", the same responses were reported only 8.9% from sighted participants. A Chi-square test found a significant association between user group and reported use of voice input ($\chi^2_{(4)} = 34.55$, $p < .001$, $\phi = .54$).

As for the participants from the interview, as shown in Table 7.3, seven participants (of 12 total) reported using voice input either *"most of the time"* or *"always"*, while three participants reported using voice input *"some of the time"*. The remaining two reported that they rarely use voice input. Four participants reported that they would use voice input for everything if possible, and five participants said that they liked using voice input because it is faster than manual input, confirming the finding in [13].

We further asked participants to specify for which if any of the ten microinteractions (see Table 7.1) they use voice input, and the top five tasks are shown in Table 7.4. In general, among the participants who reported using voice input, more percentage of VI_{audio} participants use voice input for various tasks than sighted users, and Chi-square tests found that the associations between user group and reported use of voice input were significant for "*Phone Calls*" ($\chi^2_{(1)} = 11.41$, $p = .001$, $\phi = .35$), and

Task Label: Description	Sighted	VI_{audio}
Alarm: Set an alarm or a timer	52.9%	75.9%
Navigation: Set a destination to get navigation directions	52.9%	56.9%
Phone Calls: Dial a phone number	38.2%	79.3%
Respond to a Message: Respond to a text message with a two-word reply	38.2%	70.7%
Weather: Check the weather	32.4%	56.9%

Table 7.4. The percentanges of top five tasks that participants use voice input for within each group ($N_{sighted}$ = 34, $N_{VIaudio}$ = 58) sorted by the percentages of responses for sighted participants in decending order, which shows the general trend that more percentages of VI_{audio} participants use voince input than sighted participants.

"Respond to a Message" ($\chi^2_{(1)} = 13.93, p < .001, \phi = .39$), which conflicts prior findings [13] where the number of participants who use voice input for text messages were almost the same across two user groups. Similarly, the most frequent tasks the interview participants do with voice input were text entry (N=8 of 12), making phone calls and voice search (reported by N=7 for each).

7.3.3. Barriers for Enabling Accessible Microinteractions

Tasks on a touchscreen device require longer interaction time for users with visual impairments compared to sighted users in general [15,51]. As such, to identify tasks that would be particularly more valuable to be supported for VI users as microinteractions, we asked participants to report on the durations for device access time (from device retrieval to replacement), and device usage time (e.g., completion time for ten specific mobile tasks such as dialing a phone number) and assessed the time difference between sighted and VI_{audio} participants. Furthermore, we examined situational barriers (e.g., interaction on-the-go, one-handed use) that VI users often encounter as these would prevent efficient interactions.

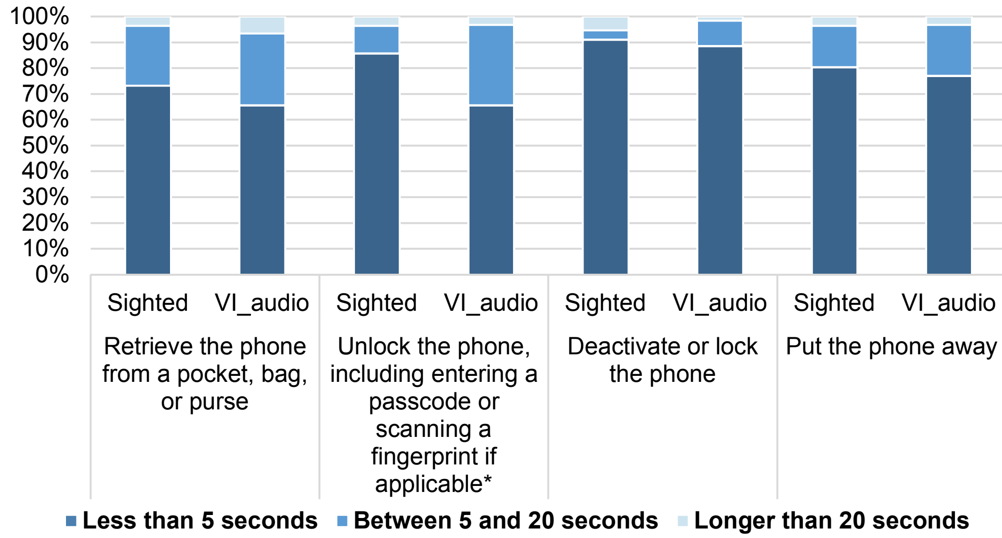


Figure 7.2. Percentages of reported duration breakdowns for each step from device retrieval to replacement within each group ($N_{sighted}$ = 56, and $N_{VIaudio}$ = 61). The mark '*' indicates that the association between the user group and the duration for the step was significant with a Chi-square test using a 2 × 3 contingency table. VIaudio participants reported slower compeletion time for unlocking their phone than sighted participants.

7.3.3.1. Duration from Device Retrieval to Replacement

We first examined whether the time taken from retrieving the device (e.g., from a pocket or a purse) to putting it way (device replacement) is different across two user groups. The duration was broken down into multiple steps as in Ashbrook et al. [11], and the reported duration of each step that are shown in Figure 7.2. Overall, sighted participants reported being faster than VI_{audio} participants for each step in accessing their phones. Chi-square tests revealed that there is a statistically significant association between user group and reported durations for *"Unlocking the Phone"* ($\chi^2_{(2)} = 7.29$, $p = .026$, $\phi = .25$); the associations for other steps were not significant.

While the percentages of participants who reported that unlocking the phone takes less than five seconds was 85.7% for sighted user group, it was only 65.6% for VI_{audio} group. Instead, more VI_{audio} participants responded for *"Between 5 and 20 seconds"*, at 31.1% compared to sighted participants at 10.7%. This result confirms prior finding from Azenkot et al. [14] that entering a 4-digit passcode by blind participants took 7.52 seconds. Further, we asked participants whether they had set up a passcode for their phone as it would impact the time taken for unlocking the phone. The percentage of participants who said "*yes*" to this question was 80.4% for sighted participants, and 68.9% for VI_{audio} user groups. A Chi-square test revealed that these ratios were also not significantly different across the two user groups, showing that this impact can be ignored when comparing device unlocking time.

Similarly, eight out of 12 participants (75%) from the in-person study reported that they have set up a passcode. Six of the eight participants reported that they prefer using the fingerprint scanner not only because it is easier (N=4) and faster (N=2) than entering a passcode, but also because it prevents from others hearing or seeing their passcodes (5 and 1 responses, respectively), confirming the VI users' concern of entering a passcode while using a screenreader in [14]. Meanwhile, the remaining participants have not set up a passcode because entering a passcode is slow (N=2), or they did not care to set it up (N=2). The results from our data confirms the findings in prior studies [14,51], and suggests that supporting authentication as a microinteraction that prevents eavesdropping would be useful for VI users.

7.3.3.2. Device Usage Time for Microinteractions

We also asked participants close-ended questions on how long each of the ten microinteractions takes on their phone (see Figure 7.3). Similar to the results above, sighted participants reported being faster than VI_{audio} participants for all types of microinteractions. We conducted Chi-square tests, and the relations between user group

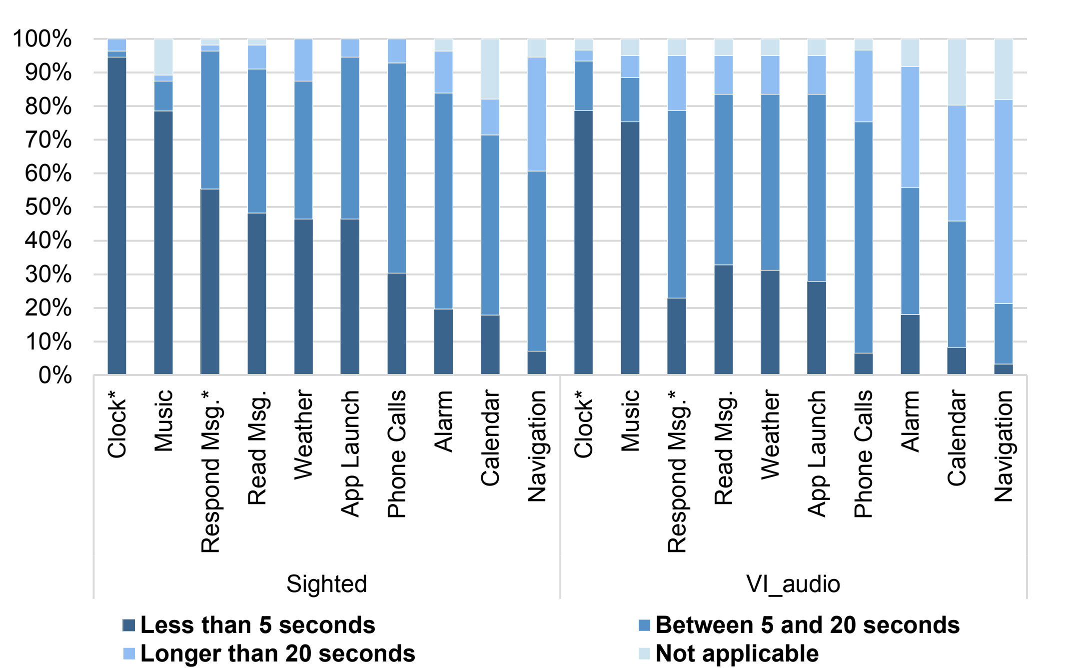


Figure 7.3. Percentages of reported duration breakdowns on smartphones within each group for both VI and sighted participants, sorted by the percentage of the sighted participants for "Less than 5 seconds" in decending order ($N_{sighted}$ = 56, $N_{VIaudio}$ = 61). This graph illustrates that sighted users tended to provide shorter task times than VI_{audio} users across the 10 microinteractions. The mark '*' indicates that the association between the user group and the duration for the task was significant with a Chi-square test using a 2 × 4 contingency table.

Task Label: Description	$\chi^2_{(2)}$	p	ϕ	$N_{sighted}$	$N_{VIaudio}$
Clock: Check the current time	6.57	.037	.24	56	59
Calendar: Check your calendar for an overview of the day's schedule	10.84	.004	.34	39	56
Navigation: Set a destination to get navigation directions	15.14	.001	.38	46	57
Respond Msg.: Respond to a text message with a two-word reply	15.84	< .001	.37	55	58

Table 7.5. For all ten microinteractions, the percentages of VI_{audio} participants who reported for "less than 5 seconds" was less than that of sighted participants. These are four tasks whose associations of user group and reported time breakdown for each task were found to be significant by Chi-square test results with 2 × 3 contingency tables.

and the reported time breakdowns were found to be significant for the following four tasks: "*Clock*", "*Calendar*", "*Navigation*", and "*Respond Msg.*" (see Table 7.5 for details). Regardless of the user group, a majority of participants reported that they can complete "*check the current time*" within five seconds while more sighted participants reported faster duration than VI_{audio} participants; 94.6% and 78.7% respectively. In addition, tasks that require long interaction time even for sighted participants (namely *"Calendar"*, and *"Navigation"*) also showed significant relation between user group and completion time.

