**Title**
Designing Multimodal Alternatives for Nonvisual Computer Interaction



**Author**
Baldwin, Mark Steven

**Publication Date**
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

# Designing Multimodal Alternatives for Nonvisual Computer Interaction

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Informatics

by

Mark Steven Baldwin

Dissertation Committee:
Professor Gillian R. Hayes, Chair
Professor Jennifer Mankoff
Professor Bonnie Nardi
Assistant Professor Stacy Branham

2020

# DEDICATION

*For Colleen. My strength, my guide, my love, my wife. Without your support from the very moment this journey started and your patience when it felt like it would never end, I would not be here. Thank you.*

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGMENTS

First, I would like to thank the Makapo Aquatics Project for all of the ways they supported my research. From embracing my work with open arms to providing access to resources and facilities. I am forever grateful for not just your support, but also your friendship. And of course, teaching me how to stay upright in an OC1. To RJ De Rama, thank you for your hospitality, for promoting our work together, and tirelessly working to help bring our academic and non-profit worlds together. To Rob and the Skvarla family, this dissertation would not have been possible without you. Thank you for always saying yes.

To my advisors, Gillian Hayes and Jennifer Mankoff, it's hard to put into words how much your mentorship has meant to me over the years. Jen, thank you for taking me under your academic wing, for spending an entire year meeting with me while I applied to PhD programs, and for introducing me to Gillian. I will never forget your devotion to my success, a commitment that I will strive to emulate as a mentor for the rest of my career. Gillian, thank you for giving me the time and space to find my path. Somehow you always knew exactly what I needed. You cheered during my highs and held out a hand during my lows. I am a better person for having learned from you both.

To my dissertation committee, Bonnie Nardi and Stacy Branham, I deeply admire the knowledge you have both created as scholars in our field. I am humbled to have had you both as members of my committee. As I progressed from proposal to defense, recalling your enthusiasm and support for my research was often the nudge I needed to persevere.

I can't imagine what this journey would have been like without my incredible academic family. The STAR group is filled with an amazing cast of past and future scholars, all of whom I am honored to call academic siblings. Jed Brubaker and Lynn Dombrowski, we only spent one short year in the lab together, but it has felt like you were with me the entire time. Oliver Haimson and Kate Ringland, thank you for teaching me how to be a more measured and thoughtful scholar. Sen Hirano, you are forever my sounding board. Thank you for finding the edge cases and bringing them into view. LouAnne Boyd, the fervor with which you pursue your goals is truly inspiring. Thank you for the countless hours of feedback and encouragement. Jazette, Elizabeth, Lucretia, and Francelli, thank you for putting up with my often erratic behavior as I juggled job search and dissertation.

To my family, thank you for instilling me with the core values that drive my work today. To my wife Colleen, this dissertation is not possible without your unwavering support. Your hard work gave me the foundation to pursue a PhD and all that it encompasses. Thank you. Mom and Dad, you taught me how to build, you let me break things and put them back together, and you let me learn at my own pace. I am not who I am today without these experiences. Christine, Lauren, Jill, and Kevin, thank you for setting a high bar for academic pursuit in our family. To Todd and Laura, you shaped the direction of my life in indescribable ways, I will always cherish your excitement and encouragement from comprehensive exam to dissertation defense.

Finally, I would like to thank Robert and Barbara Kleist, UC Irvine Grad Division, the Department of Informatics, and Billy Whitford and the Newport Aquatic Center for their support in this journey.

# CURRICULUM VITAE

## Mark Steven Baldwin

# EDUCATION

*Ph.D.*, Informatics 2014-2020
University of California, Irvine
Advisors: Professor Gillian R. Hayes, Professor Jennifer Mankoff
THESIS - Augmenting audio-only assistive technologies with additional modalities can reduce dependence on audio and improve the experience of activities for blind and low vision users of assistive technologies. *Master of Human Computer Interaction*, Human Computer Interaction Institute August 2013
Carnegie Mellon University, Pittsburgh, PA

*Bachelor of Arts in Computing*, College of Computing and Digital Media July 2012
DePaul University, Chicago, IL

# PUBLICATIONS

**Baldwin, M. S.**, Mankoff, J., Nardi, B., Hayes, G. (2020) ‘An Activity Centered Approach to Nonvisual Computer Interaction". ACM Transactions on Computer-Human Interaction (TOCHI), 27(2), 1-27.

**Baldwin, M. S.**, Hirano, S. H., Mankoff, J., Hayes, G. R. (2019). ‘Design in the Public Square: Supporting Cooperative Assistive Technology Design Through Public Mixed-Ability Collaboration." Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 155.

**Mark Baldwin**, Sen Hirano, RJ DeRama, Jennifer Mankoff and Gillian Hayes (2019). ‘Blind Navigation on the Water through Shared Assistive Technology." Workshop: Hacking Blind Navigation. ACM SIGCHI Conference on Human Factors in Computing Systems. Glasgow, Scotland.

**Baldwin, M.**, Khurana, R., McIsaac, D., Sun, Y., Tran, T., Zhang, X., Fogarty, J., Hayes, G.R., and Mankoff, J. (2019) Tangible Interfaces. In Harper, S., & Yesilada, Y. (Eds.) *Web Accessibility: A Foundation for Research.* Springer Science & Business Media.

**Mark Baldwin**, LouAnne Boyd, Jennifer Mankoff and Gillian Hayes (2018). ‘Sensory Inclusive Design for Voice Interfaces." Workshop: Accessible Voice Interfaces. ACM SIGCHI Conference on Computer Supported Cooperative Work. Jersey City, NJ.

**Mark Baldwin** (2018). ‘Activity Theory as a Framework for Nonvisual Computing." HCIC Pajaro Dunes Resort, Watsonville, CA.

**Mark Baldwin** (2017). ‘Beyond Audition: Tangible Alternatives for Nonvisual Computer Interaction." Doctoral Consortium at ACM International Joint Conference on Pervasive and Ubiquitous Computing. Maui, HI.

**Baldwin, M. S.**, Mankoff, J., Hayes, G. R., Haimson, O., Hudson, S. (2017). ‘The Tangible Desktop: A Multimodal Approach to Nonvisual Computing." ACM Transactions on Accessible Computing (TACCESS), 10(3), 9.

**Mark Baldwin**, Jennifer Mankoff, Gillian Hayes, Scott Hudson and Jeff Bigham (2016). ‘Reappropriating Desktop Computing Metaphors for Nonvisual Tactile Interaction." Workshop: Touch, Taste, & Smell User Interfaces: The Future of Multisensory HCI. ACM SIGCHI Conference on Human Factors in Computing Systems. San Jose, CA

Ringland, K. E., Wolf, C. T., Boyd, L. E., **Baldwin, M. S.**, and Hayes, G. R. (2016). ‘Would you be mine: Appropriating minecraft as an assistive technology for youth with autism." In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (pp. 33-41). ACM. ***Best Paper**.

Raymond Liaw, Ari Zilnik, **Mark Baldwin**, and Stephanie Butler. (2013, May 1). ‘Maater: Crowdsourcing to Improve Online Journalism". Student Design Competition. ACM SIGCHI Conference on Human Factors in Computing Systems. Paris, France.

# ABSTRACT OF THE DISSERTATION

Designing Multimodal Alternatives for Nonvisual Computer Interaction

By

Mark Steven Baldwin

Doctor of Philosophy in Informatics

University of California, Irvine, 2020

Professor Gillian R. Hayes, Chair

Natural and artificial acoustics of the built world provide spatial awareness, environmental context, and communicative abilities to people with visual impairments. When the acoustic soundscape is insufficient, assistive technologies – devices and computational systems designed to augment the visual modality – can be used to translate visual information to other modalities. In the field of Human-Computer Interaction (HCI), nonvisual assistive technologies seek to bridge the human to computer relationship through speech and non-speech sounds, touch, and by facilitating interactions through sighted support structures. Current research exploring technological approaches to solving nonvisual challenges in a sight-first world, often emphasize the transfer of visual information to a single modality. In this dissertation, I examine how individual modalities can be unified through novel computational interactions to alleviate sensory overload, and support a broad set of activities. I ask the questions: 1) How can tangible and mixed ability computational systems be designed to reduce audio and interactions for blind and low vision people in everyday activities, and 2) Do assistive technologies that are less reliant on auditory input improve blind and low vision user interactions? Through field studies of blind and low vision computers users, I designed and evaluated a platform for augmenting auditory computer interactions with a tangible interface. An experimental evaluation of the combined auditory and tangible control found a thirty-nine percent increase compared to traditional audio-only tools in web navigation tasks. Through a multi-year ethnographic study of a blind and low vision outrigger canoeing community, I examined the intersection of sensory modalities and mixed-ability relationships during canoeing activities. Applying a public-facing co-design participatory methodology, I worked alongside blind and sighted outrigger canoe enthusiasts to design, evaluate, and deploy a shared assistive technology to support blind paddling. Analysis of my work reveals how physicality in the world influences the auditory and tangible interactions of assistive technologies. In addition to my empirical findings, through this work I demonstrate how attending to context and the physicality of sound and touch reveal critical insights to guide the design of assistive technology.

# Chapter 1

# Introduction

> What is the world of sound? I have been spending some time out of doors trying to respond to the special nature of the acoustic world. I am impressed by the many different aspects of reality, the range and depth of the contact points between myself and something created by sound. ... The intermittent nature of the acoustic world is one of its most striking features...I have only very limited power over the acoustic world. ... This is a world which I cannot shut out, which goes on all around me, and which gets on with its own life. ... Acoustic space is a world of revelation.
>
> – *Touching the Rock, John M. Hull*

This quote, adapted from John Hull's three year account of his mid-life transition to the world of blindness (Hull, 1992), summarizes his reflection on sound and the acoustic world. In the absence of sight, Hull finds an unexpected richness in the way sound informs the world around him. For Hull, sound is an "astonishingly varied and rich panorama of movement, music and information...absorbing and fascinating." Yet, the qualities of sound that Hull reveres, often go unappreciated, misunderstood, and underused by assistive technologies designed to support blindness. Hull's account captures the underlying motivation for my research, and this dissertation: to understand how, in absence of sight, the unique characteristics of the human sensory system can be better supported by assistive technology.

As the primary sensory modality used by people with visual impairment (Hötting & Röder, 2009), the auditory channel is responsible for interpreting and understanding the visual world. Speech and non-speech sounds, processed by the auditory channel, provide spatial awareness, environmental context, and communicative abilities in place of visual information. Sounds are rendered, naturally and artificially, through the day to day encounters of human activity. For the visually impaired community, when naturally occurring sounds are unable to sufficiently supplant visual information, assistive technologies can be introduced to augment information through additional modalities. The white cane, for example, translates the textures and obstacles of the spaces surrounding a person into tangible and audible information that can be interpreted, processed, and acted upon to safely navigate complex environments. The strength of the auditory modality, as well as the relative ease through which audio is produced in technological systems, has led to an abundance of assistive technologies designed to transfer visual information to the auditory channel.

Although there is some evidence that people with visual impairment have increased auditory perception when compared to sighted people (Niemeyer & Starlinger, 1981; Lessard, Paré, Lepore, & Lassonde, 1998; Hugdahl et al., 2004), the auditory channel still requires cognitive effort to process information. Like most human perception channels, over-stimulation of the auditory channel can lead to cognitive overload (C. L. Baldwin, 2016), a side effect recognized by some assistive technology scholars (Dakopoulos & Bourbakis, 2009; Rector, Bennett, & Kientz, 2013). Others have explored ways to make auditory information processing more efficient. Rate of speech (Walker, Nance, & Lindsay, 2006), spatial sounds

(Savidis, Stephanidis, Korte, Crispien, & Fellbaum, 1996), and concurrency (Guerreiro & Gonçalves, 2014) have been shown to be effective strategies for increasing the density of auditory information, potentially decreasing processing time. However, the ephemerality of sound, sequential structure of auditory information, and concerns over disruption and privacy (M. S. Baldwin, Hayes, Haimson, Mankoff, & Hudson, 2017), as well as the broader auditory landscape (LaBelle, 2010) are seldom factored into assistive technology design.

In this dissertation, I explore how augmenting and reducing auditory interaction for assistive technology can expand the types of experiences available to the blind and low-vision community. I explore the relationship between people with visual impairment and assistive technology across two domains that each pose different challenges to the use of auditory based technologies: computer-based knowledge work and water-based leisure activities. The results of this work indicate that in a work environment, traditional auditory-focused assistive technologies are difficult to learn, inefficient, and disruptive. Similarly, in outdoor leisure activities, the results of this work indicate that natural environmental auditory cues are equally, if not more, important than the information that an assistive technology might provide to the enjoyment and utility of the experience. Taken together, these edge cases provide a way in which to think about and examine how auditory information might be reduced and augmented to improve experiences of assistive technologies more broadly.

## 1.1 Audio and Computer-based Knowledge Work

Desktop and mobile computing systems have traditionally relied on audio output as the primary interface for blind and low vision users. Using only audio output provides simplicity of implementation and relative ease in which interactions can be learned. In practice, however, this model fails to match the quality of sighted user experiences. Where sighted individuals benefit from the parallel processing provided by vision, blind and low vision users must process the same information sequentially through a single auditory modality. It is through this translation from visual to auditory that valuable information (*e.g.,* spatial arrangement and text emphasis) is lost. Additionally, the ephemeral nature of audio prevents information from persisting beyond the moment it is delivered.

Despite these challenges, blind and low vision users still engage with screen readers to serve a variety of needs, including basic home and work productivity as well as emergent forms of technical engagement. Increasingly, just as with sighted users, blind and low vision users have become dependent on some form of computer interaction for many aspects of everyday life. These uses can be particularly challenging for nonvisual computer users, because they often include images and interactive elements that do not translate well to audio output. Despite legal requirements and general support for accessibility online (Administration, 2020; Webaim.org, 2020), a variety of challenges still exist for users who must rely solely on audio output, including system inefficiency, poor user experience, and cognitive overload.

Converting visual information to auditory output through text-to-speech engines has long been the most common method for nonvisual computer interaction. Affordable computers with the ability to produce audio output are readily available, standards are in place (W3C, 2020), and there is already significant infrastructure to teach people to use screen readers through community and school-based programs. On the software side, a developer only needs to ensure that visual elements are semantically defined to make them accessible to screen readers. In practice, this is a relatively straightforward task. Thus, it is perhaps not surprising that advances in the communication of information visually, continue to be redirected as audio output for blind and low-vision users.

Even if screen readers were optimal translations of the computer experiences at the time they were developed (a time of command line interaction), the changes we have witnessed in user experiences over the last few decades have destroyed any possibility of these to be comparable interfaces today (E. Mynatt, 1997). Those dependent on screen readers are simply not benefiting from the massive advances in user experience, particularly the two-dimensional direct manipulation of the graphical user interface (GUI). Embedded within the GUI are visual and spatial cues that allow sighted users to perceive multiple actions simultaneously. These cues enable sighted users to rely on recognition rather than recall (Nielsen, 1994), providing a significant advantage over those using screen readers.

In chapter 3, I describe the ways in which screen readers are problematic, particularly for novice users, and probe how we might design improved interfaces for non-visual access. My approach takes inspiration from the way that early graphical interfaces appropriated physical metaphors with which users were already familiar, and wondered if these same metaphors might be usefully re-appropriated into a nonvisual interactive system. To answer this question, I conducted a field study of screen reader use at a school for people who are losing or have recently lost their sight. Based on the results of my field work, I developed a prototype for multimodal nonvisual computing that moves semantic information out of the audio stream and onto a tangible interface. An experimental study using this tangible interface revealed ways in which such systems might actually be easier and more efficient than screen readers, particularly for novice users. It also highlighted the complexity of introducing non-visual modalities to the interaction space of modern desktop operating systems. To deepen our scholarly understanding of these complexities, I conducted further analysis on my field work and experimental study results using an activity theory lens. I observed how these challenges exist within the fundamental structures of computation, traditionally hierarchical, and now increasingly oriented towards search and sensor streams. Rather than accept these technological orientations, a computational system structured around *the unit of human activity* (Nardi, 1996a) holds promise as a way to remove many of the obstacles blind and low vision users face.

The mismatch between activity and computational infrastructures for blind and low vision computer users can be extreme. While visual computing users take advantage of large

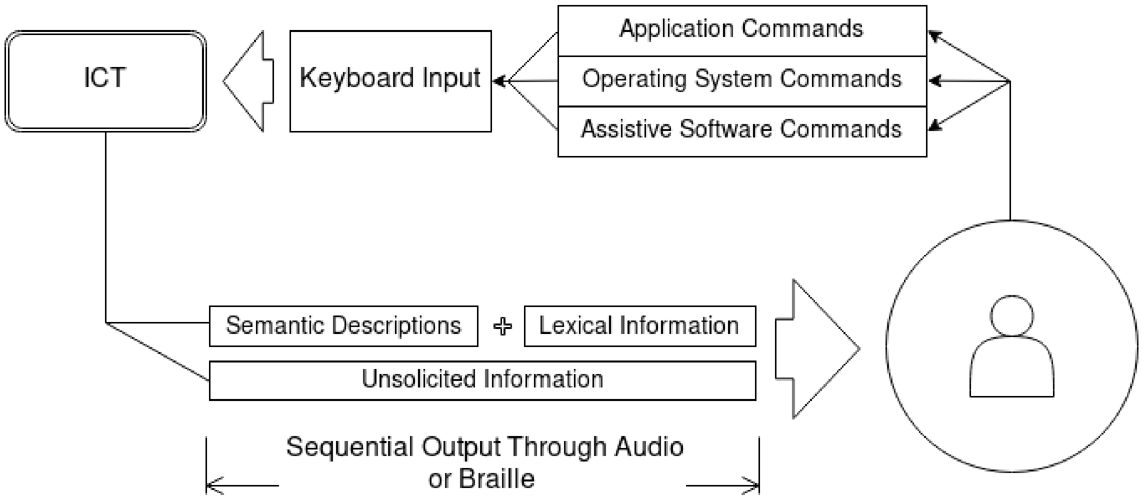



Figure 1.1: The typical workflow for nonvisual use of computational systems. A blind or low vision user must recall application, operating system, and assistive software keyboard commands to control input into an Information and Communications Technology (ICT). The ICT responds to keyboard input through a serial stream of auditory information that combines semantic and lexical information.

screens, multiple monitors, and multiple desktops to orient themselves visually to swap quickly among a variety of contexts and activities, blind and low vision users must rely on a keyboard for input and specialized software for converting visual and textual information to speech or a magnified viewport(see Figure 1.1). This approach widens the gap for blind and low-vision users between the human activity and the technological abstraction, as all of the services and data typically communicated visually (*e.g.,* progress bars, notifications, and spatial information), must also be put into the audio channel. With each technological advancement in visual communication, new efforts are required to translate information into a nonvisual medium.

Thus, I focused on how the scholarly community might reconsider accessible technologies for the desktop computing environment from the ground up as an infrastructure centered around activity. In this way, we can narrow the gaps for blind and low vision users between the goals they wish to accomplish, the tasks they are attempting to complete, and the computing infrastructure they must use. I see two distinct advantages to this approach beyond the obvious improvement of solutions for blind and low vision users. First, by simplifying abstractions to make them accessible, we may reduce the kind of overhead that is limiting adoption of an activity-centered approach in the broader community. Second, by re-articulating blind and low-vision user needs and preferences through the lens of activity theory, I contribute a conceptual framework for thinking about accessibility and assistive technologies.

## 1.2 Audio and Water-Based Leisure Activity

Disabilities are experienced through interactions with the physical and social world, and while social interactions and disability experiences can be shared, these interactions are more often othering and problematic. One might then ask what the role of shared assistive technology can be in changing this dynamic. I explored this issue by publicly developing a technology for blind outrigger canoe paddlers. I tested the boundaries of cooperative design as well as community-based research. By intentionally co-designing a shared, mixed-ability assistive technology in public settings, I was able to explore the ways in which sighted and blind technology users working together created an improved experience that neither could create alone, thus providing a roadmap for potential future cooperative assistive technology design and development.

Technology can be a powerful enabling force through which blind and visually impaired individuals experience parts of the sighted world that might otherwise not be available to them. These assistive technologies help people manage a variety of everyday tasks at home, at work, and during leisure. Assistive technologies are more often viewed as individual tools to support an individual experience (Ripat & Woodgate, 2011) than shared tools (Azenkot, Feng, & Cakmak, 2016; Bigham et al., 2010; Lacey & MacNamara, 2000). However, just as students in a class who know about a child's disability are likely to be more welcoming (Ochs, Kremer-Sadlik, Solomon, & Sirota, 2001) and how people view wearable technology more favorably when they know it is assistive (Nguyen et al., 2009; H. Profita, Albaghli, Findlater, Jaeger, & Kane, 2016), cooperation between people with and without disabilities can tremendously improve not only technology design but also experiences for both involved (Shinohara, Bennett, Pratt, & Wobbrock, 2018).

In chapter 4, I describe how I set out to understand the experience of blind and low vision outrigger canoe paddlers as a powerful example of access to leisure activities regardless of physical constraints. Working with an organization that takes blind and low vision paddlers on group outings in a six-person outrigger canoe (OC6) for both recreation and competition, I learned about the need for solo paddling. Competitions typically involve the use of an OC6 with a sighted steerperson, allowing up to five visually impaired persons to participate. To prepare for competitions, however, training in an one-person outrigger canoe (OC1) is critical to developing strength and technique. As such, we focused our investigation on supporting blind and low vision paddlers using an OC1. For example, visually impaired paddlers typically either rely on a multi-person canoe or respond to simple audible directives (*e.g.,* left, right, watch out) called out from a support boat—neither of which match the training conditions of sighted paddlers. Furthermore, the invasive stream of commands required to verbally navigate a blind paddler, are often viewed as an unwelcome disruption to the natural soundscape of the activity. In addressing these challenges, I adopted an interdependent approach to assistive technology design (Bennett, Brady, & Branham, 2018), with an eye towards the ways that working together on both the design and the experience of such technology could drive additional awareness and engagement across the boundaries of the sighted and blind paddler communities. Following a multi-month co-design experience with both sighted and blind paddlers and several rounds of prototype iterations and development, I conducted a multi-year ethnographic study of multi-modal multi-user system with blind and sighted paddlers.

## 1.3 Thesis Statement

Augmenting audio-only assistive technologies with additional modalities can reduce dependence on audio and improve the experience of activities for blind and low vision users of assistive technologies.

## 1.4 Research Questions

RQ 1. How can tangible and mixed ability computational systems be designed to reduce audio and interactions for blind and low vision people in everyday activities?

RQ 2. Do assistive technologies that are less reliant on auditory input improve blind and low vision user interactions?

## 1.5 Contribution of Dissertation

The work presented in this dissertation contributes to the literature on assistive technology for blind and low vision individuals. Little existing literature has explored how assistive technologies can reduce auditory information for blind and low vision users. My work draws attention to tangible and mixed-ability interaction as potential solutions to both the reduction and augmentation of audio in assistive technology interaction. By broadening the scope of approaches for specific nonvisual assistive technologies, I provide new ways of thinking about how assistive technology is situated within the broader context of mainstream technology and its dependence on audio. My work also contributes to the literature on technologies in support of both work and leisure activities for blind and low vision individuals.

Finally, this dissertation explicitly engages Activity Theory and activity-centric computing as frameworks for the design of assistive technologies. These approaches, when deployed in this work, led explicitly to non-audio-based engagements: tangible interfaces in the case of the workplace and mixed-ability shared control collaborative technologies in the case of water-based leisure activities. By exposing the activity as the key unit of analysis for this work, I demonstrate specific ways in which audio is problematic as well as a range of creative solutions for addressing those challenges.

## 1.6 Summary of Chapters

In chapter 2, I start with an overview of Activity Theory, the theoretical framing that I have applied to much of the work I have completed for this dissertation. I then conduct review of the literature, establishing a foundational background for my work.

The results for the work presented in this dissertation are organized into two chapters: 3 and 4. In chapter 3, I review the details of the field work I conducted at a computer training school for visually impaired adults. I detail the design, development, and evaluation of the Tangible Desktop, a desktop computer peripheral that changes how audio is presented to blind and low vision computer users. I also detail the development and evaluation of KInD, an activity centered tangible computing peripheral.

In chapter 4, I provide an overview of the multi-year fieldwork I completed with a blind and low vision outrigger paddling organization. I describe how alongside this organization, I designed a shared assistive technology to support solo canoe operation for blind paddling. I then report the results of my findings from observing twenty-four training sessions using this shared technology.

In chapter 5, I reflect on the results across all of my work. I discuss in detail I view this work contributing to scholarly work in the assistive technology design space.

# Chapter 2

# Background

Assistive tools seek to break down barriers in social, environmental, and technological activities that affect how people with visual impairments negotiate the world. Tools like the white cane, smartphone navigation, and text-to-speech systems, mediate actions to fulfill goals of the individual that they might not otherwise be able to complete. One perspective on the significance between tool mediation and human activity emerged from the work of early twentieth century Russian psychologist L.S. Vygotsky. Vygotsky viewed tool use as a distinguishing feature of human activity, one that enables humans to "master" and "triumph over nature" (Vygotsky, 1980). Vygotsky viewed the mediation of human activity through tool use as a connection to both objects of, and people in, an environment (Leont'ev, 1974). Due to the importance of tool use by people with visual impairments, Vygotsky's perspective provides a valuable lens through which we might identify characteristics and qualities for the design of assistive technologies.

In a world in which nonvisual access to vital systems and services is often calculated retroactively, Vygotsky's work brings to the forefront the importance of mediating tools as "conductor of human influence on the object of activity" (Vygotsky, 1980). Not simply for the types of individual activities that tools can support, but also for the “social intercourse" which they engender (Leont’ev, 1974). The types of human activity that emerge from individual physchological processes, can only occur through the reciprocal interaction of people (Vygotsky, 1980; Leont’ev, 1974). The white cane, for example, carries with it a sociohistorical context (derived from this reciprocal interaction) that translates to interphsychological behavior and expectations for individuals in proximity to its holder. A blind traveler aware of this behavior, therefore, is free to choose the types of activities for which the white cane best serves.

It is through the “orienting concepts and perspectives" of Vygotsky’s notion of mediated action (Nardi, 1996a; Engeström, 2009) and social consciousness, that I situate the broader scope of my work. Assistive technologies (*e.g.,* mediating tools) can and should be viewed not just as enabling tools, but also as artifacts that promote social inclusiveness and awareness. Thus, the core principles of Vygotsky’s work provide a useful language upon which the design and implementation of assistive tools can be framed.

As I use the language of activity throughout this dissertation, in the next section I offer an overview of the history and core principles of activity theory, a formal structural representation derived from Vygotsky’s work. I then explore literature related to the mediating role that assistive technology plays in the everyday lives of people with visual impairments. Finally, I conclude this chapter by reviewing how tool mediation has expanded beyond the one-to-one relationship between tool and the individual to include multiple individuals, often of varying abilities, to broaden the types of activities that people with visual impairments are able to carry out.

## 2.1 Overview of Activity Theory

Vygotsky's work on the relationship between mediating tools and human activity was refined by his colleague and former student A. N. Leont'ev, who classified the central concepts and terms that describe the general structure of activity theory (Leont'ev, 1974). Leont'ev sought to solidify Vygotsky's concepts, which he argued were far too abstract, into a more concrete, and therefore meaningful, form. One of Leont'ev's most significant contributions to the refinement of how the theoretical analysis of human activity is carried out is the concept of the object. According to Leont'ev, objects motivate activities, giving them purpose and direction. It is through the fundamental characteristic of the object that Vygotsky's early concepts of tools, goals, and operations organize into the formal structure (which I will expand upon in the next section) that has taken root in the field of Human-Computer Interaction as well as my own work on nonvisual computational systems.

Activity theory's conceptual framework for understanding the goals, motives, and needs of human consciousness is represented by the relationship between the human *subject* and an *object*, traditionally denoted symbolically as $S \rightarrow O$. The needs of the subject motivate actions towards an object, influenced by the attributes of both (Leont'ev, 1974). An object represents some thing, material or conceptual, that requires change. The actions that humans carry out upon the object are predominately driven by mediating artifacts and motivated by goals (Leont'ev, 1978). The role of the mediating artifact is to support the work of the subject as it changes the object. The classic paradigm frequently used in activity theory literature is the carpenter (subject) who uses a hammer (artifact) to hit a nail (object) (Kaptelinin & Nardi, 2006; Nardi, 1996a). The nail is transformed through the action of being hit which is carried out in service of some goal or goals required to complete the activity; for example, to practice hammering or repair a fence.

The relationship between subject, object, and mediating artifact is typically represented by

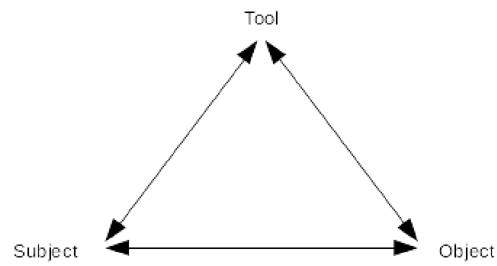



Figure 2.1: A diagram depicting the classical representation of the relationship between the subject, object, and mediating artifact (tool) in activity theory.

the model in figure 2.1.

Activities are naturally hierarchical; the completion of one activity might depend on additional related activities to be completed first. As an activity is carried out, a coordinated process of learning and acting moves the activity towards completion. Leont'ev organized this process into a three level hierarchy of activity, action, and operation (Leont'ev, 1978). Activities are carried out through actions, determined by the individual goals of the human subject. Actions are fulfilled through unconscious operations, which reflect the human subject's natural attributes. The relationship between actions and operations is described as fluid, where actions can become unconscious operations through the natural internalization that occurs through practice, and operations become conscious actions through externalizing processes, such as breakdowns (Bødker, 1991).

The application of the principles of activity theory to technological systems was a critical step in the post-cognitivist movement within human computer interaction (HCI) (Kaptelinin & Nardi, 2006). As technology became part of our everyday lives, so did its role as a mediating tool. For example, in her seminal work, Bødker established the role of the computer as mediating artifact: a user (subject) types words into a computer (mediation) to send an email (object) (Bødker, 1991). The technology as mediator paradigm has formed the basis for a significant body of literature focused on bridging gaps between human activity and computation (Clemmensen, Kaptelinin, & Nardi, 2016; Bødker & Andersen, 2005). Of particular interest to my work are the efforts focused on re-framing desktop interaction from an application-centric model to one centered on activity (Bardram, Bunde-Pedersen, & Soegaard, n.d.; Kaptelinin, 2003; Voida, Mynatt, & MacIntyre, 2007). Despite promising results, the application-centric model remains largely a siloed experience, where applications and the computational entities that they generate have little awareness of each other, leaving the user to develop individual strategies for coordination (Kaptelinin & Nardi, 2006).

One exception can be found in applications that encapsulate individual tools into a common environment. Personal information managers (PIM) such as Microsoft Outlook, are one example of a single application that operates at the level of activity (Kaptelinin & Nardi, 2006). PIMs typically coordinate messaging, scheduling, and social contacts together around a single goal, a fundamental requirement of activity. Although the PIM presents a compelling example for the integration of activity principles in computing, there is little evidence of activity centered systems being taken up in a broad and comprehensive sense. Furthermore, this is no evidence of the application of activity theory principles to the design of assistive technologies for the visually impaired community. When considering the challenges faced by users of assistive technologies (Lazar, Feng, & Allen, 2006; Lazar, Allen, Kleinman, & Malarkey, 2007; Mankoff, Fait, & Tran, 2005), it is clear more work is needed to improve their experience.

Activity theory helps us to see the ways in which the human activity and the system objects are co-constructed, updating and evolving alongside one another. The reliance that the system has on people being sighted to make these intertwined interactions work relatively seamlessly. As Nardi states, "Activity theory says, in essence, that we are what we do." Activity-centered computing has been previously proposed as a solution to numerous mismatches between user activities and system behavior and as a more general approach for social analysis and design in HCI (Kaptelinin, Nardi, & Macaulay, 1999). This work reveals how important an activity theoretical approach could be for nonvisual computer users, it draws attention to the possible benefits of applying activity theoretical concepts to assistive technologies beyond the desktop computing paradigm. Through an activity theoretical lens, we can look beyond the individual application oriented tasks of each computational system and think about how assistive systems can serve activities more broadly. This paradigm shift is an opportunity for designers and researchers to re-imagine how computation can support the activities of blind and low vision users.

## 2.2 Mediating Action Through Assistive Tools

Mediating tools designed specifically for supporting computational interaction for blind and low vision people are largely concerned with augmenting the visual channel through alternative sensory modalities. Historically, scholarly interest in nonvisual interaction with computational systems has involved novel ways to transfer information from the visual channel to the auditory channel (E. D. Mynatt & Weber, 1994; Raman, 1996; Plessers et al., 2005; Donker, Klante, & Gorny, 2002; Saha, Fiannaca, Kneisel, Cutrell, & Morris, 2019). While the recent emergence of affordable rapid-prototyping techniques have made experimental work with tangible interfaces more common (*e.g.*, (Bau, Poupyrev, Israr, & Harrison, 2010; Gupta, Morris, Patel, & Tan, 2013; Jansen, Karrer, & Borchers, 2010; Sodhi, Poupyrev, Glisson, & Israr, 2013; Weiss, Voelker, Borchers, & Wacharamanotham, 2011; Weiss et al., 2009)). In this section, I present literature in the areas of auditory and tangible interaction that has sought to expand the ways people with visual impairements engage with technology.

### 2.2.1 Application of Auditory Forms of Computer Interaction

The use of the auditory channel for the communication of information to people with visual impairments is common across assistive technology domains, from desktop computing systems to computational orientation and mobility aids (Manduchi & Kurniawan, 2018; Mulloy, Gevarter, Hopkins, Sutherland, & Ramdoss, 2014). The wide range of advances found in audio representations can be accounted for in part by the ease of prototyping audio, which is primarily a software task, despite known drawbacks (Peres et al., 2008). In desktop computing, audio requires users to make a trade-off between metaphors that are fairly closely tied to the hierarchy of graphical interfaces (*e.g.,* (Edwards, Mynatt, & Stockton, 1994)), or metaphors that are so different as to require advanced expertise and learning (*e.g.*, audio icons and contrasting voices (Raman, 1996)). Although it is commonly understood that visual impairment can lead to more efficient auditory perception (Röder, Rösler, & Neville, 2001), resolving accessibility issues through a single modality can still lead to unbalanced, complex engagements that are challenging for experts and non-experts.

The large corpus of input commands required to operate audio interfaces, for example, impose significant hurdles as users first familiarize themselves, only to become a constraint on efficiency with increasing expertise (Vigo & Harper, 2014).

Early efforts in accessibility research focused on the development of sub-systems that understood how to pair metadata with interface elements (Edwards et al., 1994; E. Mynatt, 1997). As sub-systems became natural components of user interface toolkits, attention turned to making sure that the metadata was adequately communicated. The ubiquity of 16bit computer audio, for example, saw significant research efforts in the auditory presentation of data including audio manipulation techniques like sonification and 3D audio (Gaver, 1989; Blattner, 1989; Crispien, Fellbaum, Savidis, & Stephanidis, 1996; Donker et al., 2002; Walker et al., 2006). The positive results of many of these research efforts is likely due to the capabilities of the human auditory system. Blind people are known to have significantly stronger verbal auditory encoding abilities (Lessard et al., 1998) as well as a stronger ability to identify individual sounds from concurrent speech (Guerreiro & Gonçalves, 2014) when compared to sighted individuals. Yet, few of these systems have gained traction in commercial audio-based interaction tools like screen readers. The serial, ephemeral nature of audio does not adequately translate the metaphor and expressiveness of visual information, prompting visually impaired users to search for complimentary support (Bigham et al., 2010; Lacey & MacNamara, 2000).

Assistive technologies that use auditory speech interfaces rely mainly on text-to-speech processors which take text as input and return digitized speech as output. During the conversion process, visual information such as graphical emphasis and spatial arrangement is lost (Asakawa, Takagi, Ino, & Ifukube, 2002). Work has been done to overcome this by augmenting the output with a navigation ontology (Yesilada, Harper, Goble, & Stevens, 2004), or increasing access to semantic information (Plessers et al., 2005). Additionally, speech input, non-speech input (Igarashi, Moscovich, & Hughes, 2005) and non-speech output (*e.g.,* (E. D. Mynatt & Weber, 1994; Savidis et al., 1996; Donker et al., 2002)) have been shown to enhance the interactive experience and the range of things that can be displayed. For example, sounds can be positioned in space, as with Schmandt's hallway metaphor, which used head motion to focus attention on "sound bites" passed to either the subject's left or right ear (Schmandt, 1998). Novel computational systems that make use of audio, either through sonification or speech and non-speech sounds, target the introduction or reproduction of visual information to the auditory channel. For example, graphical computer interfaces simplify interaction for sighted users using physical and visual metaphors supporting direct manipulation (*e.g.,* files and folders), but the digital objects that support these metaphors do not translate well to nonvisual use, making it dramatically more difficult compared to visual use (Borodin, Bigham, Dausch, & Ramakrishnan, 2010). Furthermore, graphical interfaces can support unconscious operation (see Section 2.1) (Leont'ev, 1978) by providing memory aids through information that can be directly perceived in a consistent and repetitive structure (Bødker, 1990), an advantage that is largely absent in audio-based interfaces. Vázquez and Steinfeld observed no preference between speech and non-speech sounds among blind photographers, suggesting that in a constrained, activity-oriented application, lightweight repetition of auditory information is beneficial (Vázquez & Steinfeld, 2012). Yet, excessive audio is known to overwhelm the auditory channel (C. L. Baldwin, 2016), regardless of ability, suggesting that audio alone might not be sufficient for communicating information nonvisually.

Avoiding auditory sensory overload is frequently signaled by researchers within the assistive technology literature as a motivating factor for non-auditory modality research (Guerreiro et al., 2015; Günther et al., 2018) or as an emergent theme from audio rich interventions (Ducasse, Macé, Serrano, & Jouffrais, 2016; Wong, Yap, Alexander, & Karnik, 2015). In a survey of electronic travel aids for the blind, Dakopoulos and Bourbakis urge designers to avoid interfering with an individual's ability to listen to environmental sounds, despite the enhanced hearing capability of people who are blind, *e.g.*, (Hugdahl et al., 2004). Similarly, Saha et al. remark how participants desired less auditory information to avoid cognitive overload while walking, stressing the importance of task-based information filtering. Yet auditory challenges (Peres et al., 2008), which are often mentioned without consideration for the individual needs of the target audience (*e.g.,* (McDaniel et al., 2014; Günther et al., 2018; Ducasse et al., 2016)), are rarely explored in depth. To the extent that challenges of audio overuse in nonvisual tools have been explored, results are narrowly scoped and not derived from the visually impaired community. For example, McGookin, Robertson, and Brewster suggest a "division of functionality" in their Tangible Graph Builder that assigns tangible and auditory interactions according to guidelines for tangible interfaces derived from a sighted study by Challis and Edwards (McGookin et al., 2010). Although McGookin et al. note that participants found the use of audio appropriate, the authors do not expand on this finding enough to be applicable more broadly. In contrast, the only robust evaluations of audio use identified in my review of the literature, Rector, Milne, Ladner, Friedman, and Kientz conduct a mixed-methods analysis of blind and low-vision experiences with exercise technology (Rector et al., 2015). The authors explicitly encourage designers to consider how auditory information is delivered in assistive technologies, drawing attention to the varying roles that audition plays across different activities. Aside from Rector et al.'s review, a dearth of literature eliciting the ways in which visually impaired subjects are affected by auditory information draws attention to the importance of the work I present in this dissertation.

### 2.2.2 Application of Tangible Forms of Computer Interaction

One way to broaden the set of metaphors available to nonvisual interaction is to move out of an audio-only realm and into the physical world of tactile interaction. The sense of touch is a rich input modality capable of perceiving a diverse range of feedback from size and shape, to texture, stiffness, and temperature (Klatzky & Lederman, 2003). Interacting with physical objects is known to be an enjoyable experience (Zuckerman & Gal-Oz, 2013; Schneider, Wallace, Blikstein, & Pea, 2013) that can strengthen human recall ability (Kuznetsov, Dey, & Hudson, 2009). The promise of tangible computer interfaces (Zuckerman & Gal-Oz, 2013), often exemplified through the works of Ishii (Ishii & Ullmer, 1997) and Weiser (Weiser, 1991), continues to be explored through a vast array of techniques and technologies (Fishkin, 2004; Sodhi et al., 2013; Weiss et al., 2009, 2011; Savage, Zhang, & Hartmann, 2012; Gupta et al., 2013; Manshad, Pontelli, & Manshad, 2012; Jansen et al., 2010; Bau et al., 2010).

The growth of tangible interactions in the physical space is increasingly more approachable because of advances in low-cost rapid prototyping tools, such as 3D printing (Klein, Adams, Dickin, & Simske, 2013) and microelectronics (Savage et al., 2012), which can support fast and inexpensive production of the tangible objects. As the types of devices envisioned by Ishii and Ullmer reached commercial viability, researchers discovered new ways of incorporating those devices into the physical world. For example, the metaDESK combines digital imagery with physical objects (phicons) using sensors and projection (Ishii & Ullmer, 1997). The emergence of tabletop multi-touch displays put the same functionality into a single device, allowing researchers to broaden the set of physical metaphors in new directions (Weiss et al., 2009, 2011). Similar efforts have ranged from input only—such as the use of capacitive touch to improve mobile phone interaction (Savage et al., 2012)—to output only—using vibrotactile feedback to add texture to touch surfaces (Bau et al., 2010). There have also been significant contributions towards developing novel forms of communication using the tactile channel. The AIREAL project successfully created the sense of movement around the body by using ultrasonic pulses of air (Sodhi et al., 2013). Similarly, Gupta et al., used air vortices to provide haptic response to gestural inputs. Using frequency modulation of an electrovibratory surface, Bau et al., created TeslaTouch to demonstrate how an otherwise smooth surface could produce a range of tactile sensations from stickiness and waxy to bumpy and rough.

Affordable digital fabrication techniques have already proven their value in the domain of assistive technology (Hurst & Tobias, 2011). Thanks to these advances, tangible interaction is increasingly accessible and appealing (Shaer & Hornecker, 2010). The blending of linear access, visual direct manipulation, location, and other dynamic mechanisms hold promise towards creating higher quality experiences throughout all forms of interactive computational spaces. The primary focus for tangible interaction in assistive technology has been text display through the refreshable tactile output of braille displays. A lot of effort has gone into converting the rich set of visual desktop metaphors into something that can be controlled and displayed in the relatively limited world of a single 24 to 40 character braille display, mouse, keyboard, and speaker system (*e.g.,* (Ratanasit & Moore, 2005; E. D. Mynatt & Weber, 1994)). However, the human perceptual system is much more sophisticated in its ability to remember and manipulate objects in the world than these traditional technologies support. Our ability to remember the location of things in space, for example, is so good that this is often used as a memory aid for remembering a list (Ängeslevä, Oakley, Hughes, & O'Modhrain, 2003). Similarly, augmenting visual memorization tasks with tactile cues has been shown to be an effective memory aid for individuals with poor recall abilities (Kuznetsov et al., 2009). For repeat tasks, spatial memory is further enhanced by proprioceptive capabilities, which are the ability to locate the relative position of body parts (Folmer & Morelli, 2012). Thus, people have an ability to easily place things in space, remember their location, and remember repeat tasks (such as reaching repeatedly to the same location).

More recently, tangible computing has gained interest as an assistive technology to augment the loss of the visual input channel. For example, TeslaTouch was repurposed as a tool for communicating 2D tactile images to the blind (Xu, Israr, Poupyrev, Bau, & Harrison, 2011). Zuckerman and Gal-Oz, demonstrated that physically interacting with tangible objects elicits a sense of enjoyment—one that is preferred over traditional forms of interaction (*e.g.*, keyboard and text-to-speech), even when the given task might take longer. This enjoyment is likely a result of the learnability of tangible interfaces due to their innate ability to be manipulated and explored (Schneider et al., 2013). However, matching the enjoyment of exploration to the complexities of a modern computing system remains a challenge. While multimodal sensory approaches that combine tactile and audio feedback have been shown to be effective at improving task completion times (Kuber, Yu, & O'Modhrain, 2010; Vitense, Jacko, & Emery, 2002; Menelas, Picinali, Bourdot, & Katz, 2014), the amount of information that can be communicated is limited. S. Brewster and Brown translated the properties defined for earcons (Blattner, 1989) (frequency, duration, rhythm, and location) to the haptic channel in what they named "tactons." Tactons explored vibrotactile feedback as a source for generating abstract patterns to communicate complex messages nonvisually. Follow up studies demonstrated that tacton patterns were distinguishable, though some patterns were more successful than others (L. Brown, Brewster, & Purchase, 2005). Prescher, Weber, and Spindler, and others (*e.g.*,(Köhlmann, Zinke, Schiewe, & Jürgensen, 2010; Völkel, Weber, & Baumann, 2008)) removed audio entirely by taking a literal approach to mapping graphical interfaces to the tactile modality with the BrailleDis9000, a 12 line by 40 character braille display capable of rendering shapes and braille characters. Prescher et al. repurposed the device to present individual windows and widgets that could be mimicked tactilely by raising specific pins, finding that although the tactile representations were difficult to grasp for some participants, most preferred the direct manipulation afforded by the device over input through keyboard commands (Prescher et al., 2010).

## 2.3 Mediating Action Through Shared Control

Since its passing, the Americans with Disabilities Act (of 1990, 1990) has brought access to vital services for the disability community; yet negative stereotypes and awareness towards people with disability pervade contemporary culture (Zychlinski, Ben-Ezra, & Raz, 2016; Scior, 2011). It is generally accepted that increased exposure can positively affect attitudes towards disability, but the mechanisms for supporting contact are less understood (Barr & Bracchitta, 2015; Tan, Wilson, Campain, Murfitt, & Hagiliassis, 2019). Exposure alone does not appear to sufficiently promote positive attitudes in mixed-ability groups (Keith, Bennetto, & Rogge, 2015). Rather, prolonged, meaningful contact between mixed-ability groups is required to create sustained positive outcomes. Keith et al. argued that influential contact that is perceived as “equally cooperative and pleasant" is necessary for a positive change in awareness (Keith et al., 2015). Drawing on conversations within the disability community, Bennett et al. encouraged assistive technology researchers to consider the “interdependent" relationship within mixed-ability groups in their work (Bennett et al., 2018). By embracing interdependence, issues that require expensive or complex computational interventions can be handled as a separate case from enabling access to activities.

The concept of shared control of assistive technologies exists in different capacities across a variety of different fields of research. However, the use of the term “shared control" is inconsistent and never explicitly perceived as a human-centered relationship. One of the earliest examples of a shared control assistive technology is the NavChair (Levine, Bell, & Koren, 1994), a wheelchair that integrates machine intelligence to assist with obstacle avoidance. Levine et al. define shared control as a relationship between a human and machine, where the human is intentionally assigned greater autonomy than traditional human-machine systems. Lacey and MacNamara apply a similar interpretation of shared control to their smart walker, which uses a Bayesian network to automate contextual shifts between automated and manual operation modes. Building upon the human-machine interpretation of shared control, Cortes et al. adopt the phrase "shared autonomy" to describe the relationship between users and assistive technology within the context of the multifaceted SHARE-it platform (Cortes et al., 2010). SHARE-it integrates input across a network of systems and services including caregivers, physicians, machine intelligence, and assistive technology users to manage functions for a particular assistive device. The authors argue that shared autonomy is a more suitable description of the functionality that SHARE-it provides. Although Cortes et al. do incorporate additional human input into the shared relationship, it is conducted passively—without the need for direct human to human contact. Similarly, though not explicitly defined as shared, other areas of research have explored the human-machine relationship through modern smartphones (Jayant, Ji, White, & Bigham, 2011; Fiannaca, Apostolopoulous, & Folmer, 2014; Vázquez & Steinfeld, 2012). For example, Jayant et al. use image recognition to enable blind users to more easily identify targets when taking pictures on an iPhone.

Assistive technologies explicitly designed to be shared through direct human to human interaction, remain underexplored in the literature. Some recent scholarly work has emphasized the value of designing with the blind and low vision community from a methodological perspective. Design practices like participatory design are useful for bringing users into the design process (Muller, 2003). Though, as Andrews argues, researchers need to adapt to the abilities of their target audience (Andrews, 2014). According to Morrison et al., in a multi-workshop study using "tactile ideation techniques", blind participants struggled with the ideation process while prototyping with physical objects in a participatory design setting. Despite enjoying the process of manipulating tactile objects, participants with less vision were taxed "in a way that did not encourage ideation", highlighting the challenges for researchers to bring blind and low vision users into the design process effectively (Morrison et al., 2017).

One way to engage blind and low vision participants in participatory design is to draw from the multiple levels of expertise within existing communities (Vredenburg, Mao, Smith, & Carey, 2002). Cooperative design or co-creation/co-design—a subfield of participatory design—positions designers, researchers, and users equally as experts of their own experience (Sanders & Stappers, 2008). Co-design principles are grounded by the principle that individuals contribute according to their own creativity and ability. Success, according to Sanders and Stappers, is dependent on not pushing individuals beyond their own level of interest (Sanders & Stappers, 2008). In my work on mixed-ability co-design (M. S. Baldwin, Mankoff, Nardi, & Hayes, Submitted), as well as the study I present in this dissertation, I build upon the principles of co-design by situating a mixed-ability design team within a public setting, similar to the work of Teal and French on Designed Engagement (Teal & French, 2016) in which participants are drawn through the public setting, rather than recruitment. In addition to the ability to uncover insights that exist outside the boundary of a design team, Teal and French suggested that public engagement through design serves to build empathy with the public as they become active participants in the design process.

Drawing upon existing practices from disability studies and participatory design, Mankoff, Hayes, and Kasnitz argued that deep, long-term engagement between researchers and disabled individuals can lead to positive results that benefit everyone involved (Mankoff et al., 2010). Furthermore, real-world deployments serve to mitigate the disparity between researchers and subjects of differing abilities (Wobbrock, Kane, Gajos, Harada, & Froehlich, 2011), and acknowledged that users with disabilities are experts of the assistive technologies that they use (Shinohara et al., 2018). However, as Kane, Hurst, Buehler, Carrington, and Williams pointed out, deployments within organizations risk small, repetitive participant pools (Kane et al., 2014). The work I describe in the next two chapters seeks to mitigate some of these risks through public-facing, inclusive design sessions.

# Chapter 3

# Augmenting Assistive Technology Through Tangible Interaction

In my work on the Tangible Desktop (M. S. Baldwin et al., 2017), with support from my co-authors, I set out to understand the ways in which screen readers, the primary source of audio in nonvisual computing environments, are used by the visually impaired community and probe how I might design improved interfaces for nonvisual access. Inspired by early graphical interfaces that appropriated physical metaphors with which users were already familiar, I wondered if these same metaphors might be usefully re-appropriated into a nonvisual interactive system.

I first conducted a field study of screen reader use at a school for adults who are losing or have recently lost their sight. Over the course of four months, I observed eight students as they were taught how to use audio-based assistive technology software and tools to perform common desktop computer tasks. Based on the insights gained from this work, I developed a high-fidelity prototype for multimodal nonvisual computing access, called the Tangible Desktop (M. S. Baldwin et al., 2017), which I examined through an experimental study with these students.