Most of all, the greatest gap was found for "*Respond Msg.*", where the percentage of sighted participants who reported the completion time to be less than five seconds was 55.4%, which was almost twice more than VI_{audio} group (23.0%). This confirms prior findings that text entry is one of the time-consuming tasks on a touchscreen-based interaction for users with visual impairments (e.g., [19,31,87]). Overall, the results suggest the duration gap is greater between sighted and VI_{audio} participants for the tasks that are either extremely simple or complex, or the task that involves text entry.

7.3.3.3. Situational Barriers

As Abdrolrahmani et al. [2] found that using a mobile device while walking on with a cane is a major concern for VI users, we asked how participants use their phone while walking during the interview, as an opportunity to solicit design considerations for supporting accessible on-body interaction for VI users. Confirming the prior findings in [2], eight participants (of 12 total) said that that they do not use the phone while walking due to safety. For example, P8 said*, "See, when I'm walking, I have to use my cane, so I don't even be on the phone. To me, that's almost like driving and being on the phone. The fact [that I] don't have any sight, just lightness, and that's not enough to tell me what's in front... […] I gotta use my cane and use my sounds at all times. So, I gotta hear, listening to stuff, and feeling with my cane, so I really don't have time to mess with the phone walking. To me, that'll put a blind person in danger. […] Safety* *comes first, and the phone is not that important, when it's me walking by myself"*. Half of the participants mentioned that they do not take their phone out while walking (where they can still use their phone with a headphone for answering a call or using voice command), and four of them specified that they are concerned of their phone being stolen. P10 said, *"Because I think it's a little bit of fear that somebody will come out and grab my phone and run away. [The phone is] Usually in my pocketbook. And now, I have something that I could put it on my neck. But, I usually don't do it. I just want to keep it safe, it [holding the phone in public] is kind of an advertisement that you have a cellphone"*.

In terms of one-handed interaction, while nine of twelve participants said that it is very important to be able to use one hand while walking on a street, only four of them reported that they were comfortable using the phone with one hand; others stated that using the phone with one hand is physically challenging (4 participants), or it increases the chance of dropping the phone (4 participants), thus they typically secure the phone with one hand and use the other hand for gestures, or do not use the phone at all when both hands are not available as mentioned above. This again confirms the findings with the issue of one-handed interaction on-the-go for VI users when interacting with hand-held mobile devices [2,135], including the findings in Chapter 4.

7.3.4. Strategies for Quick Access to Apps

To examine whether and how VI users employed app layout customization to support quick interactions, we asked interview participants if they had rearranged their app icons on their phone to find and open some apps more quickly than others. Seven of 12 participants responded *"yes*". They had rearranged the icons to specific pages, and four of them also placed their icons in folders. P6, for example, had three pages where he would place commonly used apps on the first page (home screen) without folders, his second page contains multiple folders (e.g., blindness-related, news, travel), and the last page has all the apps he has installed but not yet tried to be assigned to one of the folders in the second page.

Unlike these participants, the other five participants had not customized the location of app icons. Still, they had general idea where their icons are located. P8 said *"I've never done that [rearranging the icons]. I just figured that once they're there, they're there. […] Like I have four pages on my phone, no, six pages. So I know what basically a lot of which page some of that stuff is on. Like, my metro access, easy pay, my metro access reservation, that's on page four. My bible gateway, where I read my bible, that's on page two. Things like that. So, just working with it more and more, you remember where a lot of it is, on which page"*. This finding suggests that VI users get benefits of the spatial layout of the interface regardless of whether they customized it or not. The findings indicate the VI users' interests in customizing their interface as well as their use of spatial mapping for quick access.

7.3.5. Perceived Smartwatch Advantages and Limitations

We further asked about the perceived advantages and limitations of smartwatches to learn design implications from a mainstream wearable device for supporting on-body interaction, and the results confirmed findings in [139]. As shown in Table 7.6, besides "*Other*", the top three advantages provided were *"Quick/easy access"*, *"Small/lightweight",* and "*Portability"* for both user groups. As for the interview responses, six out of 12 participants liked that smartwatches provide quick and easy

		Sighted	VI_{audio}
Advantages	Quick/easy access	58.0%	57.1%
	Portability	16.0%	24.5%
	Small/lightweight	14.0%	34.7%
	Hands-free	10.0%	10.2%
	Less distraction	10.0%	0.0%
	Multi-tasking	8.0%	0.0%
	Discreetness	6.0%	4.1%
	Other	16.0%	20.4%
Limitations	Small size	75.5%	48.3%
	Low processing power	17.0%	18.3%
	Not a stand-alone device	11.3%	13.3%
	Short battery life	5.7%	6.7%
	Expensive cost	3.8%	6.7%
	Obtrusiveness	3.8%	3.3%
	Display not adjustable	3.8%	1.7%
	Sound-related issues	0.0%	15.0%
	Other	9.4%	5.0%

Table 7.6. Perceived advantages and limitations of smartwtches. The percentages of participants who responded to each question within each group, sorted by the popularity of responses for the sighted group from most to least popular (Advantages: $N_{sighted} = 50$, and $N_{VIaudio} = 49$; Limitations: $N_{sighted} = 53$, and $N_{VIaudio} = 53$). While the responeses were versy similar acorss uwer groips, participants from only VIaudio commented on sound-related issues.

access without pulling out the phone because it is always worn on their wrist. For example, P1 said, *"Since they [smartwatches] are wearable, they're just on me. The information is also in my hand, instead of my pocket or elsewhere on the table so that makes it more and more readily accessible"*. The next frequently reported advantage was that a watch can free their hands (reported by five participants), especially while walking. P12 mentioned that, *"'Cause a lot of times, one hand is my cane and then one hand is my daughter, so hands-free would be.. [showing two thumbs up]"*. Next prominent advantage was its portability; four participants said that the watch is easy to carry with less chance of losing or dropping. Three participants reported aesthetic part of the watch as one of the advantages.

In terms of limitations, the most prominent responses for both groups were *"Small size"* followed by *"Low processing power"*, which also included limited capacity or storage. However, while the third most-frequent response was *"Sound-*

related issues" (e.g., low volume, not being able to pair with/plug in a headphone) for VI_{audio}, none of the sighted participant mentioned this as a limitation of a smartwatch. Similarly, small size of the touchscreen was a concern for most interview participants (*N*=8); five of them specified that small screen makes input difficult (e.g., text entry, gestures for menu navigation). P1 responded *"because it [phone] has wider space so I can find the lay out of the keyboard, and I'll have a wider area of navigation between the letters and the keys. But in the case of smart watch, it's much, much smaller, so it will be compact and difficult to look at each letter."* Sound-related issues were again mentioned by three participants from the interview—they were concerned about the low volume of the sound. P10 specified *"The speaker on the phone. Like for iPhone it's big, the phone is big, so you can have the speaker is big. And for a smart watch, how big will be the speakers? It can speak maybe a tiny voice?[…] Because it will be, for somebody who is hard of hearing, that if it's a noisy place, how do you hear what your phone is saying, or what your watch is saying? 'Cause I think that will be my challenge, too"*.

7.3.6. Trade-Offs Among Design Probes

To learn design implications for supporting microinteractions through on-body interaction for VI users, the in-person sessions compared three design probes based on qualitative responses. We collected subjective ratings for *efficiency*, *easiness*, and *use in public* for the 12 participants and the results are shown in Figures 7.4 and 7.5. Separate Friedman tests to assess the impact of the different interfaces on the efficiency, easiness, and use in public ratings were not significant (*i.e.*, the data in Figure 7.4), but vote counts in Figure 7.5 suggests an affinity for the LI_{any} and LS_{palm} interfaces

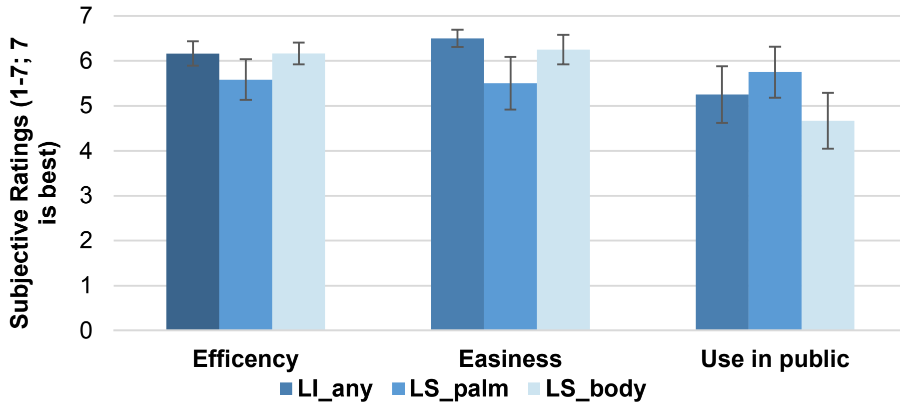


Figure 7.4. Average subjective ratings for three design probes in terms of efficiency, easiness and use in public. The gestures were performed anywehre on body (LI_{any}), five locations on only on palm (LS_{palm}), and five locations across the whole body (LS_{body}). Error bars indiciate standard errors ($N = 12$).

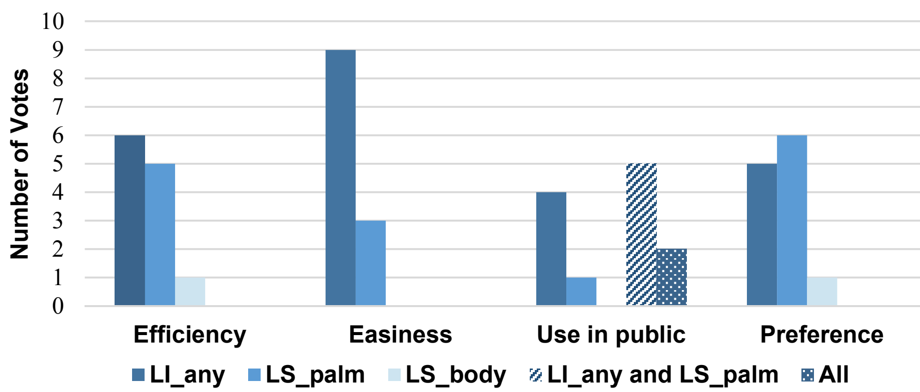


Figure 7.5. The vote counts for the designs that are the most efficient, most easy, and most comfortable to use in public as well as most preferred ($N = 12$).

compared to the LS_{body} interface. Participants' open-ended responses also provide insight into tradeoffs between the three design probes.