This work offers two complimentary contributions to scholarly thinking about accessibility. First, I describe the challenges of nonvisual computing, with a particular emphasis on how I might use alternative input techniques beyond simply audio. The in-depth exploration of this problem that I conducted with my co-researchers sheds light on underlying assumptions surrounding accessible software development and assistive technologies themselves. Second, I demonstrated the feasibility of an alternative multimodal approach and show the ways in which such systems might actually be easier and more efficient than screen readers, particularly for novice users.

## 3.1 Field Study of Nonvisual Computing

I conducted a field study in a blind and low vision computer class at a non-profit organization over a period of four months. This non-profit organization, EmpowerTech[1], provides computer based job skills training (*e.g.,* office productivity and Internet navigation skills) for people who are blind or losing their vision. The majority of students are relatively new users of assistive computing tools, such as screen readers and screen magnifiers.

Over the four-month period of this study, I participated in twelve classes averaging four hours per session for a total of 48 hours of participant observation. Being both sighted and an expert user of screen readers and other accessibility technology, I participated in the classes as a teaching assistant, doing whatever tasks were assigned by the lead instructor, a blind man who is an expert computer and screen reader user.

Class size varied from week to week, but typically consisted of eight to twelve students (see Table 3.1 for student details). At the time I joined the class, experience with screen

[1] Empowertech consented to the use of their name in this work.

| ID | Gender | Age | Technology | Condition | Time with Condition |
|---|---|---|---|---|---|
| B1 | Female | 39 | Screen Reader | Glaucoma | 9 years |
| B2 | Male | 35 | Screen Reader | Optic Neuropathy | 10 years |
| B3 | Female | 43 | Screen Reader | Retina Pigmentosa | since childhood |
| B4 | Female | 36 | Screen Reader | Retina Pigmentosa | since birth |
| B5 | Female | 41 | Screen Reader | Retina Pigmentosa | 7 years |
| B6 | Male | 37 | Screen Reader | Optic Neuropathy | 4 years |
| B7 | Female | 41 | Screen Reader | Retina Pigmentosa | since birth |
| M5 | Female | 34 | Magnification | Glaucoma | 8 years |
| B9 | Female | 45 | Screen Reader | Glaucoma | 4 years |
| I1 | Male | 52 | Screen Reader | Unknown | since birth |
| M3 | Male | 35 | Magnification | Cataracts | since birth |

Table 3.1: Detailed list of the visually impaired students and instructor observed at the field site. Technology represents the primary tool students preferred to use. Participant ID labels are coded according to technology; B group had less than three months of experience with their assigned technology, M group used magnification, and the I participant was the blind instructor.

readers and magnifiers varied across students, but none of them had received more than three months of training. Field notes were taken during the class when possible and directly after. In class instruction focused heavily on screen reader operation and keyboard commands for common computing tasks, such as file management, web browsing, and word processing. Outside of class time, the students often stayed at EmpowerTech to practice their skills, providing additional time for observation of their computer use and for informal interviews. Field notes documented seating arrangements, lecture topics, software being used, and the interactions between students and the instructor. Recurring themes during lecture included the advantages and disadvantages of various assistive technologies as presented by the class instructor and the struggles and conveniences for the students.

The training classroom was rectangular in shape with two exits on either side. Tables lined the walls of the room and provided work areas for students to sit. Each table had two or three computers, mostly PCs running Microsoft Windows 7[2] except for one table, which housed

[2]https://www.microsoft.com/en-us/windows

two Apple iMac systems[3]. Sitting next to one of the iMac systems was a magnification screen that used an optics system to increase the viewable size of any object placed beneath it. In the center of the room there was a round table that was large enough to seat four students; however, it was often pushed up against one of the rectangular tables to keep the center of the classroom open for the instructor to move more freely as he taught. Many students brought their own laptops to class making the round table the preferred place to sit.

Classes at the field site are run Monday through Thursday from 9am to 1pm. The tools used within the classroom varied widely between students. Most students brought their personal laptop with them to class. During my time at EmpowerTech, there was only one student (B3) who did not have a laptop computer. Two students used Apple MacBooks (B5,M5), the remaining students used a mixture of Windows 7 systems. Additionally, two students transitioned to new systems: B6 migrated from Windows 7 to a MacBook, and B4 upgraded from Windows 7 to Windows 8.

Over the course of my four month study, the class progressed from learning how to use a screen reader to web navigation, email, and text editing. I typically collected data through class-wide observation. On days when one or two students needed assistance beyond the scope of the lesson plan for the day, I was asked to work directly with the struggling students. Adopting the role of teaching assistant enabled me to observe the types of challenges that students struggled with first hand. As I was unable to attend to the class as a whole in these instances, I shifted the emphasis of my data collection from the class to the individual, conducting unstructured interviews with the student.

Although the classroom had specialty devices available, like a digital magnifier and braille printer, neither were ever used. The primary tools that students relied on were the Victor Reader[4] and screen reader or screen magnifier. The instructor encouraged students to choose

[3]https://www.apple.com

[4]http://www.humanware.com/microsite/stream/index.html

whichever operating system and assistive tools that they preferred. The Victor Reader, a handheld audio recording and playback device with text-to-speech capability, was used by all of the students. Students primarily used the device to capture verbal instructions from the teacher, which they played back at a later time while working on class assignments. However, the devices also supported audio output of music and electronic publications like PDF's and e-books. Although the core features of the Victor Reader could also be found on smartphones, the students that I spoke with found the tactile input through the raised buttons of the device faster and more convenient to use than a smartphone touchscreen. The most commonly used screen readers were JAWS[5] for students running Microsoft Windows and VoiceOver[6] for students on Apple MacBooks. A few of the desktop computers within the classroom were setup with multiple screen readers. In addition to VoiceOver, the iMac could also use the Dolphin screen reader. At least one of the Windows desktops had NVDA[7] installed.

Following each class, I spent time with individual students and conducted informal interviews about their experiences with technology both in and out of the classroom. Additional artifacts, including handouts and worksheets assigned by the instructor, were collected at each class period and saved for analysis. I paid particular attention to bottlenecks and challenges during instruction and work time. Between classes, I engaged with the most commonly used software in the previous class to better familiarize myself with the opportunities and challenges the tools provided and to prepare to support students during class.

Data analysis combined a mixture of inductive and deductive approaches. All field notes were read by the entire research team between each class, and discussions occurred about the notes during weekly meetings. Data were analyzed in an iterative fashion with a constant comparative approach. Each week, field notes focused both on the general themes and

[5]http://www.freedomscientific.com/Products/Blindness/JAWS
[6]http://www.apple.com/accessibility/osx/voiceover/
[7]http://www.nvaccess.org/

questions that prompted my initial research and emergent questions from analysis I conducted with my collaborators. I, along with my collaborators, first examined the data using known challenges from the literature, such as the shortcomings of screen readers (Borodin et al., 2010), software accessibility compliance (Mankoff et al., 2005; Richards, Montague, & Hanson, 2012), and user frustration (Lazar et al., 2007). Using an iterative, inductive approach, I then identified additional emergent phenomena, named, and categorized these issues. I used affinity diagramming and axial coding to understand the relationship between, across, and within these codes as well as to those themes from the literature and my original deductive coding. While completing data collection and analysis, I iterated on the prototypes described in section 3.3. As dominant codes emerged, they were incorporated into the categories presented in the next section and considered as part of the prototype development process. This intertwined process enabled me to test the boundaries of technical feasibility as ideas were broadened and then refined.

Over the course of the field study, I met regularly with the course instructor and school staff. These meetings supported three complementary goals of the research team. First, the team was interested in building a strong community partnership with EmpowerTech, including the ability to hear from them what research questions they might have or technical innovations they might wish to see. Second, I conducted preliminary analysis and presented these initial ideas to them to inform the next data collection session. Finally, as the fieldwork drew to a close, I began to shift the discussion towards larger analytic themes and design guidance for my prototype solutions, again getting feedback from the members of the community I was studying and support from the organization with whom I was partnering.

## 3.2 Effects of Audio in Computational Contexts

In my fieldwork, working alongside blind and low vision students in an active learning environment, unmasked the impact of audio in nonvisual computing contexts. In the classroom, auditorily communicated information was constant. Verbal instruction, student discussions, computer speech and sonification, as well as intermittent ambient sounds placed a consistent stream of information processing upon the students. In this section, I present the results of the analysis I conducted along with my co-authors on my fieldwork data. I describe the challenges participants encountered in using screen readers, in terms of the dual issues of keystroke memorization (input) and audio interpretation (output). These challenges led to an initial conceptualization of an alternative tangible based approach, which I describe in section 3.3.

### 3.2.1 Performance Constraints of Memorization

In lieu of the point and click interaction on which sighted computer users rely, nonvisual screen reader use is dependent on keyboard input commands. Encapsulating all of the actions required to perform visually oriented computing tasks for an audio-only interface increases the commands required to perform similar actions through direct manipulation. For example, in the fifth week of my observations, students were given a worksheet containing over forty commands for operating a screen reader and Microsoft Word, ranging from document formatting to file management. Next to each task was a blank field in which students were required to enter the task that the command accomplished. I arrived that week shortly after the students had received these worksheets, and they were working through them, with some of them audibly complaining. Complaints ranged from concern that the commands would not be useful for those who still had some sight to the concern that the commands themselves were challenging to remember. At the end of the class, I discussed the students concerns over memorization:

> When the instructor returned, I discussed what had happened with him. He acknowledged that there are a lot of commands to learn. "This is why everyone has a Victor Reader." The Victor Reader is a small recording and text-to-speech device that the students use to capture and playback the instructions. – *Field notes from fourth week*

This tension between learning what will be useful later and learning what is needed right now is a large issue in the teaching of assistive technologies (Zhou, Parker, Smith, & Griffin-Shirley, 2011; Sapp & Hatlen, 2010; Abner & Lahm, 2002). The use of recording devices, such as the Victor Reader, point to the complexities of learning audio only interfaces. The adage "recognition over recall," a fundamental tenant of Human-Computer Interaction for over 20 years (Nielsen, 1994), is clearly lost when instructions for one system must be captured and replayed through another system. This issue is further confounded when navigating the variability across software systems. During one lecture the instructor responded to a question about the best web browser to use by saying:

> "You need to learn all browsers because each one is better at a particular task. For example, FireFox is better at reading captcha's. I only use FireFox to pay bills, that is what it is good at." – *Seventh week, participant I1*

The hurdle that memorization presents is not simply a result of being a novice user; the system should be given equal responsibility. Not surprisingly, faced with the option of using GUIs, even in a difficult and challenging way, students who were able to do so, always chose to use the visual interface over memorization. This choice demonstrates that the obstacle of learning to use the memorized commands may simply be too high for most novices, an issue I address directly with my tangible approach. One participant with low vision relied on her ability to perceive the motion of the focus indicator, a small dotted outline that surrounded the selected element, as she navigated around the various elements of a web page. Detecting the motion helped to orient her location within the page after which she would rely on the audio output to select the desired element. Another participant was so adamantly opposed to using screen readers that he was willing to use a traditional glass magnifying lens coupled with a screen magnifier, a clumsy and painful solution, to avoid it. As disparate as these approaches seem, they are indicative of the students' preference for more than an audio-only interface.

### 3.2.2 Challenges of Audio-Only Output

Even with a reasonable amount of keyboard commands memorized, locating information on screen remained challenging for the students I observed. A common tactic employed by the students was to use the basic navigation controls like the tab and arrow keys to move between elements sequentially, a strategy identified by Vigo and Harper as exhaustive scanning. With each element change, an auditory processing step is then required to hear and grasp the semantic description of the selected element and its associated information. For example, while attempting to download an email attachment in a web-based email client, a common difficulty among screen reader users (Wentz, Hochheiser, & Lazar, 2013), two students were using the tab and arrow keys to bring focus to the attachment link. Unknowingly, the command inputs that they were using prevented them from reaching the panel in which the attachment link was embedded. This resulted in a cyclical process of keyboard input and audio output that would restart after all the elements had been traversed. The students looped through the content numerous times searching for the attachment link before asking for help. Using these kind of strategies enables people to recover when they have forgotten a complex command sequence or to avoid learning such things in the first place. However, the recognition over recall strategy for primarily auditory output comes with substantial costs in terms of both time and frustration.

Additionally, as described in the above example, the ephemerality of the audio stream may lead users to believe they have “missed" the link, when in fact they never encountered it. Laboring under this false belief, they take multiple passes linearly through the content. Eventually, users may recognize their inability to target the desired information, but even then, they may not grasp why. Without sight, recognizing that a particular piece of information is on another panel, which is relatively inaccessible to the screen reader in its current configuration, is nearly impossible.

Finally, the costs of this kind of traversal based seeking only grow with increased content. In the example above, each pass took an average of around 30 seconds, with the users speeding up (from 60 seconds on the first pass to 45 seconds on the second pass, and so on) as they began to recognize repeated words. Considering that established listening comprehension rates average roughly 190 words per minute (Foulke & Sticht, 1969), traversal time for content rich documents can quickly lead to unwanted delays. For example, in a similar situation observed during the sixth week of class, job searching in particular was challenging for participants as noted in my field notes:

> *B2 was attempting to search for jobs, with traversal of the results page of 20 job listings taking more than 10 minutes to complete. With a much larger content space the traversal time would increase to unusable proportions. In this particular case, the traversal time between the beginning and end of the results was long enough that B2 was not able to recognize that the traversal had looped back to the beginning.* - Seventh Week, participant B2

Not understanding content scale and excessive audio processing posed formidable constraints to efficiently accomplishing basic computer tasks.

The time cost of seeking is further confounded by the complexity of the user interface, where *complexity* represents an increase in non-linear text arrangement. During another conversation with M3 on his preference for magnification software, his dislike of the screen reader was further explained:

> *"That's another reason I don't like the screen reader, is it sometimes reads stuff I don't even need it to read."* - Twelfth week, participant M3

Where non-linear arrangement benefits sighted users by placing tertiary content to the side, the screen reader treats all text equally, sending undesired information to the user for processing. For example, a table provides a convenient, easy to process visualization of categorical data. However, when targeted by a screen reader, a detailed description of the structure and arrangement must be communicated before any of the actual data. The following quote is the text-to-speech output of a web based email inbox using the NVDA screen reader:

> *"Table with thirty-six rows and four columns, row one column one, checkbox not checked, column two link mom."*

In this example, 18 words must be spoken before the user can identify that the first email in the inbox is from "mom." Although shortcuts to navigate around the inbox are typically available, many of them are unique to the application being used, and difficult to master for all but the most advanced users. Still, the benefits of committing these immense command sets to memory is only as beneficial as the software infrastructure allows. One of the more advanced students at EmpowerTech (M5), was forced to use the seeking strategy to traverse results from a housing search when the website she was using failed to properly use heading tags to organize the search results. A simple mistake in formatting made it impossible to use the screen reader shortcuts she had learned. Here, knowing that there was an easier way, but that it failed to work, was incredibly frustrating for this student. Her solution was to print the listings in a very large font and rely on her limited vision or the help of a friend to read the results. From her perspective, this was preferable to having to repetitively seek the information.

A wide variety of challenges exist beyond simply searching through a dense, ephemeral, and linear audio stream. In particular, input is intrinsically also an act of output for sighted users, though it is rarely recognize as such. As words are typed into the screen, we see them appear letter by letter. Anyone who has ever been dogged by slow response time from a website, piece of software, or operating system can attest to the challenges of typing when whole words or sentences show up rather than character by character. Nonvisual users, however, do not have this benefit available to them. For example, in one instance, P2 struggled to understand why a job site was not able to find any jobs using his queries. The issue was easily identifiable to a sighted observer: the search field had identical search terms concatenated together. P2 was performing the task he had been trained to do, but when a shortcut designed to reduce repetition for visual users (pre-populating a text box with last used text) failed to notify P2, his workflow fell apart. These kinds of shortcuts do indeed make interaction faster and more efficient for sighted users. However, they are not visible to those using screen readers. Thus, it became clear during this fieldwork that two way communication with the computer, would be required, rather than the parallel but unidirectional flow of keyboard only input and auditory only output.

The context of the computing interaction can further exacerbate the issues of audio output. My fieldwork was particularly helpful in understanding how screen readers might work in an open work environment, the kind common to many office spaces, particularly in the United States. The training program primarily took place in one small classroom as described above. Although each student had an individual workspace, and no computers were shared except when students were explicitly collaborating, the noise in the room still rose to distracting levels during every observed session. Distractions were common as individual students entered and left the room or talked with each other. Similarly, the instructor needed to listen to the students' screen readers to make sure they were pacing appropriately with the material as well as to provide assistance when difficulties arose.

What I see from this experience is a larger issue related to distractions in general. As noted in prior work (González & Mark, 2004; Dabbish, Mark, & González, 2011; Mark, Iqbal, Czerwinski, & Johns, 2014), distractions in the workplace can be problematic for any computing user. Following a distraction, individuals must reorient to the original task, an activity that requires a certain amount of cognitive load and usually involves glancing across the workspace. In the case of nonvisual computing, however, I observed additional challenges to the reorientation task following a distraction. Not able to quickly glance through the screen, participants had to re-listen to some amount of audio. The amount required can differ, leading to users sometimes listening to a short portion of audio, backing up further and listening to a longer portion (inclusive of the short portion they just heard), and so on, sometimes doing this several times before they can find their place. At times, this can be even more problematic if the distraction led them to make an error, thereby navigating them to a different page or even a different application without their knowledge. Although sighted users bump their mouse or make errant keystrokes regularly when distracted, these are relatively easy to overcome using visual inspection. Without that option, however, I see even greater challenges. These results reinforce the need for a way to easily mark one's place when interrupted and avoid relistening to substantial portions of audio.

## 3.3 Replacing Auditory Information with Tactile and Multi-Modal Interactions

Tactile interaction makes it possible to replace the ephemeral nature of the audio stream with a permanence similar to that of a graphical display. Building on the earlier example of icons, if an icon were placed in the physical world and bound to the system state such that when an icon corresponding to an application is open, closed, or active, its physical state could change to match. In this way, inferring meaning through shape and location, the condition of the system state could be accessed directly by the user through touch without processing an audio stream.

I set out to explore this space by creating peripheral devices that target two common desktop computer activities required for almost all computer use: switching and locating. The "Tangible Desktop" system is comprised of physical implementations of the computer taskbar and application window scrollbar (see Figure 3.1), built with rapid prototyping tools (*e.g.*, 3D printer, soldering iron, and electronics hardware). In this section, I describe the potential design space for multimodal interactions as well as the prototype system I developed to evaluate that design space.

### 3.3.1 A Tangible Desktop

The peripherals I created for the Tangible Desktop were built around a motorized slide potentiometer, a device commonly found in audio mixing boards, using low-cost electronics and a 3D printer. The physical interaction of slide potentiometer enforces directional motion along a single axis. The combination of a potentiometer for data input and a motor for output in a single prefabricated unit enabled me to explore the benefits of maintaining a synchronized state between the computing system and the user's cognitive model through bidirectional control. The potentiometers were controlled through a microcontroller that relayed commands over USB to a host computer running Microsoft Windows 7. To illustrate how the Tangible Desktop works, I present the following scenario, which describes the typical workflow of a person using the Tangible Desktop:

Joanne gets ready to work by opening her laptop and connecting the Tangible Desktop. Today she is putting the final touches on her resume to be sent off to a job posting she hopes to secure. Unsure which program is currently active, Joanne places her hand on the Tangible Taskbar. Recognizing that the selection thumb is aligned with the resume icon, she slides it quickly to the jobs icon. The system reacts by giving focus to the jobs website that she had previously bound to the icon. Once the Tangible Desktop has fully loaded the jobs website, the slide thumb on the Tangible Scrollbar slides to the location of the last job posting she had read. She skips to the next job posting by moving the scrollbar thumb in a downward motion. As she moves the thumb, she feels a small detent on her fingers as the name of each job post heading is read aloud by the computer text-to-speech application. When she hears a job title that sounds interesting, she moves the thumb slower until it vibrates, indicating that she has encountered a link. The text-to-speech application reads the link aloud, "Click to Apply." Joanne presses the button on the Tangible Scrollbar to click the link and load the next page. Once the job detail page has loaded, the Tangible Scrollbar thumb automatically reorients to the top, indicating that she is now on a new webpage. She continues navigating the page using the scrollbar thumb, listens to the job description, and decides to apply for the job. She places her hand on the Tangible Taskbar and moves the taskbar thumb until is aligned with her resume physical icon. The system activates the word processing document with her resume and gives it focus. Joanne begins updating her cover letter to emphasize how her skills match the needs described in the job description.

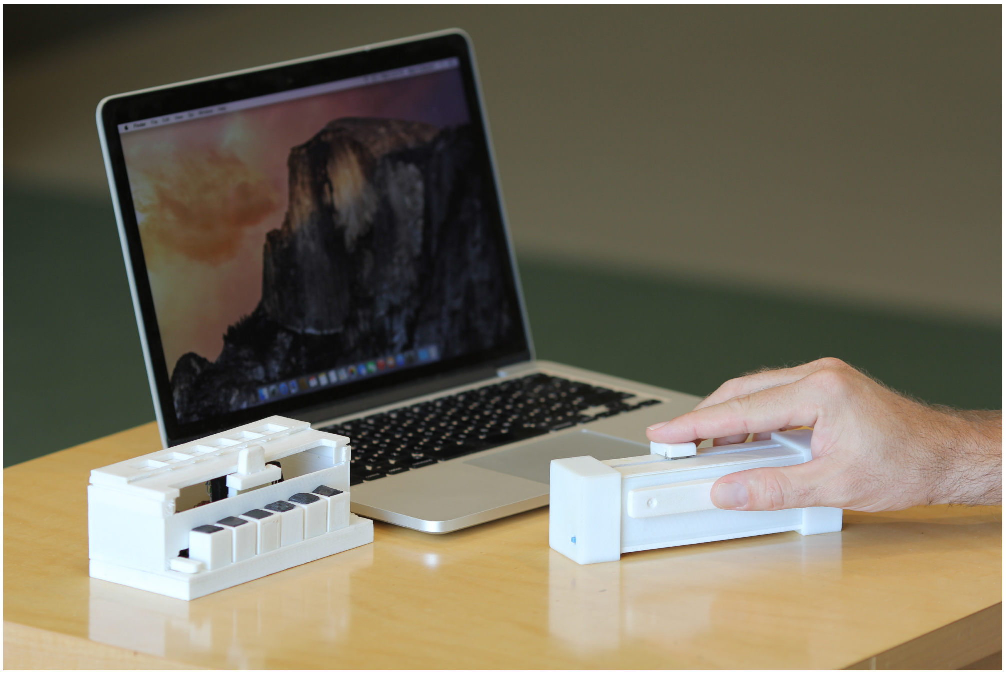


Figure 3.1: A picture of the Tangible Desktop in its standard arrangement. The Tangible Taskbar sits to the left of the laptop while a user engages with the thumb of the Tangible Scrollbar.

#### Design and Function

The Tangible Taskbar uses the potentiometer to switch between different computing entities (i.e., files, programs, and browser tabs). Entities are represented by physical icons, conceptually inspired by metaDESK's phicons (Ishii & Ullmer, 1997). However, in place of the optical and electromagentic sensors used in metaDesk, Tangible Desktop icons are implanted with RFID chips (see Figure 3.2). The physical icons were 3D printed with a two millimeter deep crown that was filled with moldable rubber. The rubber crown for each icon was given a unique pattern of indents and ridges. The surfaces allowed for tactile differentiation between icons. An inexpensive, commonly available RFID reader is mounted to the slide arm of the potentiometer, allowing the RFID chips in each icon to be read as the arm is moved. A 3D printed housing is used to enclose the hardware and provide a slotted tray for placing the physical icons. Communication software that runs on the host computer manages the binding between RFID chip identifier and the desired computer entity. A physical button on the Tangible Taskbar sends a binding command to the host computer, instructing the communication software to assign the active window to the RFID chip's unique identification number. Once a binding is created, the software transitions the bound entity to an active state whenever the identifier is read by the RFID reader. For example, if a physical icon is bound to a browser tab containing the University of California homepage, when the slide arm is moved inline with that physical icon that tab will be given focus. If the desired entity is already in focus, then the environment will not change. Likewise, if a different bound entity is given focus through external means (i.e., mouse or keyboard), the motor attached to the potentiometer will move the slide arm to the corresponding physical icon.

The Tangible Scrollbar uses the potentiometer to traverse content in an application. A small eccentric rotating mass vibration motor is mounted to the slide arm to provide vibrotactile feedback to the hand. While other projects have explored the richness of vibrotactile feedback (*e.g.*, (L. Brown et al., 2005; Kuber et al., 2010; Xu et al., 2011)), the Tangible Desktop system intentionally limited the amount of information communicated through vibration to avoid many of the complications that accompany vibratory patterns. Tactile push buttons are mounted at either end of the device to provide additional input capabilities while using the scrollbar. Like the Tangible Taskbar, a 3D printed housing encloses the hardware, and provides a graspable bar for pressing the buttons from multiple hand positions.

For my preliminary exploration, I limited the scope of the Tangible Desktop to switching tabs and traversing content in a web browser. I leveraged the extensibility of the Microsoft WebBrowser Control[8] and Speech API[9] to create a custom screen reader application. This approach enabled me to generate a testable environment without having to build fully functional drivers and software.