Overall, all participants like the idea of interacting with different locations on their body for different tasks. However, as shown in Figure 7.5, the LI_{any} interface, which deployed location-independent gestures, had the highest vote counts for efficiency and easiness, mainly because it did not require participants to learn the mapping between the locations and applications. At the same time, seven participants considered LI_{any} to be slower than LS_{palm} as they have to swipe and listen for each of the selected items. While the overall preference was almost equally split between LI_{any}

and LS_{palm}. All participants except P8 were concerned with having to interact all over the body instead of at a single location with LS_{body}, and as a result were worried about socially acceptability, as reflected in Figure 7.5. Here I present the subjective responses in detail.

7.3.6.1. Learning Curve Versus Efficiency

We examined the trade-offs between location-independent gestures (LI_{any}) and location-specific gestures (LS_{palm} and LS_{body}), focusing on efficiency and learning curve. As mentioned earlier, all participants commented that they liked having different locations mapped for different applications, mainly because directly interacting with specific locations is faster than swiping (7 responses).

We further looked into responses for each of the location-specific interfaces in detail. For LS_{palm}, for example, five participants valued the feature that allowed participants to touch their finger down and move the finger to explore different locations on palm until they found each target application (called "explore by touch"), which supports even novice interaction. In this regard, P5 stated that *"So someone who's new to it, I think they can pick it up relatively simple. I think the browsing part and being able to also open the apps but then I can go back the center of my phone and still browse through that. So, I like the look, how it's set up, I think that was pretty good"*. For LS_{body}, five participants considered the locations were easy to learn and remember, while two of them considered LI_{any} to be more difficult to learn as the menu orders are less intuitive.

On the other hand, half of the participants considered having to learn and remember the specific locations as one of the drawbacks for location-specific gestures

(LS_{palm} and LS_{body}). In comparison, participants for the LI_{any} interface may be lower as they did not have to learn the mapping between the location and the applications.

7.3.6.2. Availability of Input Locations

Six participants liked LI_{any} for being able to perform input anywhere on their body, which provides more options when one location is not available (e.g., one hand holding a cane). P12, for example, said that *"If I could do it with either hand, me holding my cane, I would choose number three [LI_{any}]. Because I have my cane in my left hand and I could always tap my other side. Or wherever or even tap anywhere. If I had a cane and I would have to stop walking to do the second one [LS_{palm}]. Because I would have to hold my cane then stop and feel around my hand. Number three, I mean number one [LS_{body}]... It would only be convenient if I stop walking. I think".* At the same time, she was concerned about her entire body becoming a touch-sensitive input surface, as it would increase the chance of unintentional input.

7.3.6.3. Touching Palm Versus Across the Body

When two *LS* interfaces are compared, all participants except one (P8) preferred the LS_{palm} over the LS_{body} interface due to social acceptability. Participants reported not feeling comfortable using the LS_{body} interface in public, concerning how they might be seen by others. Less dominant, but seven participants also mentioned another drawback for the LS_{body} interface that it requires too much movements; they did not like having to go all over the body. Instead, they wished to have a single location as in LS_{palm}. For instance, P12 said that *"If it's on this one [palm], it's easy to search through. Because if it's not here, it's somewhere in this area. But if you have it on your wrist and then your palm and on your ear, you gotta remember which on you need to touch. Because*

		Sighted	VI_{audio}
Concerns	**Technical challenges**	24.1%	10.5%
	Learning curve	11.1%	7.0%
	Social acceptability	3.7%	12.3%
	Privacy	3.7%	5.3%
	Form factor	3.7%	3.5%
	Display quality	3.7%	1.8%
	Carrying multiple devices	1.9%	5.3%
Expectations	**Larger input/output space**	11.1%	0.0%
	Quick/easy access	9.3%	5.3%
	Wider range of input	9.3%	0.0%
	Accessibility	5.6%	5.3%
	Hands-free	3.7%	0.0%
Other	**Positive**	29.6%	38.6%
	Negative	7.4%	12.3%
	Misc. or no responses	3.6%	6.6%

Table 7.7. Attitude towards on-body interaction, sorted by the total number of responses of sighted participants from largest to smallest ($N_{sighted} = 54$, and $N_{VIaudio} = 57$). Note that, the responses coded as "*Positive*" and "*Negative*" under "*Other*" are general comments without specific reasons such as "Great!" or "Not a good idea.", thus do not add up to 100%.

if I'm late for work and I'm feeling my palm trying to get the clock and I'm like oh man it's not there, it's on my wrist. It's more convenient. More convenient to either rolling on your arm or tap once on your leg and double tap and things like that". Meanwhile, remaining participants liked the physically distinctive locations of the LS_{body} interface compared to the LS_{palm} interface, as the locations on the palm are too condensed.

7.3.7. Attitude Towards On-body Interaction

To examine the general attitude towards on-body interaction for VI users and how it differs from sighted users, we asked participants from both online survey and in-person study what they think about the idea of on-body interaction after providing a brief description specified in Section 7.2.1.2. The responses from the online survey are summarized in Table 7.7. Participants across both user groups were mostly concerned with sensing accuracy and robustness for preventing accidental/unintended input (*"technical challenges": 10.5% to 24.1%*). But the next most concern was different across user groups: it was learning and getting familiar with the new interface for sighted participants (*"learning curve": 11.1%*) while it was *"social acceptability"* for the VI_{audio} group (12.3%). Moreover, participants from the sighted group provided more various expectations that none of VI_{audio} participants mentioned such as "*larger input/output space*", "*wide range of input*", and "*hands-free*". The percentages of all coded responses were not different across the user groups, except for "*Larger input space*", and "*Like the* location-specific *gestures*" where the percentage of VI_{audio} group was zero.

We also asked interview participants what they thought about the idea of on-body interaction before having participants interact with the design probes. The majority of participants (*N*=9 of 12) showed a positive attitude towards on-body interaction with more various reasons than from the survey. The most frequent reasons were quick and easy access by not having to pull out the phone (*N*=7), and hands-free interaction (*N*=4). Three participants mentioned that they like the idea of having different locations mapped for different tasks, which was found to be preferred in Chapters 3 and 4. But again, three participants showed their concern of social acceptability when performing gestures on their body in public.

We then compared the responses to how participants felt after gaining hands-on experience with the design probes. All participants liked the idea of mapping specific locations of their body for opening specific apps, regardless of their most preferred design (*i.e.*, regardless of whether they had preferred LS_{body} or not). Eight participants mentioned that at least one of the design probes was just like using their phone, half of the participants said that they liked on-body interaction as it could leave their hands free, and five participants liked the fact that they did not have to pull out their phone. However, overall, most responses were the same as they had been before using the real-time system, but new concerns related to form factors arose, particularly: having to carry multiple devices (four responses), and speaker and sensor locations (three responses). P12, for example, said that *"So will you still have like the head the earphones or will you do like a Bluetooth? Or how would that work? […] You don't want to have a boom box on your shoulder"*. Two participants (P3 and P10) commented that they would no longer have to worry about their phone being stolen. P3, for example, said *"Especially when you provide a more secure way of having a smart device computing, a system that's way more secure. No one is going to rob you for your hand"*.

While all participants could see the value of on-body interaction, three participants reported that they would like to use a physical object rather than their body (P1, P7, and P9). For instance, P1 said that *"They [smartphone and watches] are machines and I can work upon whenever I'm, whatever task I want to do. But on the body itself it's strange and body has completely different purposes and I feel I don't want to mix between the nature's activities and what I would do with the machines. And I leave my body just for natural things"*.

7.3.8. Summary

Based on the online survey and interview results, we were able to identify important tasks to be supported as microinteractions for VI_{audio} participants, and learn implications for designing on-body interaction by reflecting on the use of smartphones and expected use cases and perception of smartwatch and on-body interactions.

The most dominant microinteractions used across both sighted and VI user groups were: managing phone calls such as answering and making calls, checking notifications, checking the current time, checking the weather, and activating voice input. Specific to voice input, we confirmed that more VI participants use voice input than sighted participants, especially more frequently for phone calls and responding to text messages. Our results also showed that VI_{audio} participants reported slower completion time than sighted participants, especially for the tasks whose reported completion time for sighted participants were either extremely short or extremely long, or required text entry. In terms of perceived advantages and limitations, we found that a higher percentage of VI_{audio} group from the survey were more negative towards smartwatches than sighted participants, while there was no significant difference for on-body interaction.

From the design probe study, we found perceived trade-offs among the three interfaces. Although some participants liked LI_{any} for not having to learn any location mapping, LI_{any} was considered to be less efficient than LS_{palm} where they could directly point to certain locations instead of performing several swipes. When comparing LS_{palm} and LS_{body}, although some participants liked the contextual mapping of LS_{body} (e.g., tapping wrist for checking time), most participants preferred LS_{palm} because a set of condensed input locations on the palm were more efficient and less socially acceptable than having to go all over the body. Regardless of the trade-offs, all participants liked the idea of location-specific gestures.

7.4. Discussion

Here, I reflect on the implications of our findings, focusing on how on-body interaction can be designed to support microinteraction for people with visual impairments.

7.4.1. Microinteractions Specifically for VI Users

We identified nine tasks that would be valuable to support as microinteractions for screenreader users. The first five tasks were from frequently used tasks which include (1) answering and making calls, (2) checking notifications, (3) checking the current time, (4) checking the weather and (5) voice input. Based on the responses for specific use of voice input, additional four tasks were identified: (6) setting an alarm or timer, (7) responding to text messages, (8) getting navigation directions and (9) voice search to look up information. Especially for the tasks #7 and #8, special attention is needed when designing them as microinteractions since more percentage of VI_{audio} participants reported slower completion time than sighted participants.

7.4.2. Use of Voice Input for Enabling Microinteractions

As found in previous studies [13,135], our data confirmed that voice input is more popular for VI users than sighted users. However, unlike the finding from Azenkot et al. [13], we found that more VI_{audio} participants use voice input for tasks that involve text/numeric entry such as dialing a phone number and responding to a text message, compared to sighted participants. Additionally, "*voice search*" was a common response from interview participants as found in [13]. Furthermore, voice input was considered to be faster for five interview participants compared to manual input, and four showed strong preference, stating that they would use voice input if possible. Considering the results from our study and prior finding that showed how text entry is challenging on a touchscreen for users with visual impairments (e.g., [19,31,87]), these results suggest that VI users might use voice input strategically to relieve the inefficiencies of touchscreen-based screenreader interaction for a wide range of microinteractions. Thus, as identified earlier, voice input is critical to include in any wearable device that is designed to support microinteractions.

While voice input alone can be used for supporting microinteraction, voice input may not be always available in noisy environment or in front of others due to privacy concerns [2,13,135]. As such manual input is still needed with the wearable device (e.g., in an on-body interface) to complement voice input, for example, to support frequently used voice commands like setting an alarm or for something private (e.g., *"call Darling"*).

7.4.3. On-Body Interaction as a Complementary Means of Mobile Computing

Both sighted and VI participants considered a smartwatch to be additional device to carry that has limited or redundant capability with possibly increased input difficulties due to the small touchscreen. In contrast, only VI participants were concerned of sound-related issues of smartwatches (e.g., volume not being loud enough). Contradicting a prior study that concluded VI users have a positive attitude towards wearable devices in general [135], our participants with visual impairments showed more negative attitude towards a smartwatch than sighted participants. This suggests the issue of sound quality may be the cause of the VI participants' negative perception of a smartwatch. On-body interaction, in contrast, overcomes this limitation of smartwatches by providing an enlarged input area (e.g., palm or forearm). Furthermore, on-body interaction offers many of the same wearable device advantages of the smartwatch, such as quick access and hands-free interaction—both of which are desirable for people with visual impairments, particularly in a mobile context where their hand is occupied for holding a cane (see Chapter 4). In addition, because users are interacting with their own body, concern about their phone being stolen is eliminated. As such, on-body interaction may be more ideal for supporting microinteractions than a smartphone or a smartwatch for users with visual impairments.

7.4.4. Combination of Location-Independent and Location-Specific Gestures

We found that each of the design probes we compared was perceived to have its own advantages and limitations. Thus, here I list each of the strength and weakness as implications for designing on-body interaction for users with visual impairments.

7.4.4.1. Reduce the Learning Curve for Location-Specific Gestures

Having to learn the location where each application is mapped was found to be the main drawback of location-specific gestures. While some participants commented that the location mapping made sense and help them learn the interface at ease for LS_{body}, most of the mapping were arbitrary especially for LS_{palm}. One way to reduce the learning curve is by transferring an interface layout that users are already familiar with, as users can recall the spatial mapping of the layout [37,39]; for example, users can interact on specific locations on their body (e.g., palm, forearm or thigh) while imagining that the applications on their phone are mapped. Supporting the *"explore by touch"* feature would also reduce the learning curve as it allows novice users can eventually learn the mapping as they use the interface over time. However, to maximize the benefit from *explore by touch* feature, applications would need to be closely mapped in a single area as in LS_{palm}.

Supporting input customization would be another option, which participants expressed the interest from the design probe study. By enabling end-user customization the cost of learning location-specific gestures can be reduced as user-defined gestures are easier to remember than pre-defined gestures [77], although appropriate assistants (e.g., mixed-initiative feedback) might be necessary for a robust accuracy [83]. Input customization would also relieve the concern of social acceptability as users can choose locations that they feel are discreet and are comfortable using in public.

7.4.4.2. Support Multiple Options for Input Locations

Location-specific gestures may allow faster interaction as users can directly point to certain locations instead of performing several swipes. However, as commented by participants, it may temporarily be impossible to interact with a specific location, which means there would be no way for users to access certain apps. For example, when walking on a street with a cane, users may not be able to interact with the palm (location-specific palm-based interaction) because it requires two hands–one hand to perform gestures and the other hand to serve as an input surface (Chapter 4 and our prior study [135]). Also, interacting with the thigh would be problematic as the legs would be continuously moving while walking. In this regard, location-independent gestures may be a better option; however, as pointed out by one participant, it would also increase the chance of accidentally activating the system as the whole body would become touch sensitive. Thus, using both location-specific and location-independent gestures in combination would complement the trade-offs between location-specific and location-independent gestures with increased input flexibility than using either one.

7.5. Limitations

The primary limitation of the survey is that the data is all self-reported. Also, because the survey was conducted online, we likely had a bias toward technology savvy users. Moreover, because most participants had no experience with a smartwatch and it is highly unlikely that they had used on-body interaction either, their projections of these wearable devices might not necessarily be indicative of what users would experience in practice. Thus, an interview study on users who actually own smartwatches would be more useful for investigating the perception of smartwatches that are based on actual use. Furthermore, while gender may play a significant role in wearable technology acceptance [18,97], the visually impaired and sighted groups had different age and gender distributions. Thus, a follow up study is needed for direct comparisons of the two groups.

Regarding the design probe study, the subjective feedback might be different with a more practical form factor that is smaller and self-contained as opposed to our physical prototype. Moreover, although we stressed that our system is an early prototype to explore the *possibility* of on-body interaction, the training process may have negatively impacted participants' opinions. Also, the differences in amount of training example required and input recognition accuracy across design probes and participants might have affected users' responses differently. For example, the average number of training examples required varied from 9.08 (thigh) to 14.75 (inner wrist) for input locations, 8.22 (P3) to 17.56 (P5) for participants. Thus, robust recognition accuracies across the interfaces and participants would have been ideal. Furthermore, while we only had a two-level menu hierarchy with five main menu items where each of them had none to four sub-menus, users' subjective ratings may be different with a different layout or a greater number of items, as is the case for mouse-based visual search time in a menu [16,46]. As such, a future work is needed to thoroughly compare the task performance time with different menu lengths and layouts.

7.6. Conclusion

Our results demonstrate how current microinteractions that are designed to be quick and simple are still time-consuming for people with visual impairments, and that the perceived benefits of having smartwatches for microinteraction may not be greater than the limitation of a small-sized touchscreen, which makes input more difficult compared to the screen on smartphones. Although participants with visual impairments were less positive about smartwatches than sighted users, they were generally positive about the idea of on-body interaction, as a small input space would be no longer be an issue while providing the advantages that a smartwatch can offer (e.g., quick access, hands-free interaction). Furthermore, we also found that participants liked location-specific on-body gestures while they considered having to learn the location mapping to be a drawback. Thus, examining to what extent input customization or transferring already familiar interface layout relieves the learning curve would be an interesting future work.

Chapter 8: Conclusion and Future Work

As I stated in the Introduction (Chapter 1), the goal of this research was to support accessible mobile computing for users with visual impairments through on-body interaction as an alternative or complement to touchscreen interactions. The specific objectives of the dissertation included: (1) identification of perceived advantages and limitations of on-body input compared to a touchscreen phone, (2) assessment of performance benefits of on-body input over touchscreen input in terms of accuracy and efficiency, (3) implementation and evaluation of an on-body gesture recognizer using finger-, and wrist-mounted sensors, and (4) design implications for accessible and efficient non-visual on-body interaction for people with visual impairments.

To achieve these objectives, I have demonstrated that users with visual impairments have a greater tendency than sighted users to create location-specific gestures (e.g., pointing to a specific finger for different tasks) and static gestures by exploring on-body gestures created by 13 sighted participants and 11 VI participants (Chapter 3). Building on this preliminary result, I have also found that on-body input was considered to be especially useful for contexts where one hand is busy (e.g., holding a cane or dog leash) when compared to mobile phone with touchscreen (Chapter 4). Moreover, I have empirically confirmed the performance benefit of on-body input over touchscreen input for participants with visual impairments, extending previous pointing input results with only sighted users [38] both to this new user group and to the more complex shape-based gestures (Chapter 5). Following the design and performance studies, I have implemented an on-body gesture recognizer based on finger- and wrist-mounted sensor values for supporting location-specific gestures, and with offline performance evaluations, demonstrated that our finger-mounted input sensing system is feasible for supporting 24 location-specific gestures with the accuracies of 94.9% (Chapter 6). Finally, I have identified nine tasks that would be important to be supported as microinteractions for people with visual impairments through on-body interaction, and investigated design implications of on-body interaction for enabling microinteraction, focusing on location-specific gestures (Chapter 7).

8.1. Design Implications

Here I summarize design guidelines for supporting accessible on-body interaction for people with visual impairments from lessons learned throughout this dissertation. These guidelines should inform further work on on-body interaction with people with visual impairments. Because they have been derived from lab studies alone, they will need to be confirmed in long-term field studies.

8.1.1. Using the Hand as a Default Input Location

For supporting non-visual mobile computing through on-body interaction, the hand location may offer a number of advantages as an input surface over other locations (e.g., forearm, wrist, thigh), regardless of users' level of vision. Here, I summarize the advantages.

8.1.1.1. Familiar Experience as Smartphone-based Interaction

The hand, especially the palm side, can provide a familiar experience to users, which can also reduce the learning curve. In Chapter 3, when participants were asked to create any two-handed gestures (with the hands touching) for mobile actions, such as returning to the home screen or opening a selected application, we observed that they had a high tendency to perform gestures with one hand on the palm of the other hand as if they were using the touchscreen of their phone. Furthermore, based on the subjective responses from Chapters 4 and 7, this mimicking behavior can reduce learning cost for users. When interacting with their hand, they can transfer their knowledge of the spatial layout of the actual phone (called *transfer learning*) to their hand as if they are holding an imaginary phone (conceptualized as called *imaginary interface*) [39].

8.1.1.2. More Socially Acceptable Than Other Input Locations

The hand location can offer interaction that is easy to access and is discreet compared to other body locations such as the wrist, ear, thigh, or neck and face areas. More importantly, interacting with the hand was considered to be more socially acceptable than the other input locations we examined (Chapters 4 and 7). Considering the finding from Chapter 4, which suggested that users may more concerned about social acceptability of on-body interaction than about having high ease of use and physical comfort, the hand-based location will likely be preferred by users especially in public settings.

8.1.1.3. Performance Benefits Over Touchscreen Input for Non-Visual Interaction

The hand location can enable better performance than a smooth touchscreen of the phone when visual cues are absent. The palm was found to be faster and more accurate performance for target pointing, and more consistent and accurate for shape drawing than on a touchscreen of a phone for participants with visual impairments and blindfolded sighted participants (Chapter 5). For target pointing, specifically, participants outperformed if the targets were located on the tips of the finger than other locations on the fingers or on the palm. As such, for people with visual impairments or even for sighted people under eyes-free conditions, we recommend using the palm as an input surface for on-body interaction as opposed to a smooth touchscreen of the phone for better performance.

8.1.2. Supporting Always-Available Interaction

On-body interaction can support always-available input for people with visual impairments especially in mobile contexts when one of the hands is busy holding a cane or dog leash.