[8] https://msdn.microsoft.com/en-us/library/aa752040%28v=vs.85%29.aspx
[9] https://www.microsoft.com/cognitive-services/en-us/speech-api

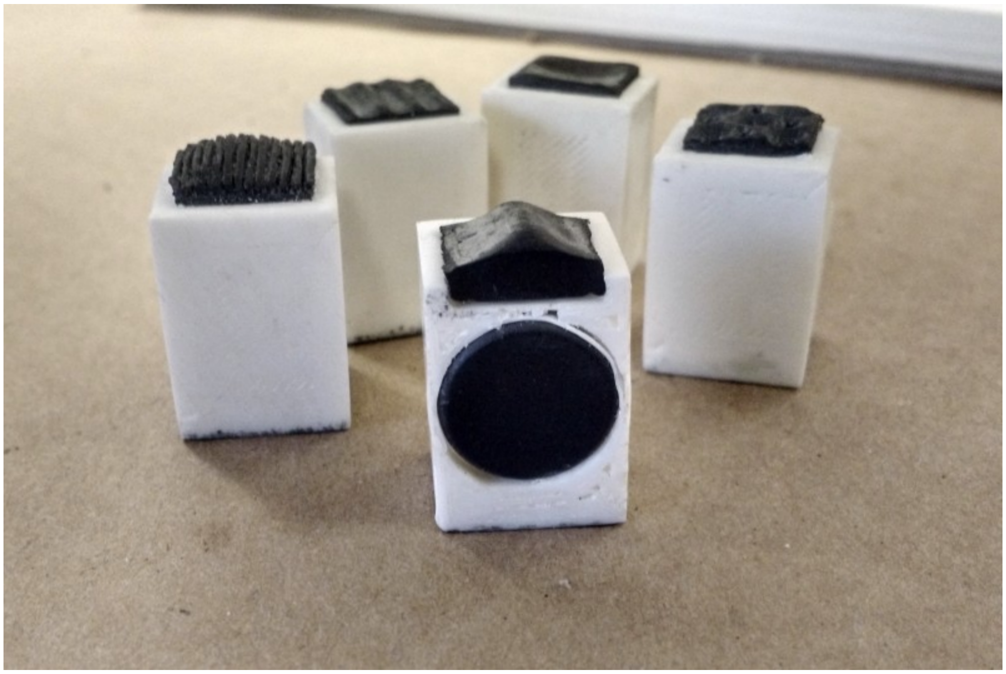


Figure 3.2: The physical icons used in the Tangible Desktop (see Figure 3.1). Each icon has an RFID tag embedded inside and a tactilely distinct rubber crown.

### Supporting Information Discovery

Information that is output from a screen reader is paired with a semantic description of the lexical information that has been requested by the user. This structure places two levels of cognitive processing on the user. First, they must comprehend the type of information; then they must process that information. For example, a document may contain multiple hierarchical levels, varying content lengths, and hyperlinks to other documents. Using a screen reader, the headings (or hyperlinks) are specified semantically (*e.g. 'heading level 1 Introduction'*), and content length must be requested using input commands. By moving this semantic information out of the audio channel and into the tactile realm the amount of audio the user must process is reduced. Furthermore, this enables the system to communicate the semantic and lexical in parallel, reducing the overall processing time (see Table 3.2).

The Tangible Taskbar is a physical representation of the desktop computer taskbar that places the basic operations surrounding a single application into a real-world, tangible object.

| Screen Reader | Screen Reader with Haptic |
|---|---|
| “Heading level one, Introduction" | “Introduction" + tap |
| “Link next page" | “Next page" + vibration |

Table 3.2: A comparison of screen reader speech output between auditory and haptic semantic descriptions. Note that haptic responses can occur in parallel with speech output.

I built the prototypes using 3D printing, moldable rubber, and radio frequency identification (RFID) tags to create tactilely identifiable icons. With the support of an RFID reader, the icons can be individually bound to a single entity within the computing environment.

The taskbar provides a space to rest individual RFID based icons, which can be bound to applications on the host computer. As an icon is placed on the taskbar, the associated application is launched. When the icon is removed, a close request is sent to the application. A slider located behind the taskbar indicates which icon has focus. A user can switch between applications by moving the slider into alignment with the desired icon. If the active window is changed through a different input (*e.g.,* keyboard or mouse), the slider automatically moves into alignment with the icon bound to the newly active window.

### Constructing Data Permanence

The Tangible Scrollbar is a navigational device similar to the scroll wheel found on most computer mice. It is used to linearly traverse and navigate content while simultaneously communicating visual and semantic information tactilely, thereby supporting the kind of two-way communication and permanence required to help users find their place in documents and remember their last place. Unlike the mouse scroll wheel, which uses a free spinning motion, the Tangible Scrollbar has distinct endpoints that physically indicate the beginning and end of scrollable content. As content is traversed, the scroll handle delivers resistance and vibration to communicate semantic information to the hand.

To accommodate varying lengths of content, a distance of travel ratio is calculated to spread haptic feedback equally between the scroll endpoints. This ratio is then used to determine how far the scroll thumb needs to travel before providing feedback. To reduce dependency on a mouse or touchpad, the Tangible Scrollbar also has two buttons that can be programmatically assigned to perform the same tasks as a mouse button.

When a web page is loaded for the first time in the custom browser, a reset signal is sent to the scrollbar, prompting it to move the scroll thumb to the home position. As the scroll thumb is moved, the web browser traverses all HTML elements that contain lexical content. The lexical data is delivered through the screen reader as speech, but the semantic data is passed to the Tangible Scrollbar and delivered haptically. When a navigable element is encountered, the user can interact with it by pressing the navigation button. If this interaction updates the content (*e.g.*, navigating to a new web page), then the scroll thumb automatically repositions itself to the top of the page. Likewise, if the previous content is restored, the scroll thumb automatically repositions itself at the last known location.

In preliminary tests, the frequency of tactile feedback had the side effect of also communicating content size. Content size is typically delivered to the user visually by changing the height of the scroll thumb in the graphical content pane. Similarly, the Tangible Scrollbar reduces the distance of travel between haptic feedback events as the content size of the page increased, thereby providing the user with a sense of size.

### 3.3.2 Experimental Validation

I conducted an experimental study coupled with qualitative interviews to understand how the Tangible Desktop compares to traditional computer interaction for sighted and visually impaired users. I recruited 16 participants (5 sighted, 8 low-vision, and 3 blind) through word of mouth (S1, S2, S3, S4, S5), university services (M1, M2), and my field site, EmpowerTech. The study was conducted in a usability lab at the University of California, Irvine (n=5 all

| ID | Gender | Age | Technology | Braille Reader | Field Study |
|---|---|---|---|---|---|
| B1 | Female | 39 | Screen Reader | – | ✓ |
| B2 | Male | 35 | Screen Reader | – | ✓ |
| B3 | Female | 43 | Screen Reader | – | ✓ |
| B4 | Female | 36 | Screen Reader | – | ✓ |
| B5 | Female | 41 | Screen Reader | – | ✓ |
| B6 | Male | 37 | Screen Reader | – | ✓ |
| B7 | Female | 41 | Screen Reader | – | ✓ |
| I1 | Male | 52 | Screen Reader | ✓ | ✓ |
| M1 | Male | 18 | Magnification | – | – |
| M2 | Male | 20 | Magnification | – | – |
| M3 | Male | 35 | Magnification | – | ✓ |
| S1 | Female | 27 | Display | – | – |
| S2 | Female | 28 | Display | – | – |
| S3 | Male | 37 | Display | – | – |
| S4 | Female | 37 | Display | – | – |
| S5 | Female | 26 | Display | – | – |

Table 3.3: Technology represents the system participants typically use. Participant ID labels are coded according to technology; B group had less than six months of experience with their assigned technology, M group used magnification, S group were the sighted participants, and the I participant was a blind assistive technology instructor.

sighted users, n=2 low vision but legally blind users) as well as the field site in Los Angeles (n=3 completely blind users, n=6 low vision but legally blind users). Table 3.3 provides a detailed view of the participants, including the crossover between my study and field work. Eight participants identify as women with an average age of 35 (SD=8.79). Three participants (M1, M2, and M3) use magnification on their personal computer systems.

### Procedure

My fieldwork demonstrated that browsing the Internet is a critical stumbling block for people with visual impairments. Thus, my study used Internet browsing performance as the primary outcome measure for understanding the potential efficiency and experience of using a traditional screen reader as opposed to the prototype multimodal system. The current

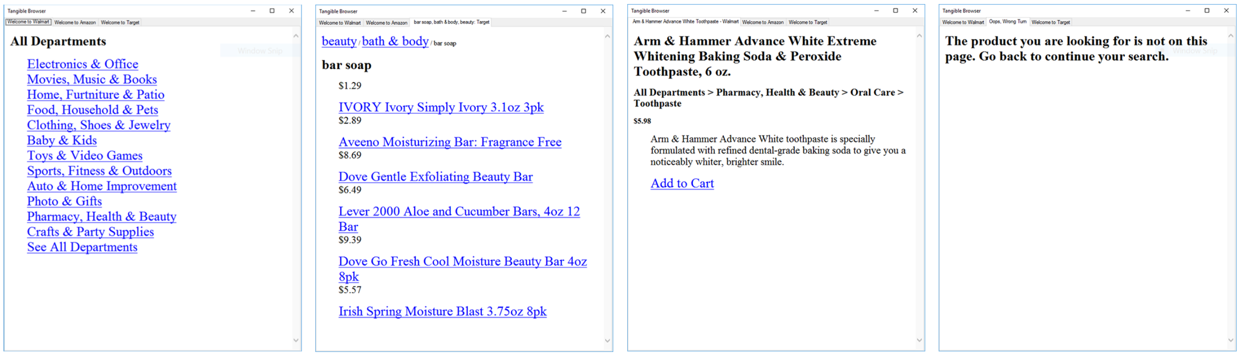



Figure 3.3: Sample of the study websites that were used, from left to right: top-level page for the Walmart site, product listing page, product detail page, and error page.

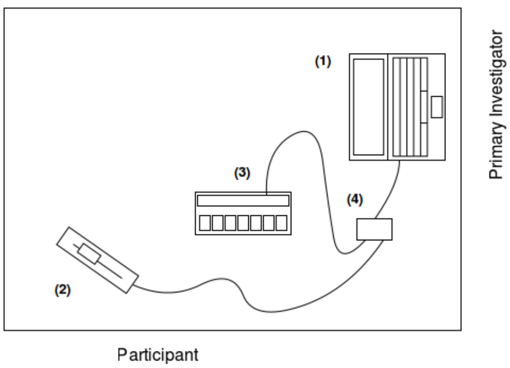



Figure 3.4: The study setup: (1) PI system running study software (2) Tangible Scrollbar (3) Tangible Taskbar (4) Controller interface

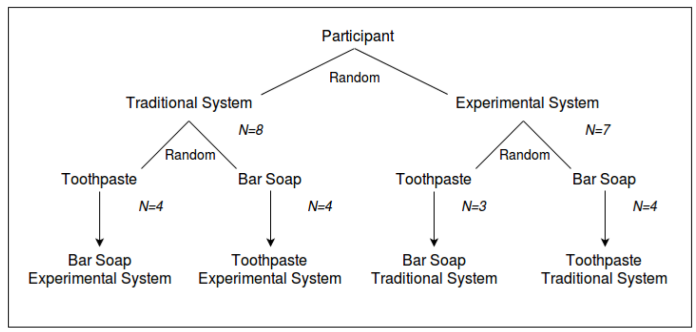



Figure 3.5: Flow diagram showing the experimental design. The starting system was randomized at the beginning of the study. Once the starting system was selected, the starting task was randomized. The study was completed using the remaining system and task. Participant B7 is excluded due to personal time constraints.

capabilities of the experimental system are limited to simplified HTML parsing, which made using live websites unreliable. I created three websites with navigational hierarchies modeled on the popular Internet shopping sites Walmart, Target, and Amazon. These sites were selected for their brand recognition as well as their incorporation into classroom assignments throughout my fieldwork. The sites that I built were modified to only include the navigational structure, an error page, a product list page, and product detail page (Figure 3.3). Each product list page contained five comparable products in addition to the target product for the task. However, because the site navigation was modeled after the commercial sites, the number of navigable elements accessible by the Tangible Scrollbar varied among target products and each website (see Table 3.4). All additional website content such as product recommendations, customer reviews, and advertisements were removed. The error page was displayed when a participant selected a product category link that did not lead to one of the task products. The two products that I used were selected from common household items (toothpaste and bar soap) available on all three Internet shopping sites. A product page was created for each target product that reflected the headings, description, and price used by the actual shopping site that sells the product. In instances in which the actual product prices were identical, I modified the price to ensure only one of my study sites had the lowest price. All the web pages that I created were constructed using accessibility standards including screen reader dependent HTML tag attributes.

Each participant completed two shopping tasks, one for each product, using the same three websites. One shopping task was completed with the experimental Tangible Desktop, and one task was completed using the participants' personal systems. All but one participant used their laptop computer to complete the task. Participant B3 did not own a laptop, so she used a familiar desktop computer in the classroom. Participants were allowed to use any assistive devices that they normally used in conjunction with their personal system. The starting system and the product to shop for were randomly assigned (see Figure 3.5). When participants performed the task using their computer they were asked to use their current typical web browsing environment.

I did not conduct a formal training session prior to the start of study. Participants were given one minute to locate the experimental system and explore it tactilely. Tactile exploration occurred naturally; I did not explicitly instruct them to do so. The PI then explained that they would move the thumbs on each device to switch between and navigate across the websites. The PI also explained the meaning of the haptic feedback indicators that the Tangible Scrollbar provided. Before beginning the task, the address of each shopping site would be spoken out loud to the participant so that they could preload each website. Once

| | Walmart | | Amazon | | Target | |
|---|---|---|---|---|---|---|
| | Toothpaste | Bar Soap | Toothpaste | Bar Soap | Toothpaste | Bar Soap |
| Home Page | 14 | 14 | 42 | 42 | 13 | 13 |
| Level 1 Page | 35 | 35 | 9 | 9 | 32 | 32 |
| Level 2 Page | 14 | 10 | 39 | 39 | 12 | 9 |
| Level 3 Page | 13 | 14 | 13 | 13 | 16 | 16 |
| Product Page | 5 | 5 | 5 | 5 | 6 | 3 |

Table 3.4: A list of the total HTML elements processed by the Tangible Desktop for each website and product used in the experimental study.

| HTML Element | Screen Reader Output | Haptic Feedback |
|---|---|---|
| <H1>...<H4> | Heading Text | Tap |
| <P> | Paragraph Text | Short Tap |
| <A> | Link Text | Vibration |

Table 3.5: The HTML elements controlled by the Tangible Desktop for the pilot study and their associated system responses. Screen Reader Output describes the type of content that was conveyed through computer audio using text-to-speech. Haptic Feedback describes the vibrotactile sensation delivered by the Tangible Scrollbar.

all three sites were loaded, the assigned product name would be spoken verbally. Participants would then be asked to find the lowest price for the product across the three shopping sites. The task was considered complete when the product with the lowest price was added to the shopping cart for the site.

When completing the task using the experimental system, navigation and interaction were performed with the Tangible Scrollbar while switching between the shopping sites with the Tangible Taskbar. I programmatically bound the three shopping sites to their own physical icons prior to the start of the study. Auditory feedback was delivered through a custom screen reader designed to only render speech for lexical information on a subset of HTML elements. Table 3.5 lists the elements that were controlled, the text that was output and the associated haptic feedback rendered by Tangible Scrollbar. As each website was explored, the Tangible Scrollbar automatically reoriented itself to the last known location of a given page. When an unvisited page was selected, the scroll thumb would move to the top of the device. On visited pages, the scroll thumb would move to the last known location prior to leaving the page. Upon completion of the study, participants were given an opportunity to describe their experience using the Tangible Desktop.

### Analysis

I used time-based task performance data and post-study interview data for my analysis. I measured overall performance by comparing the task completion times between each participant's normal computing system and their use of the Tangible Desktop. Sighted participants were averaged together and served as a baseline to isolate the task length and complexity. Low vision participants were excluded from the statistical analysis due to their preference for screen magnification. I also excluded the instructor from statistical analysis. He was a statistical outlier probably because he had been using a screen reader for more than 20 years and had a vast experience to draw from when navigating the experimental interface. Thus, participants B1-B6 were included in the statistical analysis.

The start and end of each study sub-session was extracted from the video log. Reviewing the video also enabled me to remove time corresponding to interruptions or errors that occurred. All but one participant experienced interruptions due to experimental system crash, phone calls, dropped Internet access, and questions about the task. In these situations, the I used video recording and screen recordings taken during the study to filter the delays from performance data (see Table 3.6). A delay time in milliseconds was captured by taking timestamps at the beginning and end of each interruption in the video recording. The delay time was then subtracted from the time in the data logs at the event at which the interruption began.

The group used for statistical analysis consisted of six blind or low-vision participants, recruited from the training class described in my fieldwork, and thereby primarily novice screen reader users. Two students did not complete the entire study (B1, B7). While B1 was able to complete the task using the experimental system, she was unable to complete it using her system. I have included B1's data in the results, but capped the participant system time at twenty minutes. By capping the participant system at twenty minutes, the likelihood of finding statistically significant improvement was reduced with the experimental system, thereby limiting the bias that such a choice might create. B7 was unable to complete the experimental system task due to personal time constraints. Thus, I only used her qualitative feedback in my analysis.

The experimental setup of my study introduced some limitations. Notably, the results indicate that the Tangible Scrollbar became harder to use accurately as the number of tactile interactions increased. The physical endpoints of the slide mechanism kept the distance traveled from top to bottom constant, but the distance between interactions decreased as the number of page interactions increased. During pre-study tests, I determined roughly 40-50 interactions to be the maximum that the device could traverse while maintaining a reasonable resolution. However, I did not take into account the various gripping strategies that participants engaged while using the Tangible Scrollbar. When participants used a lighter grip, the haptic feedback mechanisms were too strong and moved the thumb into an unintended position. Similarly, the custom screen reader I built for the experiment did not provide a mechanism for increasing speech rate in the text-to-speech engine, a common modifier used by screen reader users to increase text processing performance. This suggests that greater performance gains might be observed by adding support for increasing rate of speech.

Qualitative results were collected from post study interviews. Results reflect the experiences of all participants except for B7, whose time constraints prevented me from conducting an interview. I asked each participant to describe their experience using the Tangible Desktop. Additional probing questions were asked in response to participant answers that received. Follow up questions varied between participants but were patterned to elicit information on

| Participant | Time Correction (s) | Reason |
|---|---|---|
| B1 | 145 | Study system bug. |
| B1 | 12 | Participant stopped to tell story related to product used for shopping task. |
| B2 | 340 | Study system crashed. |
| B4 | 96 | Study system crashed. |
| B5 | 332 | Internet connection at study site dropped. |
| B6 | 35 | Participant stopped to make comments on product pricing. |

Table 3.6: A list of error corrections, including number of seconds that were removed from the final task time and the reason for the delay, that were made for affected participants.

| Participant | Participant System | Experimental System | Starting System |
|---|---|---|---|
| B1 | 20:00 | 09:37 | Experimental |
| B2 | 12:58 | 09:55 | Experimental |
| B3 | 14:56 | 08:26 | Participant |
| B4 | 16:32 | 12:37 | Experimental |
| B5 | 19:54 | 09:41 | Experimental |
| B6 | 14:43 | 10:03 | Participant |
| **Average** | **16:29 (16.49 mins.)** | **10:03 (10.05 mins.)** | |

Table 3.7: Task completion times in minutes for the six participants included in the statistical analysis. The starting system indicates which system the participants used to complete the first task (randomized).

the strengths and weaknesses of the system.

**Results**

My results indicate that screen reader users can work significantly more quickly using a multimodal system over traditional screen readers. Additionally, this improvement can be

| Variable | Coefficient (seconds) | Standard Error |
|---|---|---|
| **System type (1=experimental system, 0=participant system)** | **-367.90**** | **(84.19)** |
| Website order (1=Target first) | -116.60 | (152.30) |
| Intercept | 990.50*** | (56.76) |
| *Adjusted $R^2$ = 0.67, N = 6* | | * p < .05; ** p < .01; *** p < .001 |

Table 3.8: Results of linear regression model examining the effects of system used on study completion time, controlling for website order (system order was randomized).

seen within a very short time of using the system, with limited training. Comments from participants suggest that greater personalization and customization as well as additional experience with the devices would improve the experience. In this section, I describe the results of both statistical and qualitative analyses, highlighting how multimodal systems appear to improve user experience for nonvisual computer users.

On average, screen reader participants completed their tasks in just over ten minutes (n=10.05) with the experimental system as compared to over sixteen minutes (n=16.49) using their own systems, an improvement of 39.0% (*(16.49 - 10.05) / 16.49 = .390*), $t(5) = 4.94$, $p<0.01$. Table 3.7 details the time spent, in minutes, completing the task on both systems. Task completion time improved for *every* participant (see Figure 3.6 and Table 3.7). I randomized which system participants used first. Even though four out of six participants used the experimental system first (see Table 3.7), their performance was still faster on the experimental system despite having no prior experience performing the task. Because I did not randomize which website participants used first, I controlled for this factor in the regression model. I found no significant effect for website order (see Table 3.8). Even after controlling for this potential confound, I still found that using the experimental system was significantly faster for screen reader participants.

To test the breadth of such improvement, I also looked separately at results for sighted participants and those with minimal sight who use magnification to augment their screen reader use (see Figure 3.7). Sighted participants were able to complete the task faster

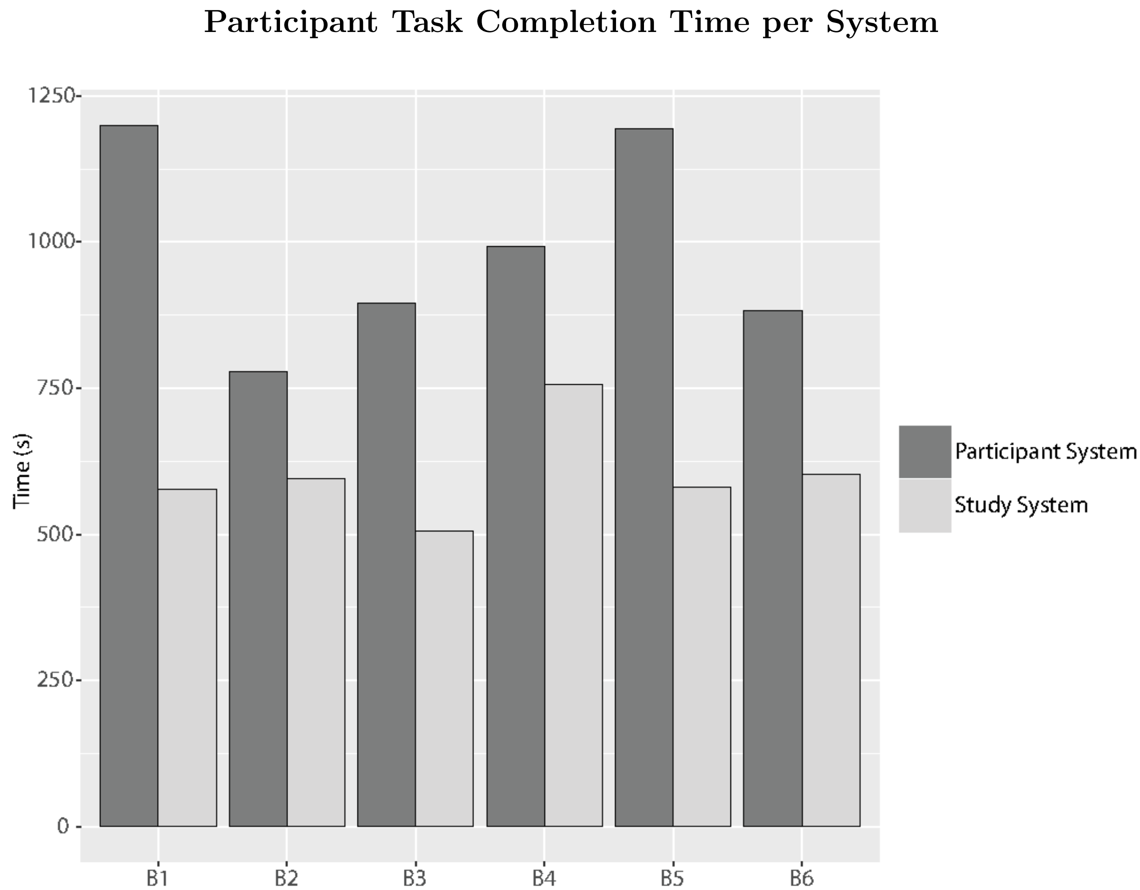



Figure 3.6: Task completion time comparison between the participant system and experimental system for each participant. Completion times are total elapsed time, so I have not included error bars.

using their traditional system. This was, as expected, due to their preexisting familiarity with visual navigation of web pages. I saw a similar but smaller trend for magnification participants.