8.1.2.1. Supporting One-Handed Interaction

While one-handed interaction was found to be a desirable feature for people with visual impairments, especially in a mobile context, participants reported that it is physically challenging to use a single hand to secure the phone and perform gestures at the same time on a hand-held smartphone. As such, users would often postpone the use of their devices or stop walking to interact with their phone (Chapters 4 and 7), as revealed in prior studies [2,135]. In contrast, one-handed interaction can easily be supported with on-body input since it can free users' hands without having to hold and secure the physical device. Thus, on-body interaction can serve as a complementary means of mobile computing for people with visual impairments for supporting interaction on the go when both hands are not available.

8.1.2.2. Supporting Multiple Input Locations and Modes

A downside of on-body interaction is that depending on the input locations used, not all input locations will be available at all times. For example, while the palm has several advantages as mentioned above, physical access can be limited when holding a cane as both hands are needed to interact with the location. The thigh location may also be problematic in practice as legs would be constantly moving while walking. In this regard, supporting multiple input locations or location-independent gestures, where gestures are not restricted to their performed locations, and thus can be performed anywhere on the body, should be considered to enable always-available interaction.

8.1.3. Enabling Location-Specific Gestures

Gestures performed on specific locations or landmarks on users' body (namely location-specific gestures) were preferred by people with visual impairments in general (Chapters 3, 4 and 7). However, careful considerations will be needed when designing location specific gestures for on-body input.

8.1.3.1. Combining Location-Independent and Location-Specific Gestures

As mentioned earlier, location-independent gestures allow always-available input as gestures can be performed anywhere on the user's body. Moreover, as opposed to location-specific gestures, less learning will be required for location-independent gestures if item navigations can be done with directional swipes (e.g., a *left-to-right* swipe) as users do not need to memorize the mapping between the locations and their corresponding items. In contrast, location-specific gestures may offer faster interaction, especially when the mapped locations are close together, since users can directly go to specific locations instead of several swipes once users become familiar with the interface. Although further investigation is needed as our study only included a relatively small set of five locations in Chapter 7, the perceived trade-offs between the two types of on-body gestures suggest that each gesture type would complement the other. Thus, enabling both location-independent and location-specific gestures is recommended for supporting both novice and expert users similar to how iOS and Android support both gesture-based navigation as well as *exploration by touch*.

8.1.3.2. Location-Specific Gestures With Customization

Appropriate on-body input locations may vary depending on users' preference or mental models. For example, some people might favor tapping their ears to activate voice input. That would allow them to memorize the mapping between the location and the application at ease, while others would tap their specific locations on the palm to invoke the same functionality because of its discreetness and easy access. One solution would be enabling end-user customization to meet these various mapping preferences, which would also reduce the learning curve as user-defined gestures are found to be more memorable than pre-defined gestures [77]. However, creating distinctive gestures in the perspective of the on-body input recognizer for achieving high accuracy would be challenging for users, as has been shown for touchscreen gesture customization [83]. As such, technical feedback should be provided when supporting end-user input customization for location-specific gestures.

8.1.4. Supporting Microinteraction

While on-body interaction may provide more efficient and always-available interaction for non-visual mobile computing for people with visual impairments as opposed to a touchscreen, some tasks might be better performed with other input devices. Text entry, for instance, would be faster with a physical keyboard (either braille or QWERTY) than a virtual keyboard on a smartphone or on the user's own skin surface. Voice input is also highly efficient for some tasks, such as searching for a specific item such as an app or contact information. In this regard, the goal of designing on-body interaction should not be to replace but to augment or complement other possible input methods by supporting a set of microinteractions that are most amenable to on-body input. For example, on-body input can be useful when voice input does not work well in a noisy environment, or when the privacy becomes a concern.

8.2. Summary of Contributions

The work I present in this dissertation makes several contributions:

- The potential and challenges of on-body input for people with visual impairments:
    - Identification of perceived advantages and disadvantages of on-body input compared to a touchscreen phone (Chapters 4 and 7).
    - Empirical evidence showing the performance benefits of on-body input over touchscreen input in terms of speed and accuracy (Chapter 5).
- Design implications of non-visual, on-body interaction specifically for people with visual impairments:
    - A characterization of the preferences for different on-body input locations (Chapter 4).
    - Empirical evidence that there is a measurable difference between subjective usability ratings for on-body interaction and touchscreen interaction for different hand availabilities: one versus two hands (Chapter 4).
- Investigation of an on-body input sensing system using finger-mounted sensors:

    - Implementation of an on-body gesture recognizer based on finger-mounted sensors such as accelerometer, gyroscope and magnetometers (Chapter 6).
    - An offline evaluation of the on-body gesture recognition accuracy comparing different input conditions such as choice of sensors and sensor placement (Chapter 6).
- Design and evaluation of on-body interaction for supporting accessible microinteractions:
    - Identification of the challenges people with visual impairments face with microinteractions and whether these differ from sighted users (Chapter 7).
    - Design guidelines based on the comparison of three different on-body interaction design prototypes for supporting microinteractions (Chapter 7).
    - Perceived trade-offs between on-body interaction versus mainstream mobile and wearable technologies (Chapter 7).

8.3. Future Work

With the investigations and designs presented in this dissertations, I have demonstrated that on-body interaction can provide an alternative or complementary means of accessible mobile computing for visually impaired users, with improved speed and accuracy as compared to touchscreen interaction. From here I have several directions for continued research aiming to support a wider range of users and applications.

8.3.1. Re-Implementation of On-Body Input Systems

To either solicit subjective feedback from participants or to demonstrate the technical feasibilities, the on-body input sensing systems I implemented and used throughout the dissertation did serve the purpose as research prototypes. However, to increase the flexibility and extensibility for future work of on-body interaction for people with visual impairments, the sensing techniques and the form factor has to be improved.

In terms of sensing techniques, because the data was collected and tested immediately while participants were seated still in our studies, the accuracy of our current on-body input sensing approach is likely to be dropped down, requiring longer calibration in realistic contexts. Also, even a finger-mounted camera-based approach may not be free from out-of-frame issues when the input location is narrow or near the edge (e.g., fingertip, wrist), or when the pointing angle of the finger is parallel to the input surface instead of with some angle (e.g., a contact made with a finger pad rather than the tip of the finger). Thus, the re-implementation of a robust sensing algorithms considering users' various input postures and locations would be necessary to compete with current mainstream devices (e.g., touchscreen-based sensing). Although, they may not be directly applied for supporting on-body interaction, leveraging state-of-art camera-based sensing approaches for recognizing hand gestures in general (see survey [99]) would be a good starting point for the next step.

The importance of designing on-body input sensing to be streamlined and not bulky (unlike our prototype system) is also important for a wide adoption of on-body interaction as users with visual impairments tend to prefer mainstream devices that are accessible rather than specific assistive devices that may stand out and incur social stigma [111]. Moreover, tactile enhancement on input locations on the body would be useful for supporting accessible on-body interaction, especially for the locations that are tactilely less distinctive than others (e.g., specific segments of fingers versus specifics location on forearm). To provide tactile feedback for users to find specific location precisely at ease, I plan to investigate into on-skin sensing techniques such as conductive on-body stickers [127] or tattoos [54].

8.3.2. Further Evaluation

Current qualitative evaluations of on-body interaction have only provided the results from single-session controlled lab settings or an online survey. Thus, the collected responses might not necessarily be indicative of what users would experience or perceive in practice. For example, while many of the questions were related to smartwatches for the online survey in Chapter 7, only 13 participants (out of 117) owned a smartwatch. As such, interviewing smartwatch owners would offer more rich and realistic data than what I have collected. Moreover, as for evaluating on-body interaction, field studies would allow me to observe and study their actual usage and subjective feedback over time, focusing on social acceptability and interaction on-the-go. To be specific, I am interested in studying the perspective of people, both onlookers and bystanders as in [96,97], of using on-body interaction in varies contexts (e.g., at a library, walking on a street). Furthermore, I plan to investigate the usability of location-specific gestures in depth. Along with the qualitative analysis I have done in Chapter 7, I would like to assess users' performance on different body locations, comparing location-independent and location-specific gestures in terms of speed and recall rate. Examining failure cases or the cause of low performance would also be valuable to get insights for achieving inclusive design. Further, while the location-specific gestures we examined involved tapping gesture only in Chapter 7, I plan to include more complex gestures (e.g., geometric shapes, multi-finger swipes), or even mid-air three-dimensional gestures for a comparison.

8.3.3. Extension to Other Applications

Although my dissertation has focused primary on supporting accessible mobile computing through on-body interaction, I would like to expend this research to authentication and text entry, which were found to be slower in prior studies and, confirmed in my study as well.

As for authentication, location-specific on-body input can be used as an authentication approach like fingerprint identification could also easily be incorporated if using our camera-based sensing approach, as shown in our prior study [114] and Chapter 6. For example, simply tapping a thumb can unlock the phone by extracting the fingerprint while users are performing a tap, which would be faster than having to enter a 4-digit passcode, which takes about 7.5 seconds [14]. Moreover, this approach would be more secure because it can prevent aural eavesdropping that is problematic for screenreader users, where the passcode would be spoken out loud in front of others [12,17,30].

Also, I would like to investigate how on-body interaction can be designed to support text entry. Handwriting-based text entry has been explored by several on-body interaction studies (e.g., [23,126]), however, symbolic gestures are not preferred by blind people [53]. Thus, instead of handwriting, I would like to study on-screen braille and QWERTY keyboard-based input methods that are location-specific (e.g., [29] or soft keyboard) versus location-independent, where the keyboard layout adaptively changes based on the relative typing locations (e.g., [29] or a braille keyboard supported by iOS VoiceOver) to learn design implications for supporting text entry though on-body interaction.

8.3.4. Extension to Other User Groups

Although I used the term "people with visual impairments," the focus of this dissertation has been supporting people who have little or no functional vision, which is only less than 10% of people with visual impairments; most of them have some functional vision [138]. As such, audio output was provided for the demonstrations of on-body interaction rather than visual output. However, for people with any level of functional vision, including people with low vision and sighted people, might prefer visual feedback than other modes of output. On the other hand, for deafblind users, neither audio nor visual feedback would be accessible. Thus, different modes of output should be supported for different users. Furthermore, I would like to investigate how always-available input enabled by on-body interaction can support situation impairments (e.g., hands or visual channels being occupied like driving), which users, both with and without disabilities, would experience more often as computing devices will be used in more diverse situations with a wide adoption of wearable technologies [108]. Overall, deploying ability-based designs principles [132] to understand specific design requirements for each of the user groups with different needs would eventually lead to a universal design [121,122].