Not only did the quantitative analysis of the tangible system improve computing performance when compared to the audio only interface of the screen reader, many participants described the Tangible Desktop as easier to use and understand. The ten participants who were visually impaired commented positively about the use of the physical devices. For example, one blind participant expressed excitement upon completing the experimental system task:

*"I think it is easier, because blind people, they do things by touch. So they are*

**Average Task Completion Time per Group**

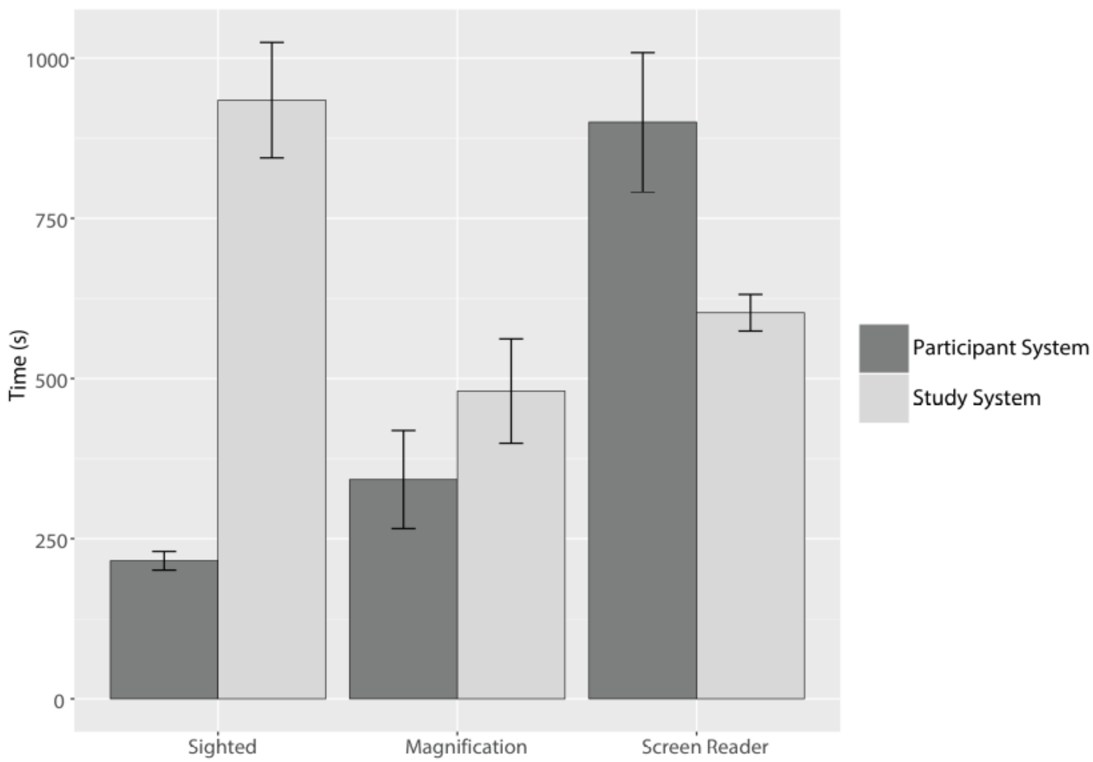



Figure 3.7: Average task completion time comparison between the participant system and experimental system grouped by technology.

*very sensitive by touch. So if they touch things they remember faster. For me I would remember faster."* - Participant B3

Even the participants who preferred magnification acknowledged the benefits afforded by a tactile interface:

*"Definitely more tactile than a software program, that was nice. There was enough pressure that I could tell what I was doing."* - Participant M2

There was an initial hesitation among all the participants as they familiarized themselves with the Tangible Scrollbar. The lack of pre-study training for the experimental system left several participants feeling like they could perform better if they performed the task again, as one participant explained at the end of the study:

*"In hindsight, if I was going to go through that again I would be better the second time around."* - Participant B6

The novelty of the device, from the automated movement of the scroll thumb to the haptic feedback that it provides, was a likely cause for this hesitation. Yet the brief period of orientation in comparison to the weeks and months of screen reader training the participants have received would seem to indicate that a tactile experience is less challenging to learn. For example, the first time participants performed a navigation event the automated reorientation motion of the scroll thumb caught them off guard. Their fingers were either still on or near the thumb when it moved, preventing it from completing its programmed movement. However, after a few navigation events, participants settled into a pattern of "click and hover," relaxing their hand slightly above the thumb until it had reached its destination.

Although all participants appeared to enjoy using the Tangible Desktop, some did experience confusion over the functionality of the navigation buttons on the Tangible Scrollbar:

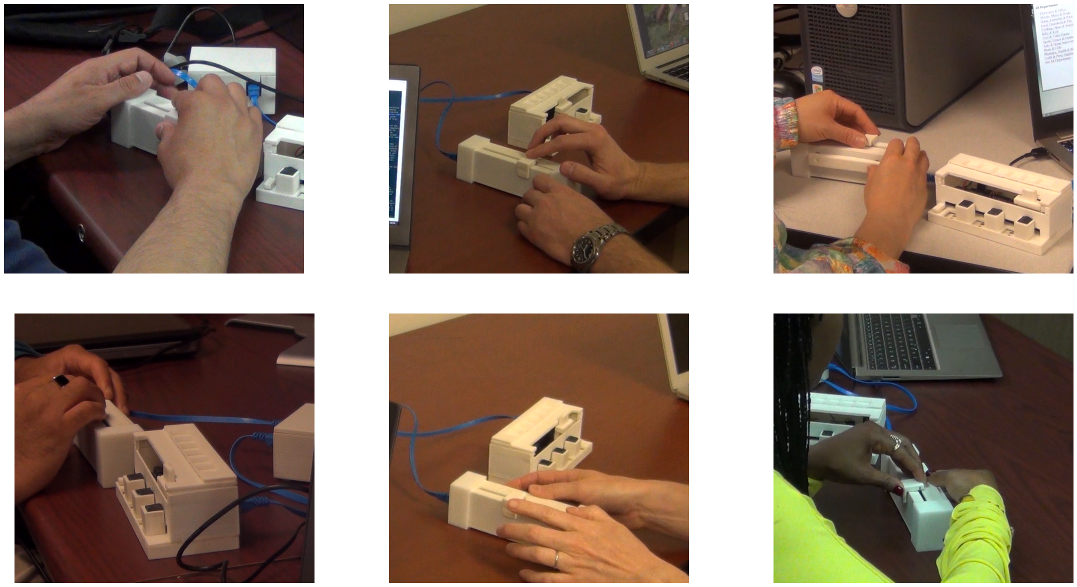


Figure 3.8: Examples of the different hand positions used by study participants.

> *"Like I don't know what the dot is or the double dot, but if I use it a couple of times, I would remember."* - Participant B3

I did not specify how to hold the Tangible Scrollbar at any point during the study. The grip style selected by each participant varied widely, and was never used as intended. The preliminary design for the Tangible Scrollbar assumed that it would be held in the palm of one hand and controlled with the forefinger and thumb of the opposite hand. Instead, the device was left in place on the desk as it was positioned prior to the start of the experimental system task (see Figure 3.8). This created some initial confusion over the functionality of the Tangible Scrollbar. The symmetrical shape of the device initially left participants confused over vertical movement direction (up vs. down) and forward and backward button navigation. Therefore, an orientation period typically occurred during the first few seconds of use.

During post-study data analysis, I observed differences in how seeking was approached between the Tangible Desktop and the participant system among the novice group. While participants generally proceeded with caution while using their system, the experimental system was used without hesitation once the initial functionality of the system was understood.

## 3.4 Simplifying the Information Stream with Activity Theory

Given the struggles I observed among students at EmpowerTech between goals and activities and the tasks they were asked to complete, I saw an activity theoretical analysis of my data as essential to a more complete understanding of the challenges and opportunities of nonvisual computing use. I began this analysis by working to understand the underlying motives (Leont'ev, 1978; Leontiev, 1981), needs, and desires of the blind and low vision individuals at my field site. For example, working in the EmpowerTech class, I observed a separation between the needs of the students and the stated goals of the class and the specific curriculum being taught. My analysis focused on the actions taken by students recorded through my fieldwork and experimental study, specifically the processes that they enacted with their goals in mind in service of their overall motives. I sought to understand the operations being conducted in service of these tasks and where the breakdowns were occurring when considered as a part of the larger whole. Taken together, this analysis indicates (1) the ways in which an activity-centric approach would be particularly supportive of nonvisual computer users, (2) some of the challenges surrounding the shift from application-centric to activity-centric computing that have made it hard to take up more broadly, and (3) multi-modal and tangible solutions that can address the needs of nonvisual computer users specifically, but also an activity orientation to computing more broadly.

Using an activity theory lens to examine my fieldwork accentuates the numerous challenges that blind and low vision computer users experience while attempting to carry out their activities. Through this lens, I pay particular attention to the relationship between actions


| *Place Marker List* |
|---|
| Menu (not visible) |
| Introduction |
| Discussion (not visible) |
| **More Information** |

Figure 3.9: In the computer display on the left, a typical web page is loaded in a browser and a link object called “More Information" has focus (denoted by the dotted box). The Place Marker List table on the right is a visual representation of user generated place markers that a screen reader stores on the computer. In this depiction, a place marker has just been created for the focused object, by issuing the four key combination: control + shift + MOD + k (MOD is the modifier key assigned by the screen reader). The remaining items in the list represent previously marked objects both on and off the viewable area of the web page. Issuing the key combination MOD + k or MOD + shift + k traverses the place marker list forward and backward, respectively. Notice that place marker order does match overall web page hierarchy.

(the tasks that humans consciously perform) and operations (actions which humans unconsciously complete) (Bødker, 1990). My analysis revealed challenges that emerged in three primary ways. First, nonvisual computer users often struggle to structure their data and services by activity. The usual techniques of establishing structure through visual arrangement are not available. Second, tracking activities, such that they can be paused and resumed, is challenging without visual markers. Support for tracking, to the extent that it exists in screen reader software, typically requires advanced knowledge of label and marker placement features with functionality varying across applications, screen reader software, and operating systems (see Figure 3.9). Finally, to understand what the system is doing without visual feedback, users must explicitly query and cycle through the services and data that are available, making the system's behaviors largely opaque. In the remainder of this section, I describe each of these challenges as they manifested in my fieldwork and as they relate to the activity theory literature. I then describe how computational systems might overcome these challenges if built with consideration for activities as the fundamental computational unit and nonvisual access as the core interface.

### 3.4.1 Organizing Computation by Activity

Humans do substantial work to translate their activities into tasks that can be completed within the hierarchical structure of a computing environment. In many cases, this work involves visual arrangements of materials. For example, when copying materials from one place to another, a sighted computer user might open two windows showing the file structure and place them side by side. Similarly, when comparing prices and selection across online retailers, a sighted shopper is likely to lay out multiple browser instances in tiles or tabs, enabling rapid scanning and comparing of options by clicking through the visible tabs. In my fieldwork (Section 3.1) and experimental study (Section 3.3.2), I observed blind and low vision computer users attempting these same activities but with radically different organizational structures and coping mechanisms. For example, consider the following vignette from my field work:

> The instructor teaches all of the students how to work with a file system, which is largely about learning keyboard shortcuts to "walk" the tree while listening to audio readouts of the folder titles and meta-data. For one assignment, students were asked to create a new folder and save a file to it. The task could be performed either by opening the file browser, creating a folder, and moving the file or using the save file dialog from within the word processing application that was used to create the file. The students encountered numerous challenges as they painstakingly attempted to complete the task. First, locating the appropriate parent directory, creating and naming a new folder, and relocating a file to the desired location must all be completed using keyboard commands to interface with multiple parts of the file system: operating system, file browser, and application. - *Field notes*

For the sighted user, the flexibility to perform basic file management operations at different contextual levels is advantageous (*e.g.,* creating a folder from within an application, rather than using a file browser.) However, the same flexibility incurs cognitive costs on visually impaired users who must balance command memorization with mental models that do not always match the system state (M. S. Baldwin et al., 2017). When mismatches occur, visually impaired users are forced to employ time consuming reorientation actions to complete their tasks (Vigo & Harper, 2013b). In the vignette above, students had to learn how to manipulate two conceptually identical mental models using distinctly different interaction patterns. Rather than remember the keyboard commands to save or move a file, they must first cognitively orient to the context (*i.e., file dialog or file browser*), then map the requisite commands that serve the context to complete the action to their mental model. For example, in the Microsoft Windows desktop environment a file dialog does not contain the menu structure present in the file browser. The absence of a menu leads to different focal order when using navigational keys to locate and traverse files and folders. Visually impaired users must negotiate these differences by either memorizing the relevant changes between contexts or through seeking behaviors such as exhaustive scanning (Vigo & Harper, 2013a).

Throughout my field work and experimental study, I observed a consistent pattern of file and application organization being carried out on the desktop. The sighted user benefits from a two-dimensional spatial arrangement in which related items can be grouped and clustered into meaningful visual relationships, whereas the nonvisual user relies on the desktop for reorientation and object location. The following observation from the experimental study (Section 3.3.2) highlights this interaction:

> After opening the first assigned web page the participant said, “Okay, hold on, let me return to the desktop", before receiving the name of the second web page to open. She pressed the windows key followed by the “m" key to return to the desktop, then repeated the key combination she had previously used to open a

new instance of Internet Explorer. — Observation from experimental study

Here the participant is setting up her system in preparation for the study. Her approach was to return to the desktop for each new web page using the same three commands. Her desktop reflected the files and programs that she used the most. By memorizing only a few shortcut commands she could quickly reorient to a familiar place–a tactic frequently employed by blind computer users (Vigo & Harper, 2013a). She could use arrow keys to navigate vertically and horizontally to the desired location. As a novice user, this behavior was likely a result of the training she had received up to that point, yet it also reveals a simplicity that students learning the file system did not have. The desktop, therefore, served as her activity, the files and application shortcuts located within represented the tools she needed to complete her tasks.

Unfortunately, because traditional desktop systems were not built to treat these behaviors as a single activity, but rather as a series of individual tasks, the responsibility of negotiating the activity hierarchy is placed on the user. In the previous vignette, the participant mixes command memorization, spatial memory, and seeking behavior, which lead to time consuming efforts to accomplish relatively straightforward tasks. Although this approach allows the participant to complete the task, it is effectively a workaround to an environment not explicitly designed for blind people (Boyd, 1990). Furthermore, not all blind people conceptualize their computing environments in the same way (Kurniawan & Sutcliffe, 2002). In general, most people, whether visually impaired or not, neither want nor find it easy to memorize commands, hence the downfall of DOS, UNIX, and other text and command based systems in the mass market (Grudin, 2008). In my fieldwork, even when users were able to memorize and use commands, they behaved inconsistently depending upon the actions they performed. An application and document approach is a system-oriented rather than an activity-oriented perspective, and it continues to be reproduced by accessible systems that simply mirror existing structures. Treating commands and the objects they act upon as a single activity with consistent behavior across them would likely improve experience for nonvisual computing users. In this case, the notion of the activity as an organizing structure becomes even more important relative to this need for consistent behavior. Additionally, when the system acts inconsistently, the challenges to receiving and understanding system feedback for nonvisual users exacerbates these issues, which I discuss in more detail in section 3.4.3 below.

Issues of consistency within an activity are not limited to the actions taken upon the various services and data within the activity. The fundamental metaphors themselves can, at times, break down within activities when the activities require use of multiple pieces of software or other computational infrastructure. For example, during the same assignment as the vignette above, one student struggled to understand the difference between the folders created in the music application iTunes and the folders he had just created on the desktop. This kind of confusion is of course not limited to blind and low vision computer users. However, without the visual feedback that drives this metaphor, the organizational underpinnings that differentiate a virtual folder (iTunes) and a physical memory-based folder (File System) make even less sense.

Finally, computer users often switch among documents, services, and applications within a single activity. A simple example might involve reading an email that asks for a meeting, switching to the calendar to check availability, and then switching back to the email client to respond. For sighted users, switching between windows is an expected and natural part of multitasking that can quickly grow to an unmanageable state in complex multi-window environments. To support this growing complexity, systems have introduced new ways of visually arranging information such as tabs, split panes, and virtual desktops. For the nonvisual user, however, these visual conveniences are lost, instead introducing additional complexity. For example, the difference between a browser window and a browser tab is conceptually insignificant for a nonvisual user, yet the behavior and interaction between the two require a different set of commands. Again from my field notes:

> The instructor often emphasizes the importance of learning how to use multiple browsers. As students were practicing browser commands later that day it became apparent how unnecessarily complicated this made switching from one website to another. Some students would open new browser windows to visit a different website, while others would open new tabs. One student asked me for help finding a webpage he had opened. The page he was looking for was open in a different browser application, yet he was using the key commands to cycle through open tabs. After a brief discussion, I realized that he had manually opened the Firefox browser to work on his assignment, then proceeded to select a link from his assignment, which opened the default browser (Internet Explorer) associated with hyperlinks in his system. - *Field notes*

Issues such as the one described in this vignette can be complicated further by skill and visual acuity. Although all of the participants that I observed were legally blind, several of them relied on their residual visual abilities to support their computer use. When participants were asked to open three web sites using their personal computer during the experimental study (Section 3.3.2, nearly all of them completed the task in a different way. The behavior of one novice participant was captured in my study observations:

> When the task required her to switch from one web page to the next, she was unsure which keys to press to make the switch. The PI instructed her to press the alt+tab key combination. When it came time to move to the third web page, she used the same key command only to be surprised when she was returned to the first web page again. The PI explained that to move to the third web page she would need to press the tab key twice. - Observation from experimental study

This instance is likely a familiar experience to sighted users of windowing interfaces, but without the visual cues displayed on screen to indicate which window is activated, nonvisual users are left to investigate further to resolve the breakdown between expectation and outcome. Here the participant expected a circular switching behavior where each key press moves to the next available window. If she had been switching between tabs, her expectation would have been met, but because she was switching between windows, the behavior did not match her mental model. A different participant, an experienced low vision computer user, explicitly stated his preference for browser tabs. Observing him carry out the web page setup task, he split duty between his screen magnifier, which kept the tabs enlarged in a narrow window at the top of his screen, and a handheld magnifying glass used to scan the content of the web page. This configuration simplified his interaction with the screen magnifier, constraining its use to horizontal movement, while he scanned the page using a physical magnifier. For this person, rather than contend with the complexities of maneuvering a magnifier around the screen, he opted to incorporate a physical tool to mediate task completion.

The challenges to organization of data, services, and applications for nonvisual users within a single activity indicate that existing computational abstractions are not sufficient for blind and low vision users. Systems in the existing research literature that have attempted to take an activity centered approach (Cornet, Voida, & Holden, 2018; Bardram & E., 2005; Bardram et al., n.d.; Voida et al., 2007; Voida, Mynatt, & Edwards, 2008; Voida & Mynatt, 2009; Rattenbury & Canny, 2007), rely on deep integration with the operating system or customized software to adapt the application-document model to activity. For nonvisual users, the adaptation layer already exists in the form of tools, such as a screen reader or screen magnification software, which restructure computational information into a more suitable format (*e.g.,* speech, sonification, and magnified graphics). Therefore, introducing activity to a nonvisual system is a matter of rethinking how existing tools present computational information, rather than introduce additional complexity through new layers of software.

### 3.4.2 Activity Tracking

Activities tend to evolve over time as people change their goals and the objects they act upon (*e.g.,* a line of text in a document, image on a web page, or higher order element like an application) and adapt to user actions. In computing environments, this behavior can result in the calling up and subsequent abandonment of a variety of documents, services, and applications. An object that played an important role at the start of an activity may never be used again. Similarly, some objects may not find utility until the final steps leading to activity completion. Managing these variations throughout the life of an activity is supported in the traditional desktop through visually oriented design cues, such as recognition and spatial arrangement, and virtual desktops. The absence of these visual cues creates a variety of challenges to the orientation of nonvisual users within their applications and data (Abdolrahmani & Kuber, 2016; Borodin et al., 2010; Lazar et al., 2006). Despite progress in the quality of screen readers and other accessibility tools, through fieldwork, I continued to see orientation within, tracking, and pausing of activities as significant challenges. In particular, students at EmpowerTech regularly struggled to restart after a pause, engaging in seeking behavior or simply restarting the entire activity from the beginning, as described in the following vignette from my field notes:

> The instructor had finished lecturing for the day and freed the students to start working on their assignments. I was observing one student who was working on a web page navigation task. After a brief hesitation she removed her headphones and asked the instructor for the command to open a new web page. Upon receiving a response, she entered the command “control-w", even though the instructor had told her “control-t", before putting her headphones back on. The action had mistakenly closed the browser, so when she proceeded with her task, the system did not respond as expected. When I asked her what happened, she told me that her computer was acting up again and probably needed to be

restarted. - *Field notes*

As demonstrated in this example, the roles that each artifact plays, and its use (or disuse) at various times are critical to contextualizing the activity, but without the system's awareness that an unattended action was executed, the user can easily get lost. These challenges are not unique to nonvisual users. In more complex activities, sighted users also struggle with task reconstruction. For example, while writing a paper, the author might simultaneously reference a spreadsheet in a separate window on the desktop. In the traditional application-document metaphor, from a system perspective, these two documents have no knowledge of each other; yet from the user's perspective they are core parts of the same activity, which she may express by laying them side by side in the visual computing environment. Once these applications are closed, the meaningful connection between the documents is lost, requiring the author to reconstitute the activity when edits are required. Thus, task reconstruction through activity tracking is a central element of various activity-based systems for sighted users. The Kimura system, for example, tracks the unique memory handles that the operating system assigns to application windows (Voida et al., 2007). Whereas in Bardram's system, applications are built on top of a framework that manages system state throughout an activity (Bardram & E., 2005). In both cases, the changes that occur to applications and documents (i.e., opening, closing, focus, and location) are captured and logged into a data store, effectively providing users with the ability to navigate backward and forward throughout the activity lifecycle. The consequence of designing these systems around sighted activity, is that they must rely on visualizations to allow the user to interact with the history. Visualizations are often difficult to translate into audio, requiring descriptive text as well as tabular representations of data that can be traversed by a keyboard. The additional complexity that these additional steps introduce makes the activity tracking format used in these systems difficult to use in a nonvisual environment.

In the vignette above, the student missed a notification that she had inadvertently closed a window because she did not hear the auditory cue–a subtle, but critical failure resulting from the ephemerality of the audio interface that does not match to the temporal patterns of the user's activities. Alternatively, an activity-centered system should support these activities–and the underlying cognitive behaviors required to maintain them–by tracking key contextual indicators and alerting the user at a time appropriate to the context of use and larger goals of the activity. Similarly, in an activity-centered system, the action of opening and closing web pages connects them through their use during the activity. By tracking and recording use, these changes can be reconstructed at any given moment throughout the life of the activity. Tracking can alleviate the challenges that surround interruption and error recovery for nonvisual users.

### 3.4.3 Operationalizing Actions

Leontiev organized human consciousness into a three level hierarchy of activity, action, and operation (Leont'ev, 1978). Activities are carried out through actions, determined by the individual goals of the human subject. Actions are fulfilled through unconscious operations, which reflect the human subject's natural attributes. The relationship between actions and operations is described as fluid, where actions can become unconscious operations through the natural internalization that occurs through practice, and operations become conscious actions through externalizing processes such as breakdowns (Bødker, 1990). In computational terms, we can conceptualize the relationship between action and operation by observing how one might learn to use a computer mouse. At first, mouse use might be action oriented, where a novice user *consciously* interacts with its various buttons and controls. Eventually, through practice, mouse use becomes operationalized, moving from a conscious to *unconscious* operational state. As an operation, focus is shifted from use of the mouse itself to performing actions that the mouse supports, only returning to an action when an attribute of the mouse changes (e.g., a broken button or dead battery). Proponents of activity-centered computing (Voida et al., 2007; Bardram et al., n.d.; Kaptelinin, 2003) have demonstrated that structuring computation around activity can lead to improved support for operationalizing actions. In practice, however, configuration and management is a common point of difficulty for users. Tasks integral to supporting activity, such as “tagging" were commonly avoided (Voida & Mynatt, 2009). Similarly, parts of the systems that required users to change their pre-existing practices were met with resistance (Voida & Mynatt, 2009; Bardram et al., n.d.). By restructuring the application-document metaphor, activity-based systems are effectively adding an additional layer of complexity that users must navigate.

From a nonvisual interaction perspective, the negative effects of restructuring application to activity is not surprising. The process of translating graphical information to auditory information for blind and low-vision users is a significant factor in keeping nonvisual interaction at the level of conscious action (M. S. Baldwin et al., 2017; Vigo & Harper, 2013b; Lazar et al., 2006). Poor translation is both burdensome and inefficient, making it noticeably intrusive to the subjects I have observed. Students as well as their blind instructor continuously faced challenges with hardware and software. As I captured in my field notes:

> One low-vision student was using screen magnification software to perform the tasks being taught by the instructor. Over the course of the lecture, her computer became increasingly unresponsive, freezing for 15-30 seconds between actions. Unable to keep up with the rest of the class, she asked the instructor for assistance. Together they spent the rest of the class time attempting to solve the problem. – *Field notes*

Technical issues like the one captured here occur frequently. In my classroom fieldwork, these issues resulted in either time lost for the student, who was dealing with the problem, or for the entire class who had to wait for the instructor to resolve it. In a workplace, these issues can interfere with accomplishment of mission critical tasks in the worst case scenario or just an employee being less efficient or perceived to be more troublesome in the best case scenario (Branham & Kane, 2015). In some cases, such issues cannot be easily resolved by the user or people nearby, leading to the entire system needing to be set to the side until a specialist can engage. In my classroom-based fieldwork, the varying inconsistencies of the tools students were attempting to learn were a constant source of distraction.

> Today I was asked to help one of the more advanced students in the class get her new Windows 8 computer setup with the Firefox and Chrome web browsers. I provided some verbal directions on where to go to download each browser, but otherwise left navigation and interaction to her. The Firefox browser downloaded and installed without issue, however, the Chrome installer interface was not detected by her screen reader, leaving her to conclude that the installation did not work. Although I could see that the installation was functioning properly by observing the installer progress bar, since her screen reader did not register the progress control, she was left without any feedback. – *Field notes*

Wildly different experiences while performing the same action can leave nonvisual users attempting to solve problems that do not actually exist. A common tactic I observed was to start over. In the example provided here, I intervened by manually providing the required auditory feedback, preventing the student from following her instinct to start the download over again. In many other instances, however, students opted to reboot the computer. An extreme, but effective, resolution to the problem which was often viewed to be successful, not because it fixed anything, rather that it allowed students to reset their frame of reference within the system.