8.3.5. One Customized Interface for Multiple Systems

Lastly, I plan to investigate end-user input customization as a continuation of my earlier work [83] to lower the barriers for learning and using technologies. The gestures customized by the users for themselves have several advantages over pre-defined gestures: improving memorability [77] and offering quick access to information [92]. Gesture customization may also be used to improve the accessibility introduced by the design choices of certain interfaces for people with different needs [7]. Having these benefits, the custom on-body gestures would be useful to be supported as a part of Global Public Inclusive Infrastructure (GPII) [16], specifically for Automated Personalization Computing Project (APCP). For instance, one set of gestures can be performed on users' own bodies instead of physically interacting with an input device for any new or different interface that users encounter, without having to learn or switch to different types of input modes for different interfaces (e.g., public kiosks, home appliances).

8.4 Final Remarks

In this dissertation, I have demonstrated the following thesis statement:

> *On-body interaction can provide an alternative or complementary means of accessible mobile computing for visually impaired users, with improved speed and accuracy as compared to touchscreen interaction.*

I have shown that participants with visual impairments have a positive attitude towards on-body interaction in general mainly because of quick access and ability to keep the hands freer than when using a smartphone. While a smartwatch can also offer these two benefits, more VI participants were negative about it due to its small touchscreen.

[16] This Project is being led by Trace Center (directed by Dr. Gregg Vanderheiden) in the iSchool at the University of Maryland, College Park.

Furthermore, in terms of performance of on-body interaction compared to touchscreen-based interaction, I have shown that VI participants were faster and more accurate for target pointing tasks, and more consistent for shape drawing tasks on their hand, than on a smooth screen of a phone.

I have demonstrated the advantages of on-body interaction, in terms of both qualitative feedback and empirical performance gain over touchscreen interaction, in support of mobile computing for people with visual impairments. However, for on-body interaction to truly support accessible mobile computing, it needs to become a mainstream technology, rather than an assistive technology which is used by only people with visual impairments. Without widespread adoption of on-body interaction, it may stand out and incur social stigma [111]—this concern of social acceptability was raised by a number of participants in my study. Thus, based on the lessons I have learned from people with visual impairments throughout this dissertation, I plan to continue this research with the goal of supporting on-body interaction as a mainstream technology for a wider range of users, with and without disabilities. It will also be important to determine how on-body interaction fits into a larger ecosystem of mobile computing technologies to best complement the strengths of devices such as smartphones, smartwatches, other emerging wearable devices.

Bibliography

1. Julio Abascal and Antón Civit. 2000. Mobile communication for people with disabilities and older people: New opportunities for autonomous life. *Proceedings of the 6th ERCIM Workshop*, 255–268.
2. Ali Abdolrahmani, Ravi Kuber, and Amy Hurst. 2016. An Empirical Investigation of the Situationally-Induced Impairments Experienced by Blind Mobile Device Users.
3. Dustin Adams, Lourdes Morales, and Sri Kurniawan. 2013. A qualitative study to support a blind photography mobile application. *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '13*: 1–8. http://doi.org/10.1145/2504335.2504360
4. Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. 2004. Face Recognition with Local Binary Patterns. In *Computer Vision - ECCV 2004*. 469–481. http://doi.org/10.1007/978-3-540-24670-1_36
5. Roland Aigner, Daniel Wigdor, Hrvoje Benko, et al. 2012. Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for hci. *Microsoft Research TechReport MSR-TR-2012-111*.
6. Farshid Amirabdollahian, Rui Loureiro, and William Harwin. 2002. Minimum jerk trajectory control for rehabilitation and haptic applications. *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, 3380–3385.
7. Lisa Anthony, Yoojin Kim, and Leah Findlater. 2013. Analyzing user-generated youtube videos to understand touchscreen use by people with motor impairments. *Proc. CHI 2013*: 1223–1232. http://doi.org/10.1145/2470654.2466158
8. Lisa Anthony and Jacob O Wobbrock. 2010. A lightweight multistroke recognizer for user interface prototypes. *Proc. of Graphics Interface 2010*, Canadian Information Processing Society, 245–252.
9. Daniel Ashbrook. 2010. Enabling mobile microinteractions. *PhD Thesis*, May.
10. Daniel Ashbrook, Patrick Baudisch, and Sean White. 2011. Nenya: Subtle and Eyes-free Mobile Input with a Magnetically-tracked Finger Ring. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2043–2046. http://doi.org/10.1145/1978942.1979238
11. Daniel L Ashbrook, James R Clawson, Kent Lyons, Thad E. Starner, and Nirmal Patel. 2008. Quickdraw: The Impact of Mobility and On-Body Placement on Device Access Time. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*: 219–222. http://doi.org/10.1145/1357054.1357092
12. Dmitri Asonov and Rakesh Agrawal. 2004. Keyboard acoustic emanations. *IEEE Symposium on Security and Privacy*, 3–11.
13. Shiri Azenkot and Nicole B Lee. 2013. Exploring the use of speech input by blind people on mobile devices. *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility*, ACM, Article No. 11.

14. Shiri Azenkot, Kyle Rector, Richard E. Ladner, and Jacob O. Wobbrock. 2012. PassChords: secure multi-touch authentication for blind people. *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 159–166.
15. Shiri Azenkot, Jacob O Wobbrock, Sanjana Prasain, and Richard E Ladner. 2012. Input Finger Detection for Nonvisual Touch Screen Text Entry in Perkinput. *Proceedings of Graphics Interface 2012*, Canadian Information Processing Society, 121–129.
16. Gilles Bailly, Antti Oulasvirta, Duncan P Brumby, and Andrew Howes. 2014. Model of visual search and selection time in linear menus. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3865–3874.
17. Yigael Berger, Avishai Wool, and Arie Yeredor. 2006. Dictionary attacks using keyboard acoustic emanations. *Proceedings of the 13th ACM conference on Computer and communications security*, 245–254.
18. Kerry Bodine and Francine Gemperle. 2003. Effects of Functionality on Perceived Comfort of Wearables. *Proceedings of the 7th IEEE International Symposium on Wearable Computers*, IEEE Computer Society, 57--.
19. Matthew N Bonner, Jeremy T Brudvik, Gregory D Abowd, and W Keith Edwards. 2010. No-look Notes: Accessible Eyes-free Multi-touch Text Entry. *Proceedings of the 8th International Conference on Pervasive Computing*, Springer-Verlag, 409–426. http://doi.org/10.1007/978-3-642-12654-3_24
20. Robert W Van Boven, Roy H Hamilton, Thomas Kauffman, Julian P Keenan, and Alvaro Pascual--Leone. 2000. Tactile spatial resolution in blind Braille readers. *Neurology* 54, 12: 2230–2236.
21. Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2: 77–101.
22. John Canny. 1986. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8, 6: 679–698. http://doi.org/10.1109/TPAMI.1986.4767851
23. Liwei Chan, Yi-Ling Chen, Chi-Hao Hsieh, Rong-Hao Liang, and Bing-Yu Chen. 2015. CyclopsRing: Enabling Whole-Hand and Context-Aware Interactions Through a Fisheye Ring. *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, ACM Press, 549–556. http://doi.org/10.1145/2807442.2807450
24. Liwei Chan, Rong-Hao Liang, Ming-Chang Tsai, et al. 2013. FingerPad: Private and Subtle Interaction Using Fingertips. *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, ACM, 255–260. http://doi.org/10.1145/2501988.2502016
25. James C Craig and Keith B Lyle. 2001. A comparison of tactile spatial sensitivity on the palm and fingerpad. *Perception & psychophysics* 63, 2: 337–347.
26. Dimitrios Dakopoulos and Nikolaos G Bourbakis. 2010. Wearable obstacle avoidance electronic travel aids for blind: a survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 40, 1: 25–35.
27. Niloofar Dezfuli, Mohammadreza Khalilbeigi, Jochen Huber, Florian Müller,

and Max Mühlhäuser. 2012. PalmRC: Imaginary Palm-based Remote Control for Eyes-free Television Interaction. *Proceedings of the 10th European Conference on Interactive Tv and Video*, ACM, 27–34. http://doi.org/10.1145/2325616.2325623

28. Julien Doublet, Marinette Revenu, and Olivier Lepetit. 2007. Robust GrayScale Distribution Estimation for Contactless Palmprint Recognition. *IEEE Conference on Biometrics: Theory, Applications, and Systems 2007*, IEEE, 1–6. http://doi.org/10.1109/BTAS.2007.4401935
29. Leah Findlater and Jacob Wobbrock. 2012. Personalized Input: Improving Ten-Finger Touchscreen Typing through Automatic Adaptation. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*: 815–824. http://doi.org/10.1145/2207676.2208520
30. Denis Foo Kune and Yongdae Kim. 2010. Timing attacks on pin input devices. *Proceedings of the 17th ACM conference on Computer and communications security*, 678–680.
31. Brian Frey, Caleb Southern, and Mario Romero. 2011. Brailletouch: Mobile Texting for the Visually Impaired. *Proceedings of the 6th International Conference on Universal Access in Human-computer Interaction: Context Diversity - Volume Part III*, Springer-Verlag, 19–25.
32. Florence Gaunet, Yves Rossetti, and others. 2006. Effects of visual deprivation on space representation: Immediate and delayed pointing toward memorised proprioceptive targets. *PERCEPTION-LONDON-* 35, 1: 107.
33. F. Gemperle, C. Kasabach, J. Stivoric, M. Bauer, and R. Martin. 1998. Design for wearability. *In Proc. ISWC '98*, IEEE Comput. Soc, 116–122.
34. Daniel Goldreich and Ingrid M Kanics. 2003. Tactile acuity is enhanced in blindness. *The Journal of Neuroscience* 23, 8: 3439–3445.
35. T. M. Greiner. 1991. *Hand Anthropometry of U.S. Army Personnel.*
36. Tiago Guerreiro, Paulo Lagoá, Hugo Nicolau, Pedro Santana, and Joaquim Jorge. 2008. Mobile text-entry models for people with disabilities. *Proceedings of the 15th European conference on Cognitive ergonomics the ergonomics of cool interaction - ECCE '08*: 1. http://doi.org/10.1145/1473018.1473067
37. Sean Gustafson. 2012. Imaginary Interfaces: Touchscreen-like Interaction Without the Screen. *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, ACM, 927–930. http://doi.org/10.1145/2212776.2212867
38. Sean G Gustafson, Bernhard Rabe, and Patrick M Baudisch. 2013. Understanding Palm-based Imaginary Interfaces: The Role of Visual and Tactile Cues when Browsing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 889–898. http://doi.org/10.1145/2470654.2466114
39. Sean Gustafson, Christian Holz, and Patrick Baudisch. 2011. Imaginary Phone: Learning Imaginary Interfaces by Transferring Spatial Memory from a Familiar Device. *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, ACM, 283–292. http://doi.org/10.1145/2047196.2047233
40. Chris Harrison, Hrvoje Benko, and Andrew D Wilson. 2011. OmniTouch: Wearable Multitouch Interaction Everywhere. *Proceedings of the 24th Annual*

ACM Symposium on User Interface Software and Technology, ACM, 441–450. http://doi.org/10.1145/2047196.2047255

41. Chris Harrison and Scott E Hudson. 2009. Abracadabra: Wireless, High-precision, and Unpowered Finger Input for Very Small Mobile Devices. *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, ACM, 121–124. http://doi.org/10.1145/1622176.1622199
42. Chris Harrison, Shilpa Ramamurthy, and Scott E Hudson. 2012. On-body Interaction: Armed and Dangerous. *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, ACM, 69–76. http://doi.org/10.1145/2148131.2148148
43. Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: Appropriating the Body As an Input Surface. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 453–462. http://doi.org/10.1145/1753326.1753394
44. Eve Hoggan, Stephen A Brewster, and Jody Johnston. 2008. Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1573–1582. http://doi.org/10.1145/1357054.1357300
45. Sture Holm. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*: 65–70.
46. Anthony J Hornof. 2001. Visual search and mouse-pointing in labeled versus unlabeled two-dimensional visual hierarchies. *ACM Transactions on Computer-Human Interaction (TOCHI)* 8, 3: 171–197.
47. Daniel J. Hruschka, Deborah Schwartz, Daphne C. St.John, Erin Picone-Decaro, Richard A. Jenkins, and James W. Carey. 2004. Reliability in Coding Open-Ended Data: Lessons Learned from HIV Behavioral Research. *Field Methods* 16, 3: 307–331. http://doi.org/10.1177/1525822X04266540
48. Allyn Jackson. 2002. Communications-The World of Blind Mathematicians. *Notices of the American Mathematical Society* 49, 10: 1246–1251.
49. Lei Jing, Yinghui Zhou, Zixue Cheng, and Tongjun Huang. 2012. Magic Ring: A Finger-Worn Device for Multiple Appliances Control Using Static Finger Gestures. *Sensors* 12, 5: 5775–5790. http://doi.org/10.3390/s120505775
50. Shaun K Kane, Jeffrey P Bigham, and Jacob O Wobbrock. 2008. Slide Rule: Making Mobile Touch Screens Accessible to Blind People Using Multi-touch Interaction Techniques. *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 73–80. http://doi.org/10.1145/1414471.1414487
51. Shaun K Kane, Chandrika Jayant, Jacob O Wobbrock, and Richard E Ladner. 2009. Freedom to Roam: A Study of Mobile Device Adoption and Accessibility for People with Visual and Motor Disabilities. *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 115–122. http://doi.org/10.1145/1639642.1639663
52. Shaun K Kane, Meredith Ringel Morris, and Jacob O Wobbrock. 2013. Touchplates: Low-cost Tactile Overlays for Visually Impaired Touch Screen Users. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 22:1--22:8.

http://doi.org/10.1145/2513383.2513442

53. Shaun K Kane, Jacob O Wobbrock, and Richard E Ladner. 2011. Usable Gestures for Blind People: Understanding Preference and Performance. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 413–422. http://doi.org/10.1145/1978942.1979001
54. Hsin-Liu Kao, Christian Holz, Asta Roseway, Andres Calvo, and Chris Schmandt. 2016. DuoSkin: Rapidly Prototyping On-Skin User Interfaces Using Skin-Friendly Materials. *The 15th International Semantic Web Conference*. http://doi.org/10.1145/2971763.2971777
55. David Kim, Otmar Hilliges, Shahram Izadi, et al. 2012. Digits: Freehand 3D Interactions Anywhere Using a Wrist-Worn Gloveless Sensor. ACM Symposium on User Interface Software and Technology.
56. Adams Kong, David Zhang, and Mohamed Kamel. 2009. A survey of palmprint recognition. *Pattern Recognition* 42, 7: 1408–1418. http://doi.org/10.1016/j.patcog.2009.01.018
57. Gordon Kurtenbach and William Buxton. 1994. User learning and performance with marking menus. *Conference companion on Human factors in computing systems CHI 94*: 258–264. http://doi.org/10.1145/191666.191759
58. Kerrie Laguna and Renée L Babcock. 1997. Computer anxiety in young and older adults: Implications for human-computer interactions in older populations. *Computers in human behavior* 13, 3: 317–326.
59. Byron Lahey, Audrey Girouard, Winslow Burleson, and Roel Vertegaal. 2011. PaperPhone. *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, January: 1303. http://doi.org/10.1145/1978942.1979136
60. Steven Landau and Lesley Wells. 2003. Merging tactile sensory input and audio data by means of the Talking Tactile Tablet. *Proc. EuroHaptics' 03*: 414–418.
61. Gierad Laput, Robert Xiao, Xiang'Anthony' Chen, Scott E Hudson, and Chris Harrison. 2014. Skin buttons: cheap, small, low-powered and clickable fixed-icon laser projectors. *Proceedings of the 27th annual ACM symposium on User interface software and technology*, 389–394.
62. Dar-Shyang Lee and Sargur N. Srihari. 1995. A theory of classifier combination: the neural network approach. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, IEEE Comput. Soc. Press, 42–45.
63. Gordon E Legge, Cindee Madison, Brenna N Vaughn, Allen M Y Cheong, and Joseph C Miller. 2008. Retention of high tactile acuity throughout the life span in blindness. *Perception & psychophysics* 70, 8: 1471–1488.
64. Sinikka Lehtinen-Railo and Jyrki Juurmaa. 1994. Effect of visual experience on lundocational judgements after perspective change in small-scale space. *Scandinavian journal of psychology* 35, 2: 175–183.
65. G Julian Lepinski, Tovi Grossman, and George Fitzmaurice. 2010. The design and evaluation of multitouch marking menus. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*: 2233–2242. http://doi.org/10.1145/1753326.1753663

66. Barbara Leporini, Maria Claudia Buzzi, and Marina Buzzi. 2012. Interacting with Mobile Devices via VoiceOver: Usability and Accessibility Issues. *Proceedings of the 24th Australian Computer-Human Interaction Conference*, ACM, 339–348. http://doi.org/10.1145/2414536.2414591
67. John P. Lewis. 1995. Fast Template Matching. *Vision Interface* 95, 120123: 15–19.
68. Rong-Hao Liang, Shu-Yang Lin, Chao-Huai Su, Kai-Yin Cheng, Bing-Yu Chen, and De-Nian Yang. 2011. SonarWatch: Appropriating the Forearm As a Slider Bar. *SIGGRAPH Asia 2011 Emerging Technologies*, ACM, 5:1--5:1. http://doi.org/10.1145/2073370.2073374
69. Shu-Yang Lin, Chao-Huai Su, Kai-Yin Cheng, Rong-Hao Liang, Tzu-Hao Kuo, and Bing-Yu Chen. 2011. Pub - Point Upon Body: Exploring Eyes-free Interaction and Methods on an Arm. *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, ACM, 481–488. http://doi.org/10.1145/2047196.2047259
70. Christian Loclair, Sean Gustafson, and Patrick Baudisch. 2010. PinchWatch: a wearable device for one-handed microinteractions. *Proc. MobileHCI*.
71. S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan. 2011. Estimation of IMU and MARG orientation using a gradient descent algorithm. *2011 IEEE International Conference on Rehabilitation Robotics*, IEEE, 1–7.
72. Roberto Manduchi and J Coughlan. 2008. Portable and mobile systems in assistive technology. In *Computers helping people with special needs*. Springer, 1078–1080.
73. Heinrich Braun Martin Riedmiller. 1992. RPROP - A Fast Adaptive Learning Algorithm. *Proceedings of the International Symposium on Computer and Information Science VII*.
74. Denys J. C. Matthies, Simon T. Perrault, Bodo Urban, and Shengdong Zhao. 2015. Botential: Localizing On-Body Gestures by Measuring Electrical Signatures on the Human Skin. *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15*, ACM Press, 207–216.
75. David McGookin, Stephen Brewster, and WeiWei Jiang. 2008. Investigating Touchscreen Accessibility for People with Visual Impairments. *Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges*, ACM, 298–307. http://doi.org/10.1145/1463160.1463193
76. Pranav Mistry, Pattie Maes, and Liyan Chang. 2009. WUW - Wear Ur World: A Wearable Gestural Interface. *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, ACM, 4111–4116. http://doi.org/10.1145/1520340.1520626
77. Miguel A Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of Pre-designed and User-defined Gesture Sets. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1099–1108. http://doi.org/10.1145/2470654.2466142
78. Suranga Nanayakkara, Roy Shilkrot, Kian Peen Yeo, and Pattie Maes. 2013. EyeRing: A Finger-worn Input Device for Seamless Interactions with Our Surroundings. *Proceedings of the 4th Augmented Human International*

Conference, ACM, 13–20. http://doi.org/10.1145/2459236.2459240
79. Loris Nanni and Alessandra Lumini. 2008. Local binary patterns for a hybrid fingerprint matcher. *Patt. Recog.* 41, 11: 3461–3466.
80. Masa Ogata and Michita Imai. 2015. SkinWatch: skin gesture interaction for smart watch. *Proceedings of the 6th Augmented Human International Conference on - AH '15*, ACM Press, 21–24.
81. Masa Ogata, Yuta Sugiura, Yasutoshi Makino, Masahiko Inami, and Michita Imai. 2013. SenSkin: adapting skin as a soft interface. *Proceedings of the 26th annual ACM symposium on User interface software and technology*, 539–544.
82. Masa Ogata, Yuta Sugiura, Yasutoshi Makino, Masahiko Inami, and Michita Imai. 2014. Augmenting a Wearable Display with Skin Surface as an Expanded Input Area. In *Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments*, Aaron Marcus (ed.). Springer International Publishing, Cham, 606–614. http://doi.org/10.1007/978-3-319-07626-3_57
83. Uran Oh and Leah Findlater. 2013. The Challenges and Potential of End-user Gesture Customization. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1129–1138. http://doi.org/10.1145/2470654.2466145
84. Uran Oh and Leah Findlater. 2014. Design of and Subjective Responses to on-Body Input for People With Visual Impairments. *Proceedings of ASSETS 2014*, 8.
85. Uran Oh and Leah Findlater. 2015. A Performance Comparison of On-hand versus On-phone Non-Visual Input by Blind and Sighted Users. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 10.
86. Uran Oh, Shaun K Kane, and Leah Findlater. 2013. Follow That Sound: Using Sonification and Corrective Verbal Feedback to Teach Touchscreen Gestures. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 13:1--13:8. http://doi.org/10.1145/2513383.2513455
87. João Oliveira, Tiago Guerreiro, Hugo Nicolau, Joaquim Jorge, and Daniel Gonçalves. 2011. Blind People and Mobile Touch-based Text-Entry: Acknowledging the Need for Different Flavors. *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '11*: 179. http://doi.org/10.1145/2049536.2049569
88. Goh Kah Ong Michael, Tee Connie, and Andrew Beng Jin Teoh. 2008. Touch-less palm print biometrics: Novel design and implementation. *Image and Vision Computing* 26, 12: 1551–1560. http://doi.org/10.1016/j.imavis.2008.06.010
89. OrCam Technologies Ltd. OrCam - See for Yourself.
90. Antti Oulasvirta. 2005. The fragmentation of attention in mobile interaction, and what to do with it. *Interactions* 12, 6: 16–18. http://doi.org/10.1145/1096554.1096555
91. Antti Oulasvirta, Sakari Tamminen, Virpi Roto, and Jaana Kuorelahti. 2005. Interaction in 4-Second bursts: The Fragmented Nature of Attentional Resources in Mobile HCI. *Proceedings of the SIGCHI conference on Human*