The difficulties that I observed with screen reading and screen magnification tools stand in stark contrast to the students' use of the Victor Reader that I described in section 3.1. While the use of screen translation tools were continuously conscious interactions, students operated the Victor Reader without issue. They spoke positively about their interactions with the Victor Reader and made regular use of it. One student in particular made a habit of transferring all of her documents to the device to read, rather than relying on her desktop screen reader.

Use of the Victor Reader yields valuable insights into the frictionless interaction that nonvisual technology can provide when translation of a graphical interface is not a dependency. Realistically, nonvisual interaction with graphical systems simply is not possible without some amount of translation to alternative modalities. However, the model that is used can be shaped to fit the more natural practices of human activity, making the process less complex. For example, at the level of activity, there is no need to consider where a document is stored or how it is saved; those details are managed by the system. The benefit of this reduction in cognitive tasks grows with the complexity of activity, essentially unifying the actions of saving and storing across many documents into one single action. To make systems truly lead to unconscious operations for nonvisual users, however, they must be reconsidered in light of the activities they are meant to support. Shifting the translation of information between modalities from a model of individual elements within documents within applications and folder hierarchies to one that considers information flows across activities could enable this kind of invisibility in use for blind and low vision users.

### 3.4.4 Case Study of a Tangible Activity-based Platform

In this section, I have described how an activity theoretic analysis of nonvisual computer use reveals new ways of approaching computer interaction for blind and low vision users. My analysis suggests that an activity-centric model can improve the computing experience by transferring critical organizational and structural task management behaviors from the user to the system. To understand how a nonvisual, activity-centered system might function

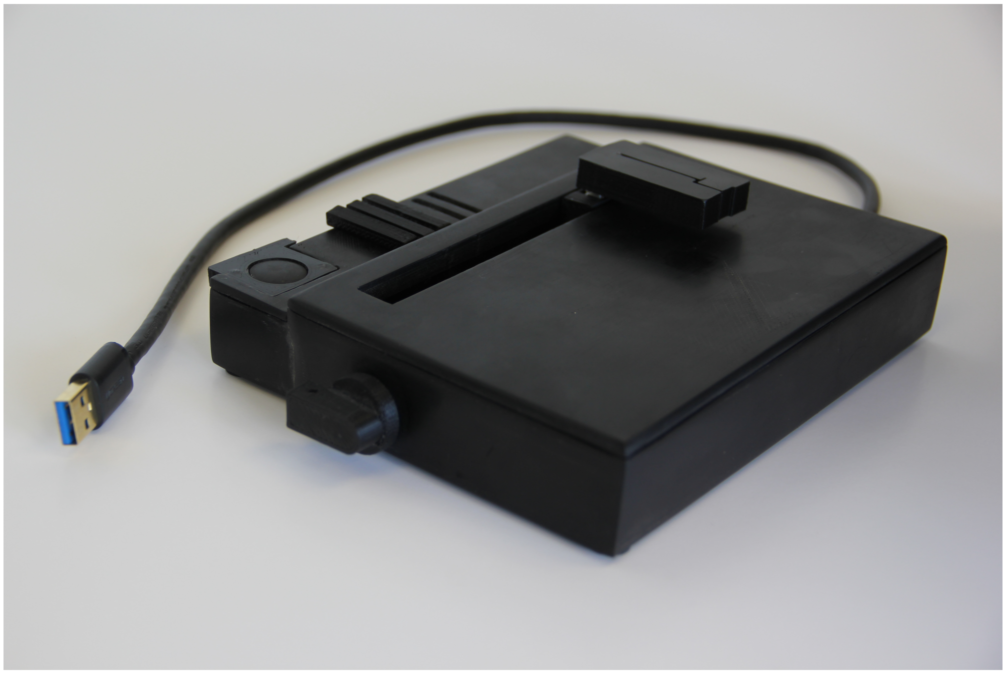


Figure 3.10: The Kinesthetic Interaction Device (KInD). On the left side, the activity context dial is in the application content position. The thumb sits in the center of the device, ready to move to the left or right.

in practice, I designed an exploratory platform for basic file and application management. The platform consists of an activity oriented application programming interface (API) and tangible interaction device called the Kinesthetic Interaction Device, or KInD (see Figure 3.10). The API serves as an intermediary layer on top of the Google Drive cloud based file management system (Google, 2019) and can be controlled through either a command line interface (CLI) or KInD. KInD shares many of the characteristics that I introduced with the Tangible Desktop (see Section 3.3.1), but goes further by providing richer haptics and increased portability. KInD can be programmed to support different types of interactions that benefit from proprioceptive and tactile input and output. The activity API provides text-to-speech output to communicate information auditorily. In this section, I first describe the functionality of the activity API, including an overview of its CLI and KInD interfaces. I will then conclude with a discussion of my findings from two design sessions conducted to elicit feedback from blind and low vision computers users.

### Activity Interaction

The activity API overlays a set of activity-oriented interactions on top of Google Drive (Google, 2019) to support a platform independent interface for nonvisual activity. To demonstrate one way that activity-centered concepts can benefit nonvisual computing, the activity API implements a small subset of the types of interactions that a complete nonvisual activity-centric platform might require. The API is accessed and manipulated through two complimentary interfaces, a command line interface and a prototype computer peripheral, both of which I describe in detail in the next two sections.

**Activity CLI** The Activity CLI provides a natural language point of interaction for the activity API. The CLI was designed to be used as a standalone interface or in conjunction with either the Google Drive web interface, or KInD (described in the next section). There are five primary commands interpreted by the CLI to control the API: create, move, list, tell, and help. The create and list commands are combined with additional parameters recognized by the API to perform operations. For example, the create command is combined with an application type to create a new file in Google Drive (*e.g.*, “create spreadsheet" or “create document"). The move command enables file movement from activity to another. The remaining commands are used to provide context about system state such as which activity is active, the applications within the current activity, and activities in the system.

**KInD: Kinesthetic Interaction Device** Similar to the Tangible Desktop (see Section 3.3.1), KInD utilizes kinesthetic resistance, vibrotactile touch, and proprioception to translate input and output of computational information without audio descriptions. However, unlike the Tangible Desktop, which was distinctly designed to support an application-centric

| | Activity Contexts |
|---|---|
| *Context Level* | *Description* |
| 1 | Activities |
| 2 | Applications and Documents |
| 3 | Application Menus |
| 4 | Application Content |
| 5 | Activity History |

Table 3.9: Table of Activity Contexts

interaction model, the design and interaction models supported by KInD have been structured to support activity-centeric interaction. KInD makes use of three physical interactions that enable users to accomplish their goals nonvisually. First, KInD represents activities through tactilely differentiated tangible tokens to support computational organization in the physical world. Second, KInD tracks the systematic changes that occur during an activity, enabling a user to tangibly move between tasks temporally rather than through a method of auditory repetition used by traditional screen readers. Finally, KInD implements tangible spatial structure to encourage operationalization through proprioceptive memory rather than auditory seeking (Vigo & Harper, 2013b). Each physical interaction are described in greater detail below.

**Activity Management** Activity *tokens* (see Figure 3.10) are tangible icons that represent an activity. KInD only requires that a token be placed on it for that token's activity to become active. At that point, the associated documents and applications become available for use. An advantage of this approach is that non-active tokens can be physically arranged in whatever manner the user desires–allowing one's proprioceptive abilities to be leveraged for activity organization (*e.g.*, work on the left, entertainment on the right). Transferring the representation of activity to the physical world has the advantage of introducing the effects of unconscious operation that I previously described through classroom use of the Victor

Reader (see Section 3.4.3). Like the uniquely shaped buttons on the Victor Reader, activity tokens can convey meaning through shape, size, and text. For example, the sample tokens created for this platform could be 3D printed with braille text or different shapes, allowing tokens to be molded to the mnemonic strengths of the individual, reducing the burden of command recall.

**Interaction History** KInD is situated between the intentions of the user and the execution of those intents in the computer system, allowing the infrastructure to capture input and output events as they occur. Events are logged in the host computer's file system to provide a history of interaction, and associations between activity token and entities. The primary responsibility of the interaction history is to support the dynamic nature of an activity. As system entities are opened they become an active part of the activity. When they are closed, they move out of the activity space, but remain in its history. This allows the platform to restore the last known state of an activity upon activation. Furthermore, it introduces a historical archive of the activity that can be used for retracing steps and error recovery (*e.g.,* accidentally closing a document).

**Contextual Change** To accommodate the movement across activities as well as their associated entities, KInD functions within different contexts (see Table 3.9). The top context sits at the top level of an activity, supporting the movement of entities in and out of an active activity. The next context functions within an activity, providing support for movement across entities associated with the activity. The third context supports control at the application level, allowing KInD to be used to traverse and select features of the active application. Finally, the fourth context supports movement across the content of the application (*e.g.,* traversing a web page or document). Switching between contexts is managed by a

| | *Task* | *Category* |
|---|---|---|
| 1. | Place a token on KInD. | Task Completion |
| 2. | Using the context dial, select application context. | Change |
| 3. | Open a document. Then open spreadsheet. | Task Completion |
| 4. | Move context dial to activity. | Change |
| 5. | Move document to the right side. | Change |
| 6. | Remove the current token, place a new token. | Task Completion |
| 7. | Move context dial to application. | Change |
| 8. | Open email. Then open spreadsheet. | Task Completion |
| 9. | Move context dial to activity, explore. | |
| 10. | Move context dial to application, explore. (switching mode) | |
| 11. | Pin email. | Change |
| 12. | Switch token | |
| 13. | Unpin document. | Change |
| 14. | Move slider to find email. | Change |
| 15. | Return to original document. | Change |

Table 3.10: Design session tasks

physical rotary dial on the left side of KInD. The dial uses detents and audible confirmation to indicate when it has been rotated into a new context. As with the activity tokens, the physical manipulation enables proprioceptive abilities to quickly identify the desired context.

**Virtual Docks** Organization within an activity, potentially involving tens of individual entities (*e.g.*, application file, email, or web page), is too complex to be handled through physical manipulation of an activity token. Instead KInD implements virtual docks to sort and store the entities assigned to an activity. Docks are virtual, in that they are managed through a software subsystem, but are accessed through physical interaction using the KInD slide bar. As new entities are added to an activity a new dock is created. A haptic detent is generated by the slide potentiometer to indicate traversal across an entity. If an activity contains two entities a detent is felt at the midpoint of the slide traversal, with three entities two detents are felt equally divide along the slide, and so on. Individual entities can be arranged within these slots in a manner most suitable for the user. For example, a user might prefer to keep web pages in the right most slots, a presentation document in the far left, and everything else in between, retaining the advantages of proprioceptive recall found in activity tokens. Just as the flow of work can be dynamic, entities can flow in and out of activities depending on user need. For example, a user might want to take a break from a work activity and browse the news, a goal preferred not be associated with the work activity. KInD supports this type of activity flow by enabling docked entities to be pinned outside the activity specific scope. Pinning enables individual entities to be moved between activities or simply detached from an activity entirely.

#### Exploring the Activity API through Workshops

I conducted two design sessions with 6 blind and low vision participants (see Table 3.11) to evaluate how the Activity API and KInD align with user needs and expectations. Design sessions were organized into four phases to probe participant reaction to the activity-centered interaction model. In phase one, participants were asked to describe the types of tasks they use their computer to complete. Responses from phase one were used to personalize the task structure for the remaining phases. In phases two-four, participants were asked to carry out a series of file management tasks (see Table 3.10) using the Google Docs web interface, the Activity CLI, and KInD, respectively. Phase four began with a training period, where I described KInD's features (see Figure 3.11). Participants were encouraged to interact with KInD, to familiarize themselves with the tactile, haptic, and proprioceptive features utilized during task completion. Once each participant acknowledged they were comfortable with KInD's functions, each file management task was read aloud as the participant performed the task. After the fourth phase was complete, a semi-structured discussion with participants was conducted.

The design sessions provided useful insights into the structure and design of the activity

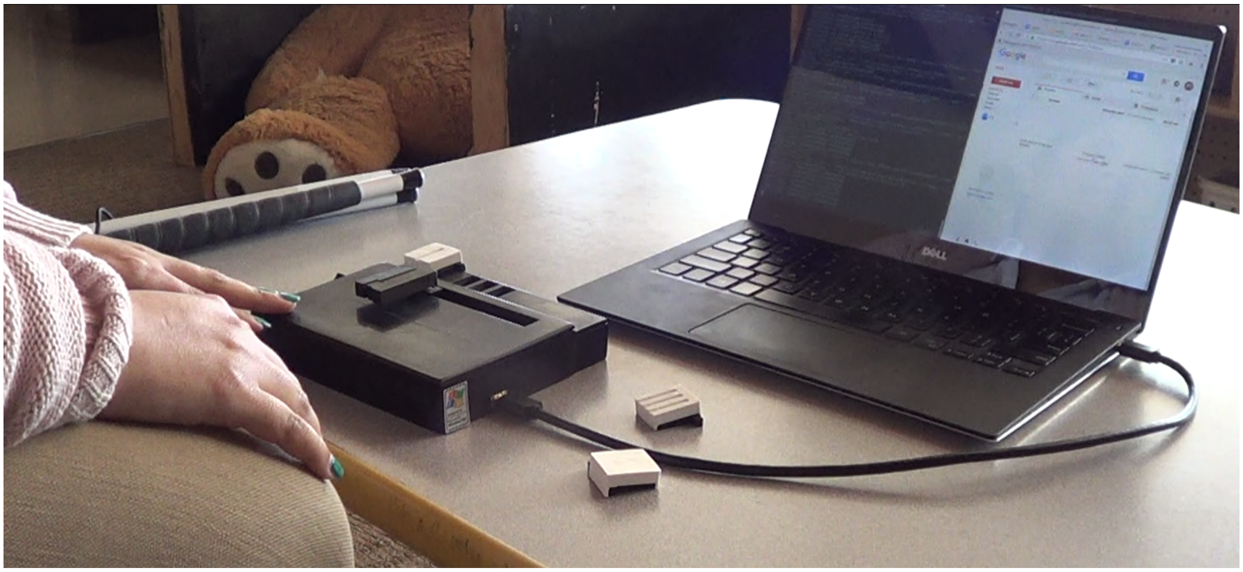


Figure 3.11: An example of the setup used during design sessions with KInD. Activity Tokens have been positioned near KInD by the participant for quick retrieval.

platform. Overall, participants responded positively to the CLI and KInD interfaces to the API. While all participants moved fluidly through task completion, each expressed a unique perspective on their experience, ranging from indifference (P6) to enthusiasm (P2,P3). P6, who primarily uses her computer for audio editing, admitted that her specific use case made it difficult to imagine how orienting towards activity would change her workflow. Whereas P3, whose job required her to record and manage a variety of different meeting notes, found the activity structure to be a positive change from her existing practices, comparing and contrasting the activity platform with her typical workflow. Other participants adopted similar patterns to explain how they thought the CLI or KInD might improve their own task completion. For example, P1 pointed out that he prefers to use his smartphone over desktop for most computing tasks due to its simplified interaction model, yet explained that KInD would make it easier for him to engage with his daughter on their desktop computer. P4 and P5 expanded on P1's experiences by explaining that the spatial movements used by KInD to manipulate activities aligned more closely with their smartphone interaction patterns. Similarly, P3, P4, and P5 described their challenges with command memorization and content traversal, noting that the natural language of the CLI interface (P3), and the

| | | **Participant Data** | | | |
|---|---|---|---|---|---|
| *ID* | *Session* | *Experience* | *Gender* | *Visual Impairment* | *Age* |
| P1 | 1 | Proficient | Male | Low Vision | 35 |
| P2 | 1 | Expert | Male | Low Vision | 43 |
| P3 | 2 | Proficient | Female | Blind | 28 |
| P4 | 2 | Expert | Female | Blind | 30 |
| P5 | 2 | Expert | Female | Blind | 29 |
| P6 | 2 | Proficient | Female | Blind | 30 |

Table 3.11: Study participants

tactile cues of KInD (P4, P5) could serve to lower dependency on command memorization.

Interestingly, P2, P4 and P5, who self-rated as expert computer users, also expressed benefits of the activity platform in terms of improvement for *other* users in addition to themselves. P2 envisioned the activity platform as an onboarding tool. He stated that although the feature set was small, its constrained options could help prevent users from getting lost. P1 agreed stating that he found KInD less intimidating than his desktop computer. These types of responses evoke sentiments of sociality similar to those observed by Morrison et al.. They view their own experiences with desktop computing to be sufficient, but recall their own difficulties in learning, as well as the challenges faced by others, as problematic enough to desire an alternative.

As a preliminary investigation into the application of activity theory to a nonvisual tangible interface, this case study provides one potential path forward. Although the informal structure of the design session limits claims about the effectiveness of an activity-centered interaction model, the responses to the simplicity of the system are compelling. Participants viewed the activity system as both useful and simplistic enough to support their own work as well as the work of novice users, aligning with the action–operation process discussed in Section 3.4.3. Therefore, a nonvisual model oriented towards activity can support a constrained interaction space, but does not prevent more experienced users from accessing the lower level system. Just as the Activity CLI and KInD provide a constrained interaction space on top of Google Drive, the full featured web based interface remains available for users who need or prefer it.

## 3.5 Summary

In this chapter, I have introduced two systems that restructure how computational information is represented in a nonvisual context. Through the Tangible Desktop, I demonstrated how physical peripheral devices can be used to represent visual desktop metaphors, removing the need to communicate semantic information auditorily. My activity theory analysis of blind and low vision computer use established how an activity-centered interaction model can be applied to transfer many audio dependent tasks from user to system. I demonstrated one possible approach to using this model through the Activity API and the KInD peripheral. I view this work as a first step towards adopting new ways of thinking about how nonvisual computing systems might be designed to lower barriers to entry for blind and low vision computer users.

However, I remain cognizant of the limitations of the data used throughout this work. Both the field work and experimental data used to inform my analysis, for example, derive from novice computer users in the midst of learning assistive tool use. While these findings suggest that tangible and activity-centered nonvisual computing can improve the computing experience for the visually impaired community, the inclusion of a majority of novice users leaves questions surrounding the effectiveness of these approaches with more advanced users.

Related literature describes experiences similar to those that I have reported across a range of skill levels, suggesting that a broader generalizability of activity-centered principles is viable. Billah, Ashok, Porter, and Ramakrishnan, for example, identified similar technical challenges to those I observed including inconsistencies across software tools and operating systems that forced users to learn multiple paths towards completing a task (Billah et al., 2017). Similarly, Potluri et al. place discoverability and navigation as two high-level challenges blind and low vision programmers encounter with their tools (Potluri et al., 2018). Both Billah et al.'s and Potluri et al.'s work included participants from a range of professions and skill levels, indicating that the challenges I articulate in this chapter manifest in similar ways, regardless of experience. Furthermore, the diversity of efforts to uncover workarounds to desktop computing and ICT more broadly, including those reviewed in Chapter 2, highlight the value of rethinking the nonvisual computing experience.

# Chapter 4

# Augmenting Assistive Technology Through Shared Mixed-Ability Interaction

Results from the Tangible Desktop and KInD studies presented in Chapter 3 demonstrate how isolating the semantic and lexical elements of an auditory stream enables information to be broadcast simultaneously to both kinesthetic and auditory channels, thereby reducing the amount of auditory information that a nonvisual user must process. While this strategy works well in computational environments, the semantic—lexical divide might not be suitable for other types of information streams. In some cases, semantic information is represented by naturally occurring auditory phenomena and not desirable for translation to other channels. A walk in nature, for example, where semantic information is represented by environmental sounds, would radically change a person's experience if not interpreted auditorily. In real-world contexts like leisure, fitness, and social events, the process by which auditory information should be augmented is less clear. While tangibility has been used to augment physical activities such as swimming (Muehlbradt, Koushik, & Kane, 2017), running (Avila Soto et al., 2017), and general outdoor navigation (Dakopoulos & Bourbakis, 2010) for people with visual impairments, these approaches apply tangibly communicated information in service to the intervening technology, rather than towards a structured taxonomy.

To understand the relationship between assistive technology and auditory information more broadly, I conducted a multi-year field study and design engagement with a group of blind and low vision outrigger paddlers. Together, we cooperatively developed a remotely controlled steering system for a one-person outrigger canoe (M. S. Baldwin, Hirano, Mankoff, & Hayes, 2019). Following a participatory design process, we leveraged public insight and *in situ* evaluation, to understand the role of auditory information in the paddling experience. Our cooperative design work led to the creation of an assistive navigational control system shared between a sighted and visually impaired dyad.

In this chapter, I present the results of the field work, iterative development of the Cooperative Outrigger Paddling system (CoOP), and long term evaluation of CoOP. I demonstrate the various ways that our *in situ* design process fostered investment and interest across the sighted and visually impaired paddling community, revealing numerous design insights that inform how to navigate the use of auditory information in assistive technology design. Results are presented according to four high-level themes that emerged from analysis of the data and inform the research for this dissertation: 1) public-facing co-design supports development of assistive technology; 2) do-it-yourself (DIY) approach supports co-design; 3) the physicality and interdependence of multi-modal design presents new challenges and opportunities for designers; and 4) auditory information influences the design and operation of assistive technology.

## 4.1 Methods

The work presented in this chapter comprises the results of an ongoing observational study and co-design effort with the Makapo Aquatics Project[1]—an organization that focuses on providing paddling opportunities for individuals with visual impairment. The goal of this study was to investigate questions of how mixed-ability individuals interoperate through the design and use of a shared assistive technology. I report on the collection and analysis of data gathered over the course of a twenty-five month period working alongside a mixed-ability team of athletes, coaches, and support crew. Activities during this time were centered around paddler training, outrigger paddling races, and design and evaluation of the CoOP system. During evaluation and practice sessions, I accompanied the team in an observational role as a passenger on one of two support boats typically deployed during training or CoOP evaluation sessions. The focus of my observations was multifaceted. I was primarily interested in understanding the relationship between coach and paddler, how their interactions were mediated by CoOP, and the role of auditory information. I also attended to the physical interactions between participants and the CoOP system, seeking to identify the types of design constraints that arise in an assistive technology shared between mixed-ability pairs. Data was collected through workshops, unstructured interviews with athletes, coaches and support crew, observations, and digital artifacts. The study was conducted at the Newport Aquatic Center, a facility dedicated to supporting paddle-based water activities such as rowing, kayaking, paddle boarding, and outrigger canoeing. The Newport Aquatic Center serves as a home for numerous external groups, including Makapo.

[1]The Makapo Aquatics Project consented to the use of their name in this dissertation.


Figure 4.1: A Makapo paddler in a one-person outrigger canoe (OC1) with version three of CoOP attached.

### 4.1.1 The CoOP System

In this section, I provide a brief description of CoOP to establish context for the results and discussion that follow. To work around the challenges faced by Makapo, I, along with my collaborators, designed CoOP, an assistive system through which the rudder (*i.e.,* steering) of an OC1 is remotely controlled by another person (*e.g.,* a coach), freeing blind and low vision paddlers from managing directional control and enabling increased focus on technique. CoOP sits on top of the rudder (see Figures 4.4 and 4.5) and manipulates it using special motors (*i.e., common servos* ), which are “plug and play" with a standard handheld RF transmitter. CoOP is attached to the OC1 using a GoPro Suction Cup Camera Mount (see Figure 4.6) (Inc.", 2019).

#### Design and evaluation of CoOP

Over a twenty-five month period from January 2018 to February 2020 at the Newport Aquatic Center, I engaged in a participatory design and deployment practice with Makapo. During 23

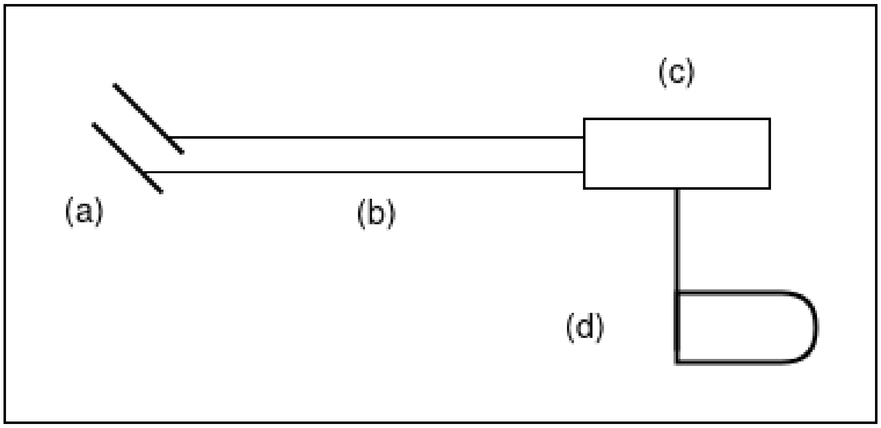



Figure 4.2: Diagram depicting how the rudder system works in an outrigger canoe. The pedals (a) are connected to the tiller cap (c) by a guide wire (b). The rudder (d) is attached to the tiller cap by a small shaft that runs bisects the stern of the canoe.

sessions the participatory design team iteratively developed, evaluated, and deployed CoOP, an assisted steering system for one-person outrigger canoes (OC1). Project development took place in 14 cooperative sessions from January to October 2018. Project deployment began in November 2019 and continued through February 2020 for a total of 9 sessions. Sessions were held at the Newport Aquatic Center with a core group (n=16 including myself and one additional researcher (see Table 4.1), for roughly two to four hours each.

In total, I conducted 76 hours of sessions over 24 days including semi-structured (n=7) and *ad hoc* (n=11) interviews, focus groups (n=2), and evaluations of the prototype (see Figure 4.3). Interviews and focus groups were audio recorded when possible, otherwise notes were taken. Deployment sessions were audio recorded, and periodically supplemented with video recording. The natural progression of the project ranged from planning to real world operation, which is categorized in detail below. Participation in sessions was opportunistic; sessions were scheduled around having enough participants from the core group who were available, and various others at the Newport Aquatic Center came in and out of the session areas as they desired. Because of the nature of people's availability, some members are more heavily represented in the results.

| *Identifier* | *Role* | *Vision* | *Organizational Affiliation* |
|---|---|---|---|
| ResearcherS1 | Researcher | sighted | UCI |
| ResearcherS2 | Researcher | sighted | UCI |
| DirectorLV1 | Director, Coach | low vision | Makapo |
| PaddlerLV1 | Paddler | low vision | Makapo |
| PaddlerLV2 | Paddler | low vision | Makapo |
| PaddlerB1 | Paddler | blind | Makapo |
| PaddlerB2 | Paddler | blind | Makapo |
| PaddlerB3 | Paddler | blind | Makapo |
| PaddlerS1 | Paddler | sighted | Makapo |
| PaddlerS2 | Paddler | sighted | Makapo |
| CoachS1 | Coach | sighted | Makapo |
| CoachS2 | Coach | sighted | Makapo |
| CoachS3 | Coach | sighted | Makapo |
| CoachS4 | Coach | sighted | Makapo |
| SpecialistS1 | O&M Specialist | sighted | Makapo |
| SupportS2 | Support Boat | sighted | Newport Aquatic Center |
| TechnicianS3 | Boat Repair Technician | sighted | Newport Aquatic Center |

Table 4.1: This table details members of the core design team for CoOP. They are organized by role within the design team, their visual acuity, and organizational affiliations. Identifiers are assigned using their role, visual status, and a numerical indicator.

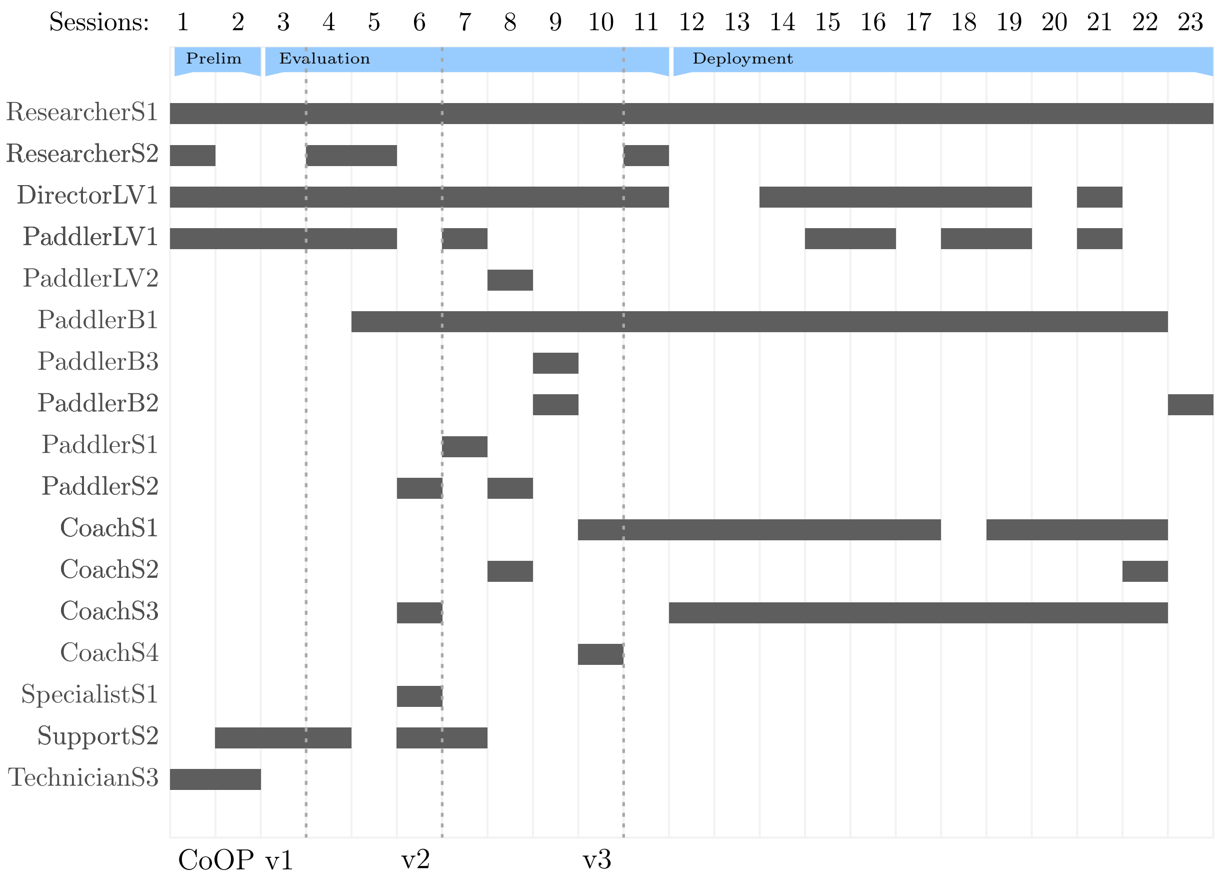



Figure 4.3: Timeline visualizing the involvement of the core team across all sessions. Grey horizontal bars represent the sessions attended by each member of the core team. The vertical dotted lines represent when new iterations of CoOP were implemented.

**Preliminary Sessions: 1,2** Preliminary sessions consisted of informal meetings with Makapo to gain a better understanding of the problem space. Because the research team had no experience with outrigger paddling, the first session was primarily spent discussing the details of the activity. Common goals, philosophies of enabling technology, and typical paddling skill sets were identified. The second session was spent exploring the type of canoe that would be utilized for the project. Through our exploration of the canoe, we identified guidelines and constraints of canoe modification, learned basic canoe operation, and agreed on a strategy for implementation. The research team spent the three month period leading up to the third session developing a low-fidelity prototype based on the data collected during the preliminary sessions (see Figure 4.4). The photographs and measurements of the canoe that were collected informed preliminary system design. Preliminary designs were first sketched on paper, followed by implementation using OpenScad[2] 3D modeling software, and finally physical creation (see Figure 4.5) using a combination of 3D printed and off-the-shelf parts.

**Evaluation Sessions: 3-14** The evaluation sessions commenced a four month period of iterative prototype development and testing. Sessions typically occurred every other week depending on participant availability and time constraints surrounding prototype development. Sessions were attended by stakeholders from UCI, Makapo, and the Newport Aquatic Center (see Table 4.1). Due to the public space in which prototype evaluation was conducted, the team also received solicited and unsolicited insights from nearby paddlers and Newport Aquatic Center employees. Evaluation sessions consisted of an iterative cycle of *in situ* tests to refine the design.

The physical dimensions of the outrigger canoe constrain any assistive technology designs.

[2]https://openscad.org

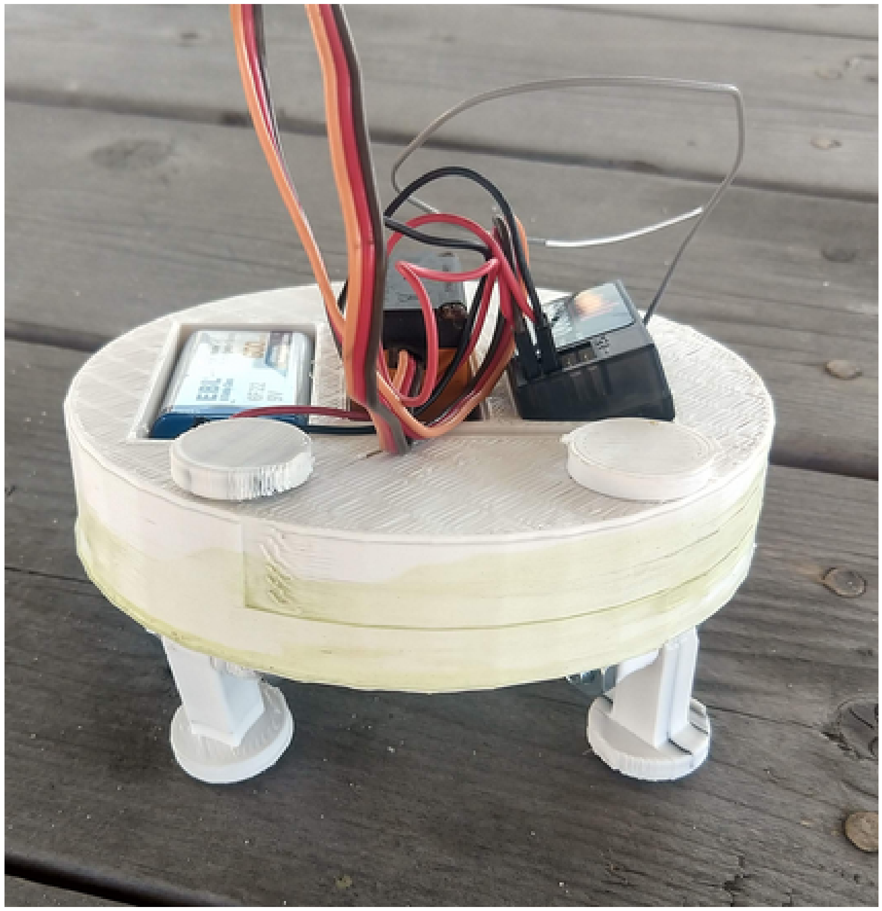


Figure 4.4: The first iteration of the remote control system that used adhesion as a mounting strategy.

At roughly ten feet in length, transferring the OC1 between the design lab and the Newport Aquatic Center was not practical. Additionally, Makapo was actively using the OC1 during the development phase, further complicating any plan to work off-site with the OC1. The OC1 that the team used for testing was stored on a rack located outside. At the beginning of each session the canoe would be removed from the rack and placed on stands in a public area where incoming and outgoing paddlers cared for their canoes.

The final three sessions centered on preparation for a yearly paddling race sponsored by the Newport Aquatic Center. Sessions were attended by a core group that included CoachS1, PaddlerB1, and at least one member of the research team. These sessions provided CoachS1 with the opportunity to learn how to use CoOP, give PaddlerB1 additional training, and

Figure 4.5: A close-up of version two of the CoOP system mounted to the rudder assembly and the transmitter used to control the rudder (right corner).

evaluate the robustness of CoOP.

**Deployment Sessions: 15-23** The deployment sessions focused on evaluating long term use of the CoOP system. Makapo coaches applied real world training scenarios to identify best practices for paddlers and support staff who used CoOP.

### 4.1.2 Data Collection

Data for this study comes from application of traditional ethnographic methods (*e.g.,* (DeWalt, DeWalt, & Wayland, 1998)) alongside the participatory design and evaluation of a shared assistive technology to support blind paddlers. I have engaged with the paddling community for over two years, throughout which the Newport Aquatic Centerand Makapo have welcomed my presence as a researcher. Staff for both organizations are familiar with my role, and freely engage me in dialogue related to blind paddling. Observations made while interacting with the paddling community are recorded through memoing (Emerson, Fretz, & Shaw, 2011) upon conclusion of an event. I avoid using visible recording devices such as a notepad, camera, or recorder, outside of suitable contexts (*e.g.,* others performing the same task) to communicate my desire for mutual collaboration. When formal capture methods are required, I request permission and explicitly state the purpose of data collection.

Observational data was collected from aquatic related activities at the Newport Aquatic Center, including: visually impaired athlete practices, sport sanctioned race events, evaluation of the CoOP system, and social engagements surrounding all paddling activities. Although my primary role was that of researcher, as I increasingly became part of the community, I adopted additional responsibilities in support of the activities I observed. These activities typically included canoe preparation, CoOP setup, and support boat operation.

Participatory design data from the development of the CoOP system was collected alongside observational data. Team members discussed device function during paddling sessions. Discussions were captured through notes and audio recordings. I collected videos and photographs of device use. At the end of each paddling session, team members reflected on their experiences and outlined changes for the next iteration.

### 4.1.3 Analysis

In the evaluation sessions, working alongside various members of the core team (see Figure 4.3), I performed repeated member checks and escalating *in situ* testing to continually improve the design of CoOP. I then analyzed interviews, observations, and field notes using inductive coding and memoing to identify needs and considerations for prototype iteration.

Additionally, I conducted an analysis on the design process, for which the themes that were generated were refined through discussions with collaborators.

Interview and observation data were analyzed using a deductive approach to identify patterns and themes. Deductive coding centered on categories of verbal and non-verbal engagement, enjoyment, cooperation, coordination, course corrections, reduced audio, and shared control design.

## 4.2 Results

In this section, I present the results of the field work and design, surrounding the iterative development of CoOP. I demonstrate how the *in situ* design process fostered investment and interest across the sighted and visually impaired paddling community. Results are presented according to three high-level themes that emerged from my data analysis: 1) public-facing co-design and development of assistive technology; 2) DIY within the context of community; and 3) the physicality of multi-modal design.

### 4.2.1 Public-Facing Co-design of Assistive Technology

Conducting the majority of design research in public at the Newport Aquatic Center served as a convenient location for the team to coordinate as well as allow researchers to conduct evaluations without the need to pull the test OC1 from active use. Although a lab-centered approach may have shortened the iterative design cycle, my thematic analysis revealed that many design decisions were a direct result of the visibility of the work to the general paddling community. In this section, I describe 1) how the public setting for the work enabled the collection of both solicited and unsolicited feedback, and 2) how multiple levels of expertise contributed to the design and development of CoOP.

### Solicited and Unsolicited Feedback

The team did not intentionally solicit feedback from the general paddling community, but conducting design and evaluation on the wash deck (*i.e.,* where equipment is washed down after use) at the Newport Aquatic Center exposed the process to a steady flow of paddling enthusiasts. In the second session, preliminary functional tests on the wash deck caught the attention of a few paddlers working nearby. Upon observing that the prototype was unable to turn the rudder, one of the paddlers commented:

> I used to be really into racing RC trucks, spent a lot of money in my youth, you need to get a high performance LiPo battery. They can deliver a lot more power than what you are using. You should also think about looking into a high-torque waterproof servo. – *Session 2, Unsolicited Community Member*

Although receiving unsolicited feedback from the public during an evaluation of a non-functional prototype was unexpected, the input I received prompted a deeper investigation into the capabilities of the radio-controlled watercraft ecosystem, such as through discussions with employees at a local hobby store. In the same session, a different community member overheard our discussion on harness mounting strategies and commented:

> You know what would work are those 3M stick pads, the ones that act like Velcro. I have used them to hold stuff to my canoe in the past. – *Session 2, Unsolicited Community Member*

As with the earlier community member, this input prompted a deeper conversation between team members, eventually leading to an entirely different mounting strategy. The team decided that straps wrapped around the hull of the OC1, an approach that I had originally rejected due to concerns over drag in the water, would be an acceptable alternative to adhesive mounts (see Figure 4.4). Recognizing the value of unsolicited feedback, the design team adopted an open design stance, inviting anyone who expressed interest in the work to express ideas. This shift in approach follows what Sanders and Stappers describe as a blurring of roles, where the researcher shifts from translator to facilitator (Sanders & Stappers, 2008). Embracing the broader community as experts of their own range of experiences, led to valuable opportunities for input and feedback on the current iteration of CoOP. From my field notes during the following session:

> Even with the straps, the fit was still a little bit too loose. One of the paddlers cleaning his canoe next to us suggested that we pick up some pipe insulation foam. He said, "It's cheap, easy to cut, and compresses enough to give you some flexibility [in the design]." – *Session 3, Field Notes*

Even if I could have eventually resolved the various design issues that arose during early sessions without these inputs, involving the broader paddling community in the process not only expanded how challenges were resolved but also increased community investment in the work. At the sixth session, PaddlerB1 and his father were attaching the harness to the OC1. After struggling to thread the straps into tension clips, an alternative mounting discussion took place:

> As we were cleaning the canoe after our test, PaddlerB1's father asked if I had thought about using a suction cup to hold the harness on to the canoe instead of straps. I expressed concern over durability, but he assured me that the brand of suction mount that he used would provide more than enough resistance to the forces we were experiencing. – *Session 6, Field Notes*

Here, a community member's father, reflecting on his own experiences with suction cup mounts outside of the paddling world, contributed a design insight that ultimately proved to be the right mounting solution for our final iteration of CoOP. His insight drew from both his individual experience (*i.e.,* his knowledge of suction cup mounts) and his collective experience with PaddlerB1(*i.e.,* facing a particular challenge)—a central tenet of the co-design process. Adopting community insight in this way can lead to positive feedback loops between the design team and community members that encourages deeper investment (Könings, Seidel, & van Merriënboer, 2014). After integrating the suction cup mount into the design, PaddlerB1's father exhibited increased engagement in the design process, which in turn encouraged the design team to solicit his input on new iterations.

Soliciting feedback from users or the community at large is a critical step in any user-centered design process (Abras, Maloney-Krichmar, Preece, et al., 2004), as it ensures that the product works appropriately for a target audience within a defined setting. However, it gains additional value within a public co-design process, allowing for the generation of edge-case scenarios and enabling more holistic design thinking. In preparation for final prototype evaluations, I was testing the strength of the new suction cup mount in the water with DirectorLV1 when PaddlerLV1 returned from a Makapo practice:

> During mount testing, DirectorLV1 asked if the harness would float. We detached the system and tossed it in the water. It floated. PaddlerLV1, standing nearby, commented, "If that were to fall off while I was in or near the canoe, there is no way I would be able to see it." DirectorLV1 agreed and suggested wrapping neon green flotation bands (typically used to keep personal items from sinking) around the top of the enclosure. He had one with him so we ran a quick test with CoOP. Both DirectorLV1 and PaddlerLV1 noted that it was much easier to see. – *Session 8, Field notes*

PaddlerLV1 and DirectorLV1 can and do paddle without sighted guidance; yet their reduced visual acuity makes it difficult to differentiate between lower contrast objects. At the time, the CoOP enclosure was being 3D printed with black ABS filament. I switched to neon green filament for the next iteration to increase the contrast between CoOP and the water (see Figure 4.6). While the use of fluorescent colors for marine equipment is not unusual, PaddlerLV1's unsolicited feedback *in situ* led me to think about how colors further shaped the design process for CoOP. Color might not be something a designer would think to consider for an assistive device for a blind user. Public and cooperative contexts, however, benefit from such consideration. Too frequently, assistive technologies that are well designed with the primary user in mind, like this one, do not consider other users who may not share the same disabilities as the primary user.

**Multiple Levels of Engagement**

Receiving feedback from the community proved to be a useful guide for adapting the design, as is the case in public engagement (Teal & French, 2016) and co-design (Sanders & Stappers, 2008) efforts. I was able to access this level of engagement through my repeated presence in the public community space. For example, in the following excerpt from field notes, I describe the degradation of the 3D printed tiller cap, the mount point that sits between CoOP and the canoe rudder, and the way in which we solved the issue using direct manufacturer expertise:

> Today DirectorLV1 introduced me to a design engineer from one of the canoe manufacturers who was visiting the Newport Aquatic Center. After a brief demo of the system, I expressed some concern over the durability of an ABS tiller cap. We brainstormed on design and material selection for CoOP's custom tiller cap. He suggested adding a particular type of washer for reinforcement. - *Session 5, Field Notes.*

The expertise accessed in this particular case was *ad hoc* and opportunistic. These are not the characteristics of a structured design process that is replicable, drawing attention to the challenges of the overarching DIY approach to some assistive technology. However, it also draws our attention back to the necessity, in this case, of accessing a dedicated community already engaged in the cooperative leisure activity at stake. In other cases, such dedicated expertise may be less essential to functional design requirements, yet foster idea generation in unexpected ways. For example, in one session, a Makapo employee was able to address an engineering challenge from her experience working with children with visual impairments:

> Today we asked the Makapo paddlers to attach the CoOP harness to the OC1. Noticing some of the challenges they were having with strap alignment and the tension clips, SpecialistS1, Makapo's orientation and mobility specialist, suggested using Velcro straps. She explained how the textural variations helped with fastening and alignment. – *Session 6, Field Notes*

Drawing on her background in orientation and mobility work, SpecialistS1 introduced a different perspective to the design process. Rather than focus on design or engineering requirements, SpecialistS1 considered the experience of manipulating objects without sight. As the lead for the kids program at Makapo, SpecialistS1 was not initially involved with the CoOP project. The project's public orientation, however, led to days where design sessions and kids paddling activities overlapped. SpecialistS1's insight introduced a critical design requirement to the process that the team incorporated into future iterations, reflecting the type of positive results championed by co-design practitioners (Sanders & Stappers, 2008). Textured thumb screws, a tactile power switch, and the suction cup mount described in the previous section, were all changes that reflected SpecialistS1's insight. Assistive technology researchers looking to engage in co-design may need to consider intentionally including this type of expertise on their design teams, and the development of supports for the larger non-research community to cooperatively create such solutions is an important area of future inquiry.

The benefits of disabled-abled co-design go well beyond the improvement the design team witnessed to the technology itself. Drawing from the expertise that surrounded the project—including both blind and sighted co-designers—established a community wide interest in the project. As the director of Makapo explained:

> People are excited. They keep hearing about what you are building and want to learn more. I think this has the side effect, you know, of drawing more people over when we are testing. – *DirectorLV1*

Notably, in the above quote, the "they" to whom the director is referring are not the people directly participating but a broad range of people who are involved with the Newport Aquatic Center, or who are local to the area but not involved, and so on. In this way, the co-design exercise organically grew wider than the direct participants, broadening awareness about the potential of the technology and the activities of blind paddlers. Novelty of the project likely led to a part of the curiosity and engagement; however, the month's following the development of CoOP indicate that awareness extends beyond initial curiosity, allowing further engagement.

Every Fall, the Newport Aquatic Center sponsors an open water race for human-powered vehicles that is designed to promote water sports and bring the various paddling communities together. Makapo typically participates in the OC6 category with a sighted steerperson. For this particular race, Makapo received permission to allow PaddlerB1 and CoachS1 to enter the race using CoOP. When the race was set to start, the race director focused everyone's attention on PaddlerB1:

> Waiting for the official race kick-off, I am on the support boat alongside DirectorLV1, PaddlerB1's dad, and the film crew. SupportS2 is driving the boat and

> CoachS1 is steering CoOP. We are surrounded by approximately sixty race participants as the race director prepares to start the race. Using a megaphone, he informs everyone that PaddlerB1 will be the first blind paddler to compete in an OC1. PaddlerB1 receives a round of cheers from the other paddlers. – *Race Day, Field Notes*

The director's announcement served to make sure the race participants understood why PaddlerB1 would be followed by a support boat, but it also brought attention to the significance of PaddlerB1's participation. After the race, while cleaning the OC1 on the wash deck, the team was greeted by numerous race participants interested in learning about the CoOP system. These events directly reflect the public co-design nature of the project. Our visible presence at the Newport Aquatic Center, and involvement with the community, brought merit and meaning to the project that would likely be absent if the team simply showed up with CoOP on race day.

In the months that followed the race, I observed two changes that reflected an increasing awareness and engagement of this community through better accommodations for Makapo's blind paddlers: the relocation of canoe storage to a more accessible location, and the introduction of a portable walkway. When I returned to the Newport Aquatic Center for the first time after the race, I noticed that Makapo's OC6 canoes were no longer stacked in a hard to reach area of the beach. From my field notes:

> At the Newport Aquatic Center today, noticed that the Makapo OC6's are no longer crammed between the other paddling crew canoes on the beach. They are now located on the outer edge of the area with a lot more spacing around them. – *Field Notes*

The change was significant. The coordinated actions of launching and stowing the OC6 that

I described in the previous section were a constant challenge for Makapo paddlers when the canoes were stored in the original location. The proximity to other boats, support stands, ropes, and debris, added unnecessary obstacles for the sighted team members to attend to when maneuvering the canoe. The additional space provided by the new location removed the need to navigate these obstacles. Similarly, in a meeting a few weeks later I learned that a staff member at the Newport Aquatic Center expressed interest in continuing to make the wash deck and beach area more accessible, from my field notes:

> Met with Makapo today to discuss the installation of portable walkways on the beach designed to support wheelchair access. DirectorLV1 mentioned that the Newport Aquatic Center has expressed interest in supporting some type of installation to make traversal easier for blind paddlers. – *Field Notes*

While I am unable to broadly claim that the public co-design orientation of this work influenced staff members at the Newport Aquatic Center, our observations align with positive changes in attitude associated with mixed-ability exposure (Barr & Bracchitta, 2015). In an interview conducted with DirectorLV1, I explored his thoughts about the changes:

> I think it [public co-design] has had an impact, more so with the staff at Newport Aquatic Center than with the general public—they need to see more, but for the staff, they are more aware of what we are doing here and we are starting to see that effect more and more. – *DirectorLV1*

The public co-design activity introduced much more than just the feedback into the design process—it brought attention to and made more people outside of the direct community consider some of the issues that individuals with disabilities struggle with on a day-to-day basis.

Figure 4.6: The design evolution of the CoOP system in order of iteration from left to right. Early iterations of CoOP version 1 are depicted in the first three pictures from the left, followed by versions 2 and 3.

Although physical constraints of canoes had forced the project to be conducted in a public space, it appeared to have the unintended and positive effects of promoting engagement, shaping the project in ways likely different from how the original partnership between the research team and the community organization would have, and spreading information about the project to a disparate audience. Researchers doing community engaged work should leverage this by thoughtfully incorporating public space projects as an intentional part of their process, even without the forced constraint.