factors in computing systems - CHI '05: 919. http://doi.org/10.1145/1054972.1055101
92. Tom Ouyang and Yang Li. 2012. Bootstrapping Personal Gesture Shortcuts with the Wisdom of the Crowd and Handwriting Recognition. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2895–2904. http://doi.org/10.1145/2207676.2208695
93. Achille Pasqualotto and Michael J Proulx. 2012. The role of visual experience for the neural basis of spatial cognition. *Neuroscience & Biobehavioral Reviews* 36, 4: 1179–1187.
94. Stefania Pizza, Barry Brown, Donald McMillan, and Airi Lampinen. 2016. Smartwatch in vivo. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*: 5456–5469. http://doi.org/10.1145/2858036.2858522
95. John C. Platt. 1999. Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. *Advances in Large Margin Classifiers* 10, 3: 61–74.
96. Halley Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K. Kane. 2016. The AT Effect: How Disability Affects the Perceived Social Acceptability of Head-Mounted Display Use. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*: 4884–4895. http://doi.org/10.1145/2858036.2858130
97. Halley P Profita, James Clawson, Scott Gilliland, et al. 2013. Don'T Mind Me Touching My Wrist: A Case Study of Interacting with On-body Technology in Public. *Proceedings of the 2013 International Symposium on Wearable Computers*, ACM, 89–96. http://doi.org/10.1145/2493988.2494331
98. Shanaka Ransiri, Roshan Lalintha Peiris, Kian Peen Yeo, and Suranga Nanayakkara. 2013. SmartFinger: Connecting Devices, Objects and People Seamlessly. *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, ACM, 359–362. http://doi.org/10.1145/2541016.2541064
99. Siddharth S Rautaray and Anupam Agrawal. 2015. Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review* 43, 1: 1–54.
100. Zhou Ren, Junsong Yuan, Jingjing Meng, and Zhengyou Zhang. 2013. Robust part-based hand gesture recognition using kinect sensor. *IEEE transactions on multimedia* 15, 5: 1110–1120.
101. Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 887–896. http://doi.org/10.1145/1753326.1753458
102. André Rodrigues, Kyle Montague, Hugo Nicolau, and Tiago Guerreiro. 2015. Getting Smartphones to Talkback : Understanding the Smartphone Adoption Process of Blind Users. *ASSETS 2015: The 17th International ACM SIGACCESS Conference of Computers and Accessibility.*: 23–32. http://doi.org/10.1145/2700648.2809842
103. David A Ross. 2001. Implementing assistive technology on wearable

computers. *Intelligent Systems, IEEE* 16, 3: 47–53.
104. Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-defined Motion Gestures for Mobile Interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 197–206. http://doi.org/10.1145/1978942.1978971
105. Jaime Sánchez and Fernando Aguayo. 2007. Mobile messenger for the blind. *Universal Access in Ambient Intelligence Environments*: 369–385.
106. Munehiko Sato, Ivan Poupyrev, and Chris Harrison. 2012. Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 483–492. http://doi.org/10.1145/2207676.2207743
107. M Satyanarayanan. 2005. From the Editor in Chief Swiss Army Knife or Wallet ? 2–3.
108. Andrew Sears, Min Lin, Julie Jacko, and Yan Xiao. 2003. When computers fade: Pervasive computing and situationally induced impairments and disabilities. *HCI International*, 1298–1302.
109. Neil Selwyn, Stephen Gorard, John Furlong, and Louise Madden. 2003. Older adults' use of information and communications technology in everyday life. *Ageing and society* 23, 5: 561–582.
110. Roy Shilkrot, Jochen Huber, Wong Meng Ee, Pattie Maes, and Suranga Chandima Nanayakkara. 2015. FingerReader: A Wearable Device to Explore Printed Text on the Go. *Acm Chi'15*: 2363–2372. http://doi.org/10.1145/2702123.2702421
111. Kristen Shinohara and Jacob O Wobbrock. 2011. In the Shadow of Misperception: Assistive Technology Use and Social Interactions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 705–714. http://doi.org/10.1145/1978942.1979044
112. T. Starner, J. Auxier, D. Ashbrook, and M. Gandy. 2000. The gesture pendant: a self-illuminating, wearable, infrared\ncomputer vision system for home automation control and medical\nmonitoring. *Digest of Papers. Fourth International Symposium on Wearable Computers*: 87–94. http://doi.org/10.1109/ISWC.2000.888469
113. Lee Stearns, Ruofei Du, Uran Oh, et al. 2016. Evaluating Haptic and Auditory Directional Guidance to Assist Blind Persons in Reading Printed Text Using Finger-Mounted Cameras. *ACM Transactions on Accessible Computing* xx, x.
114. Lee Stearns, Uran Oh, Bridget J Cheng, et al. Localization of Skin Features on the Hand and Wrist from Small Image Patches.
115. Joseph C Stevens and Kenneth K Choo. 1996. Spatial acuity of the body surface over the life span. *Somatosensory & motor research* 13, 2: 153–166.
116. Joseph C Stevens, Emerson Foulke, and Matthew Q Patterson. 1996. Tactile acuity, aging, and braille reading in long-term blindness. *Journal of experimental psychology: applied* 2, 2: 91.
117. Emi Tamaki, Takashi Miyak, and Jun Rekimoto. 2010. BrainyHand: A Wearable Computing Device Without HMD and It's Interaction Techniques. *Proceedings of the International Conference on Advanced Visual Interfaces*, ACM, 387–388. http://doi.org/10.1145/1842993.1843070

118. Catherine Thinus-Blanc and Florence Gaunet. 1997. Representation of space in blind persons: vision as a spatial sense? *Psychological bulletin* 121, 1: 20.
119. Hsin-Ruey Tsai, Cheng-Yuan Wu, Lee-Ting Huang, and Yi-Ping Hung. 2016. ThumbRing: private interactions using one-handed thumb motion input on finger segments. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*: 791–798.
120. Gregg C Vanderheiden. 1996. Use of audio-haptic interface techniques to allow nonvisual access to touchscreen appliances. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 1266.
121. Gregg C Vanderheiden and J Bern Jordan. 2012. Design for people with functional limitations. *Handbook of Human Factors and Ergonomics, Fourth Edition*: 1407–1441.
122. Gregg Vanderheiden and Jim Tobias. 2000. Universal design of consumer products: current industry practice and perceptions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 6–19.
123. Francisco Vega--Bermudez and Kenneth O Johnson. 2001. Differences in spatial acuity between digits. *Neurology* 56, 10: 1389–1391.
124. Ramiro Velázquez. 2010. Wearable assistive devices for the blind. In *Wearable and Autonomous Biomedical Devices and Systems for Smart Environment*. Springer, 331–349.
125. Julie Wagner, Mathieu Nancel, Sean G. Gustafson, Stephane Huot, and Wendy E. Mackay. 2013. Body-centric design space for multi-surface interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, ACM Press, 1299.
126. Cheng-Yao Wang, Min-Chieh Hsiu, Po-Tsung Chiu, et al. 2015. PalmGesture: Using Palms as Gesture Interfaces for Eyes-free Input. *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15*, ACM Press, 217–226.
127. Martin Weigel, Tong Lu, Gilles Bailly, Antti Oulasvirta, Carmel Majidi, and Jürgen Steimle. 2015. iSkin: Flexible, Stretchable and Visually Customizable On-Body Touch Sensors for Mobile Computing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
128. Martin Weigel, Vikram Mehta, and Jürgen Steimle. 2014. More Than Touch: Understanding How People Use Skin As an Input Surface for Mobile Computing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 179–188. http://doi.org/10.1145/2556288.2557239
129. Samuel White, Hanjie Ji, and Jeffrey P. Bigham. 2010. EasySnap. *Adjunct proceedings of the 23nd annual ACM symposium on User interface software and technology - UIST '10*: 409. http://doi.org/10.1145/1866218.1866244
130. Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 143–146. http://doi.org/10.1145/1978942.1978963
131. Jacob O Wobbrock, Meredith Ringel Morris, and Andrew D Wilson. 2009. User-defined Gestures for Surface Computing. *Proceedings of the SIGCHI*

Conference on Human Factors in Computing Systems, ACM, 1083–1092. http://doi.org/10.1145/1518701.1518866

132. J O Wobbrock, S K Kane, K Z Gajos, S Harada, and J Froehlich. 2011. Ability-Based Design: Concept, Principles and Examples. *ACM Transactions on* 3, 3: 1–36. http://doi.org/10.1145/1952383.1952384
133. Jiahui Wu, Gang Pan, Daqing Zhang, Guande Qi, and Shijian Li. 2009. Gesture Recognition with a 3-D Accelerometer. *Ubiquitous intelligence and computing (UIC 2009)*, 25–38. http://doi.org/10.1007/978-3-642-02830-4_4
134. Xing-Dong Yang, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2012. Magic Finger: Always-available Input Through Finger Instrumentation. *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, ACM, 147–156. http://doi.org/10.1145/2380116.2380137
135. Hanlu Ye, Meethu Malu, Uran Oh, and Leah Findlater. 2014. Current and Future Mobile and Wearable Device Use by People with Visual Impairments. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 3123–3132. http://doi.org/10.1145/2556288.2557085
136. Yang Zhang, Junhan Zhou, Gierad Laput, and Chris Harrison. 2016. SkinTrack. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*: 1491–1503. http://doi.org/10.1145/2858036.2858082
137. *WebAIM. Screen Reader User Survey # 4 Results. 2012*.
138. 2014. Visual Impairments and Blindness. *World Health Organization*.
139. 2016. Consumer views on wearables beyond health and wellness. *Report, An Ericsson Consumer Insight Summary Report*, ERICSSON CONSUMERLAB.