### 4.2.2 DIY within the context of community

A common challenge when designing for people with disabilities is gaining access to the community for conducting user studies (Kane et al., 2014). My partners were cautious about allowing their community to become subjects of an experiment that may give undue hope or stress without resulting in something that they could use in the foreseeable future. By emphasizing a do-it-yourself (DIY) approach, I was able to clearly demonstrate the steps for reproduction and giving my partners a sense of reciprocity. As noted in my field notes:

> Today we conducted the first water test. I joined DirectorLV1 and SupportS2 on the support boat while PaddlerLV1 paddled in the OC1. Afterwards, as we inspected the system on the wash deck, DirectorLV1 was pulling in anyone who

> walked by to share the news, emphasizing how inexpensive it was to build. – *Session 4, Field Notes*

The initial test process showed my partners that the relationship would not be one-sided; They would not be limited to answering my research questions. Parting ways is an inevitable outcome for nearly all community research projects (Hayes, 2014), so communicating my intentions early was an important step in the co-design process. Transparency about working progress—articulating precisely how I produced the prototype, including cost, tools, and part suppliers—gave Makapo confidence that they would be able to use and continue the project independent of my involvement. In the months following the conclusion of our deployment sessions, Makapo has continued to use the CoOP system without my involvement and acquired a 3D printer to produce replacement parts and additional devices.

The paddling community—as far as it is represented at the Newport Aquatic Center—exhibits attributes reflective of DIY practices, which have been shown to support assistive technology work (Hurst & Tobias, 2011; Hamidi, Baljko, Kunic, & Feraday, 2014). Canoes are personalized, modified, customized, and reappropriated to satisfy paddler interests. I frequently observed DIY behavior throughout my time with Makapo. When asked about canoe personalization, DirectorLV1 explained:

> Our canoes have a personality, they are considered part of the family. The modifications people make, the stickers, custom colors...it comes from a long history of use. The outrigger canoe has deep history with the pacific island culture, and I think those historical practices just emerge when people get into the sport. – *DirectorLV1*

Bringing a DIY approach to assistive technology design into a community that engages in similar practices influenced how CoOP was perceived. Attaching an obviously handmade device onto a canoe was not seen as out of the ordinary with the context of established community practice. On the wash deck, every outrigger canoe looks different, even when they are the exact same model from the same manufacturer. Although the most common forms of personalization can be observed through color and semi-permanent attachments, the canoes also undergo permanent physical modification, as with the following example from my field notes:

> The canoe we will be evaluating CoOP on looks pretty beat up. It's not painted, has areas where the buoyancy material is exposed, and has random beer can stickers on its hull. – *Session 1, Field Notes*

Although the canoe was fully functional, it was far from "normal" in its visual appearance compared to the backdrop of outrigger canoes, rowing sculls, and kayaks that surrounded it. Despite its looks, the canoe is an employee favorite. The hull had been cut apart and reshaped a number of times to fulfill performance curiosities of paddlers, eliciting a desire for creative expression and identity that echo sentiments common to crafting communities (Bardzell, Rosner, & Bardzell, 2012). The primary canoe the team eventually used for testing went through a similar journey:

> Makapo acquired a new canoe to use for CoOP evaluation. The canoe had fallen off of a car on the highway and received so much damage that the owner decided to replace it rather than repair it. The canoe repair team at Newport Aquatic Center rebuilt the fiberglass hull and repainted it. – *Session 2, Field Notes*

In both cases, the repairability of the particular style of canoes made them conducive to modification. The modification reflected back on the behavior and attitudes of the broader community. For paddlers, DIY is an expected, normal part of canoe ownership. Strapping plastic parts to the back of the canoe, as the design team regularly did while testing CoOP, was not out of ordinary on the wash deck. Throughout the prototype design process the only critique I received came from a non-paddler:

> PaddlerB1 brought a large group of family members out for our evaluation today. One relative provided some critique on my bolt size choices for the mounting legs, noting that they were probably overkill for the application. I later learned that he was a retired aviation engineer. – *Session 5, Field Notes*

These kinds of comments hint at the contrast between a DIY stance and one more aligned with professional production. For the DIY community, successful operation of the artifact takes precedence over decisions like bolt size. Rather than constant scrutiny over design decisions, those around the design team during prototype development accepted the status of the system at each new session and only tried to assist when we were met with a challenge.

Making structural changes to canoes is one of numerous ways that the paddling community engages in DIY behavior, and one that from my observation appears to be reserved for a subset of paddlers with a particular skill set. For others, DIY manifests itself in non-permanent modifications and customization. At the end of Session 7, DirectorLV1 and PaddlerS1 were discussing their plans to have custom decals created for their canoes. They talked about options for ordering, colors, designs, and meaning behind their choices. Other paddlers maintained a variety of attachments used for carrying artifacts such as phones, fitness trackers, water bottles, and cameras. Although many of these attachments were commercial products, they imbued each canoe with a distinct appearance.

While the personalizations I observed are certainly not unique to the paddling community, they demonstrate the heterogeneity of paddling culture. The perception of an assistive technology attached to a canoe is that of typical paddler behavior rather than a beacon for disability. This allows Makapo paddlers to maintain control over the visibility of their disabilities. In one example from early in my field work, I observed how PaddlerLV1 wore a bright orange shirt with the Makapo logo on the front and the words “Blind Paddler" on the back in large black block letters. After I had spent more time working with Makapo I noticed that the PaddlerLV1 had stopped wearing the shirt. When asked about the purpose of the shirt, the DirectorLV1 noted:

> We had those shirts made up for a race a few years ago. There are no requirements that it be worn, we just thought it would be helpful for other race participants. It’s also a good way to let them know what we can do! – *DirectorLV1*

Rather than a requirement, the shirt is representative of how Makapo perceives itself within the broader paddling community. They did not need to inform other paddlers of their disability, but did so when they were proud to declare their participation in an outrigger canoeing event. As others have demonstrated, control over disclosure is a desirable feature for assistive technology (Shinohara & Wobbrock, 2011; H. P. Profita et al., 2018).

### 4.2.3 Physicality and Interdependence as Design Requirements

Physicality impacted the paddling experience with and without assistive technology. My design goals were frequently modified for CoOP as I observed the existing practices of Makapo’s paddlers. Makapo’s blind and low vision paddlers appropriated naturally occurring phenomena from the environment to support the loss of auditory cues in the six person canoe. Auditory cues are highly valuable to the Makapo weekly children’s program:

> Today I joined the Makapo team and the [kids paddling program] in the triple hull canoe. I sat in the front behind PaddlerB3 and PaddlerB2. PaddlerS1 was the steerperson and called out [paddle commands]. She teased PaddlerB3 and PaddlerB2 whenever they started to paddle out of sync. So PaddlerB3 started

> audibly calling his strokes. PaddlerB2, listening to the calls, could then sync his water entry point with PaddlerB3. – *Session 5, Field Notes*

The steerperson is the sixth person in an OC6 who is responsible for steering and calling out pace. For Makapo paddlers who are visually impaired, the sounds of paddles entering and leaving the water combined with calls from the steerperson provide feedback on their individual pace. The absence of these auditory cues makes gauging pace more challenging. When asked if the audible calls to synchronize paddles was typical behavior, PaddlerB3 and PaddlerB2 explained that they usually do not have to, but were having a hard time differentiating paddle sounds in the triple hull with so many people (twenty kids and adults). While synchronization is not necessary in an OC1, blind and low vision paddlers appropriate audible sounds in other ways. According to CoachS1, sighted paddlers will use the wake created by the canoe, movement of objects in a stationary position (*e.g.,* fishing boats, buoys, and landmarks), and GPS based fitness devices to measure pace. For Makapo's visually impaired paddlers, these visual indicators are either entirely absent or too difficult to see to be worth using. At Session 7, the design team was evaluating CoOP with PaddlerB1 for the day. PaddlerLV1, who was also at Newport Aquatic Center that day, decided to join using PaddlerS1's OC1. When we returned, he described how difficult it was to manage pace without the "sucking sounds." Curious about what he meant, I probed further:

> When I'm paddling hard in [their usual OC1 model], I can hear suction sounds from the drainage holes in the canoe. So I use that sound to know how fast I am going. But in [a different OC1 model], there are no sounds. – *PaddlerLV1*

In the absence of the visual landmarks available to sighted paddlers and audible feedback from teammates in an OC6, PaddlerLV1 learned to rely on the unique sounds generated by a component of the canoe designed to drain water from the foot area. Such a phenomenon is not uncommon for people with visual impairments who frequently rely on background noise for contextual cues about their environment (Koutsoklenis & Papadopoulos, 2011). However, acknowledging the value of ambient sounds, taking steps to identify them, and integrating them into the design process for assistive technology presents an opportunity to enrich the usefulness of the devices we create. Learning about the reappropriation of the use of drainage holes inspired me to reflect on design decisions for CoOP differently: one constant side effect of using a digital servo motor to turn the canoe rudder is the constant high pitched hum the motor generates when it is engaged. As testing progressed, the hum was frequent enough that I grew concerned over potential damage to the motor. During post-session discussions about the issue, PaddlerLV2 and PaddlerB2 explained their interpretation of the hum:

> It's kind of funny, cause when you turn you want to make sure you are paddling on the correct side of the canoe, so when you turn on the motor or whatever, I can hear it squeal, and prepare for a turn. – *PaddlerLV2*

> Yeah, I know when I hear the motor noise start that the boat is going to turn. – *PaddlerB2*

Recognizing the value of the audible noise CoOP emitted, I embraced the servo noise as a feature rather than an undesirable artifact. In the final version of CoOP, I integrated a relay between the battery and the servo motor, allowing the guide to turn the motor on and off as needed, while isolating the servo noise to moments of steering activity. In subsequent tests, participants noted that a direction change was easier to anticipate when the hum only occurred as the guide was preparing to steer. These observations of the physical experiences of paddling were not limited to auditory cues. The introduction of CoOP to the canoe transferred steering responsibility from the paddler to the guide, leaving the pedals unnecessary except during a CoOP failure. Yet as paddlers spent time in the canoe, they discovered that they could use the movement of the pedals as feedback for directional changes.

Without verbal communication, paddlers learned how to interpret secondary effects of CoOP to anticipate course changes and direction.

The physicality of CoOP extends to more than just allowing users to understand and anticipate its state. As all paddlers have Newport Aquatic Center membership in common, the typical social interactions one might expect from similar organizations occur throughout the day as paddlers come and go. The Newport Aquatic Center operates with a consistent ebb and flow: paddlers arrive, pull their canoes from storage, prepare them for use, carry them to the waterfront, paddle, and reverse the process before departing. The prep and stow routines practiced by Newport Aquatic Center members are an essential part of the paddling experience, serving as a rally point for social engagement. My observations of established customs and practices at the Newport Aquatic Center demonstrated how important it was for CoOP to fit into this paddling culture. As I observed in my field notes:

> Today PaddlerLV1 helped PaddlerB1 take down the OC1 and clean it. PaddlerLV1 took the lead carrying the bow up to the rinse area while PaddlerB1 followed with the stern. I have seen PaddlerLV1 carry the canoe by himself before, so the split duty was not about weight, but rather including PaddlerB1 in the process. – *Session 6, Field notes*

Blind and low vision paddler participation in the same community practices encouraged me to consider CoOP not just as as an artifact for the Makapo organization, but as a tool to extend and enrich the social engagement of Makapo paddlers within the larger paddling community. These kinds of shared activities increase awareness and inclusion for people with disabilities in society more generally, and can improve social and emotional intelligence for children without noted disabilities as much as for those with (Ochs et al., 2001). Sharing canoe carrying was just one of several procedural activities during which I observed Makapo paddler participation. PaddlerLV1 and PaddlerB1 frequently teamed up during sessions as the primary paddlers helping the design team evaluate functionality of CoOP. During other sessions PaddlerLV1, CoachS1, and SpecialistS1 guided PaddlerB1 through the process of attaching the outrigger to the canoe.

The role that canoe maintenance played for paddlers shaped how I thought about design for CoOP. Attaching, cleaning, and removing CoOP from the canoe needed to parallel other aspects of the canoe preparation cycle such that the Makapo paddlers could take ownership of the process. Having established that a portion of CoOP would be shared with sighted paddlers during preliminary sessions (see Section 4.1.1), I wanted to ensure that all other aspects of the system were accessible to Makapo paddlers. Working through these issues *in situ* served as a critical step towards identifying design decisions to consider both independence and interdependence as design goals for CoOP.

Assistive technology enables people with disabilities to perform tasks that would otherwise be difficult or impossible to complete. Ingrained in this definition is the idea that assistive technologies should allow the people who use them to live with greater independence. In preliminary sessions with Makapo, I probed the extents to which the system needed to support independent navigation. Drawing upon knowledge of recent research in blind navigation, alongside my collaborators, I discussed the use of sensors for obstacle avoidance, haptic feedback mechanisms, and computer vision. Although Makapo expressed interest in exploring autonomous technologies, they emphasized the importance of simply getting their paddlers on the water in OC1's, as DirectorLV1 described:

> We reached out to a local [robotics club] for help a few years ago that said it would cost us a few thousand dollars. The thing is, for us, we really just need a way to get our paddlers in a canoe by themselves. Training in an OC1 is the best way to improve your performance in an OC6. Since it's going to be a training situation, there will be a sighted coach with them anyway, so we really don't

> need anything fancy. – *DirectorLV1, Session 1, Field Notes*

For Makapo, the desire for an entirely independent experience was secondary to the need to build a competitive outrigger paddling team. Team driven cooperation is an intrinsic quality of outrigger paddling, regardless of the individual abilities of team members. That team driven culture influenced Makapo's perception of assistive technology. As much as they might benefit from a fully autonomous solution, they realized a practical approach was the best way to accomplish their goal. Knowing that a sighted coach would be present for OC1 practices meant that some responsibility could be transferred from the paddler to the coach, thereby limiting the amount of technology required to support the paddler, and effectively leading the team towards building a shared assistive technology.

As part of the team-based culture, the practice of sharing manifested itself in numerous ways. As I described in Section 4.1, the Newport Aquatic Center is home to a wide variety of aquatic activities. Over the course of the eight months I spent at Newport Aquatic Center, I observed row and outrigger paddling crews of various sizes working alongside each other to setup, carry, clean, and break down their canoes. A practice reflected with Makapo as well, from my field notes:

> CoachS2 was the first out of the canoe, followed by PaddlerS2. PaddlerS2 held the canoe while CoachS2 retrieved a wheeled carriage used to roll the canoe up the beach. CoachS2 aligned the carriage and instructed the remaining four paddlers to step out of the canoe. CoachS2 and PaddlerS2 verbally guided the blind paddlers into positions around the canoe. The blind paddlers used differing tactile cues of the canoe to navigate towards the requested positions. In concert, they lifted the canoe onto the carriage and pushed up the beach to its storage location. – *Session 11, Field Notes*

When Makapo's OC6 group practices, they verbally coordinate positions around the hull of the canoe, push the canoe down to the water at the start of practice, and bring up to the beach after practice. In both situations, blind and sighted paddlers work together, sharing responsibility for the care and operation of the equipment. From this perspective, designing an assistive technology that shares control between sighted and visually impaired paddlers fits naturally within the context of the paddling experience at Makapo. As PaddlerB3, explained:

> It's not much different than what we do in the OC6, it's somebody else's job to steer, I just focus on my stroke. – *PaddlerB3*

Here, PaddlerB3 is describing an experience similar to all paddlers, regardless of ability, participating in a coordinated team-based activity. The role of steering is singular, assigned to a team member with a particular skill set. For PaddlerB3, that an assistive technology fills that role in an OC1 is irrelevant, the experience remains the same. His primary goal is to perform his assigned role as effectively as possible for himself and for his team. DirectorLV1 expressed a similar sentiment, stating that opportunities for contribution are rare:

> When I paddle in a six-person outrigger canoe, it's one of the few times as a blind person where I know my sighted teammates are relying on me, and that doesn't happen very often, so it's really empowering. – *DirectorLV1*

For DirectorLV1, the contribution that he makes to the success of his team is an empowering activity. He has found empowerment through contribution, rather than by executing his own independence. I am not asserting that independence is not a worthy goal for assistive technology, but that assistive technology can help achieve empowerment in unexpected ways.

### 4.2.4 Impact of Natural and Artificial Auditory Information on Blind and Low Vision Activity

Visually impaired paddlers depend on verbal and non-verbal auditory information to complete canoeing activities on the water. Analysis of my cooperative fieldwork indicates how auditory information shapes paddler actions in pursuit of their goals. While in most cases paddlers routinely complete their goals, I identified a tension between paddler enjoyment, their dependency on auditory information, and the issues that arise when auditory communication breaks down. In this section, I describe how these tensions arise over the course of a training session and the steps that both paddler and coach take to return to a functional state. I begin with an exploration of the underlying motivations for participation in outrigger paddling. Understanding the connections that develop between paddler, coach, and activity highlight the importance of auditory information on the water. I then describe how paddlers and coaches make use of the auditory channel through verbal and non-verbal communication. Finally, I explore the challenges that arise on the water when auditory communication breaks down.

#### Immersion in an Unconventional Activity

Whether alone or with others, outrigger canoeing is inherently a social activity through which paddlers bond over the quirks and customs of the sport. A common phrase articulated by DirectorLV1 when discussing Makapo's mission, *"You have to be a little bit nuts to do what we do."*, simultaneously referring to the extremes paddlers are willing to undergo and Makapo's mission of making the sport accessible to the visually impaired community. For Makapo members, outrigger paddling represents a source of personal growth, learning, and agency that they struggle to find anywhere else. During an interview with Makapo staff, one coach described their approach with blind members of the racing team:

> We give emotional support, but I am the kind of person that pats you on the back and kicks you in the ass. We get people who aren't athletes and we are building them to be athletes. You are pushing yourself to get to 5 miles as fast as you can. We want to ingrain a competitive aspect into them. In one instance, while we are in the middle of practice, as they reach the 2 mile mark they are just wanting to get out of the boat. We are training for [race in Hawaii], I canâĂŹt leave you on some beach if you get tired. PaddlerB2 finally hit a wall. He went two miles, emotionally broke down, but after that he was 10 times better than he ever was. — *CoachS1, Interview*

I frequently observed the behavior that CoachS1 describes here from multiple coaches. Paddlers are pushed, because that is what the sport requires, rather than coddled because of their disability. Perceived overprotection (*e.g.,* (Cimarolli, 2002; Cimarolli, Reinhardt, & Horowitz, 2006)) and lack of social support (*e.g.,* (Papadopoulos, Papakonstantinou, Koutsoklenis, Koustriava, & Kouderi, 2015; Alma, Van der Mei, Groothoff, & Suurmeijer, 2012)), which are often reported by the blind community as a source of distress, suggest that the equity conveyed by Makapo coaches is viewed positively by blind and low vision members of the racing team. When we started integrating CoOP into regular practices, PaddlerB1 embraced the experience. Towards the end of one training session, CoachS2 noticed our group while paddling on her own and joined PaddlerB1 for the final mile. During post-session discussion, PaddlerB1 commented:

> I really liked when CoachS2 gives me tips, like when she said 'keep your back straight, you are arching your back too far.' – *PaddlerB1, Interview*

For PaddlerB1, his enjoyment of the activity translated to the agency he acquired from solo paddling. He often stated how he liked the *"freedom"* and *"independence"* of solo paddling compared to his experiences in the six person canoe. The one-to-one mapping between effort (*e.g.*,paddle stroke) and output (*e.g.*, canoe speed) resonated with PaddlerB1, as he described during the same discussion:

> Yeah it was pretty amazing, because when I do my paddling, it just feels, you know I can get myself around pretty quick, pretty fast, doing you know getting more power, putting more pressure on the paddle to get more power, that's what causes [the canoe] to go much faster... ...I felt really good about being able to switch from left side to right side whenever I like to. – *PaddlerB1, Interview*

In a six-person canoe, paddlers are required to switch paddle sides (effectively giving different muscle groups an opportunity to recover) after a predetermined stroke count. Here PaddlerB1, who at the time of our discussion was still new to solo paddling with CoOP, acknowledged the satisfaction he received when presented with the ability to make that determination for himself. The control that solo paddling offers was one of the primary motivators for the development of the CoOP system, and represents a unifying theme carried throughout its development. As DirectorLV1 described during one discussion over ownership and deployment of the CoOP system:

> Independence is very important. The blind paddlers could give [CoOP] to a sighted person, but that's not what our coaching style is about. We want to provide as much independence as possible. – *DirectorLV1, Interview*

Although the experience of paddling in the one person canoe was preferred by PaddlerB1, finding enjoyment on the water is not limited to solo paddling. For PaddlerLV1, spending time in the one person canoe was primarily in service of self-improvement, motivated by a desire to strengthen his abilities for the six person canoe team. A goal that his coaches observed during one of our post training session discussions:

> It is common if you are in a six [person canoe] at some point, to make the jump [in performance], you have to get on a one man to figure it out. I think for our paddlers it's hard to explain the concepts of what they need to feel. So getting on the one [person], I have seen PaddlerLV1 getting better. – *DirectorLV1, Interview*

In this quote, DirectorLV1 reflects on his own experiences as a low vision paddler and coach, recognizing that most of Makapo's visually impaired paddlers who find solace in the six person canoe, need time to internalize how both solo and team activities function in support of each other. PaddlerLV1 observed the benefits of solo paddling through his own self improvement and found enjoyment in the balance of both styles. PaddlerLV1, who is also low vision, regularly paddles in the one person canoe without the use of the CoOP system. Though he is typically with a sighted partner, also in a one person canoe, PaddlerLV1 makes use of known routes where high contrast landmarks serve as navigational aids. Additionally, PaddlerLV1 appropriates auditory indicators to augment his reduced visual acuity. Conversely, PaddlerB2, who through our time together had limited opportunities to experience solo paddling, explained his perspective:

> I like the freedom of the one [person canoe], but maybe it's security, I like the security [of the six person canoe]. And like I said, the teamwork and people working together. I wasn't in any sports before, ever, and so I don't have anything, whether it be a single person sport or a team sport, to compare it to, but I really like the idea of a team sport and you don't get that with the one, I like the freedom of the one, but if someone was to present me with the six or the one I would take the six. – *PaddlerB2, Interview*

The feeling of security that PaddlerB2 describes here refers to increased stability of the six person canoe, which is much less likely to roll over than the one person canoe. Earlier in our discussion, PaddlerB2 described how we was overly focused on preventing the one person canoe from flipping. He found the effort required to maintain his center of balance interfered with his enjoyment of the experience. While he acknowledges the value of the one person for training purposes, his enjoyment is primarily drawn from the team setting, an experience notably absent throughout his life as a blind adult. Despite Makapo's desire to broaden their paddlers experiences beyond the six person canoe, they recognize their members attraction to the team experience:

> Well there is also the social aspect [to paddling], there is a desire to get them down here and get on the water. Be on the water. – *DirectorLV1, Interview*

At the core of the teamwork that PaddlerB2 described earlier is a sociality that emerges from spending hours in a small canoe together, working towards a common goal. Paddlers receive instructional feedback from coaches, banter over performance, and interpret and respond to natural and artificial sounds in their environment, all of which contribute towards shaping an enjoyable experience. The perspectives highlighted so far summarize the myriad ways through which Makapo paddlers identify with the sport of outrigger canoeing. Enjoyment is reflected through the multiple facets of activity the sport provides, all of which depend on the acquisition and interpretation of auditory information. In the next section, I describe the ways in which paddlers are dependent on verbal and non-verbal auditory information.

### Reliance on Verbal and Non-Verbal Communication

The different types of auditory information that make canoeing enjoyable for Makapo paddlers also make the activity possible. The interpretation of auditory communication from verbal and non-verbal information is critical for paddler safety, learning, and situational context. As the dominant source of sensory input and output, the auditory channel bears the majority of responsibility for information processing in blind and many low-vision individu-

Figure 4.7: A typical training session configuration. The photograph is taken from the support boat using the CoOP system. In the distance, a coach drives a second support boat used to observe paddlers at different locations along the training route.

als. Although the absence of visual information has been linked to increased perception and information processing through the auditory channel (Röder et al., 1999; Röder et al., 2001; Collignon & De Volder, 2009), auditory information must still be audible, comprehensible, and contextually relevant. In this section, I describe how real world conditions can challenge the assumptions of assistive technologies that substitute sensory input with auditory information.

Paddling is generally a safe activity, regardless of individual ability. The greatest risk to paddler well-being arises in the water. In calm water conditions, solo paddlers can recover from an overturn by righting the canoe while treading water and lifting themselves back into the seat by using the outrigger for stability. When water conditions are rough, paddlers wear inflatable life vests and an ankle leash to keep their bodies connected to the canoe. Recovering from an overturned canoe is largely an inconvenience that disrupts the activity, or as PaddlerB2 exclaimed during a training session in February, *"I do not want to huli [overturn canoe] in this cold water and then be all wet."* So when discussing safety in the context of outrigger paddling, the primary objective of the Makapo coaches is to keep paddlers upright by focusing on navigation and balance.

Paddlers rely on frequent auditory communication for safety. Risks can arise unexpectedly, requiring coaches to visually assess evolving situations on the water and communicate those situations clearly to the paddler. Some risks to safety can be assessed by paddlers, for example, at the end of one particularly windy practice session, CoachS1 explained how wind can affect the canoe:

> [The wind matters] a lot, if you are going against the current or in heavy waves, there are times it can pop your ama (canoe outrigger) up, and send you into a huli. *−CoachS1, Interview*

As paddlers gain experience in the one person canoe, they are able to assess wind conditions and prepare themselves for unstable conditions. However, as CoachS1 points out, waves also factor into the risk of overturning. Coaches use verbal communication with paddlers to prepare them for inbound waves. When a wave or set of waves are large enough to raise the canoe outrigger, coaches will instruct the paddler to *"lean left"* or *"paddle on the left"* to ensure that their body position puts as much weight as possible on the outrigger, offsetting the risk of an overturn. In heavy wave conditions, the successful communication of this information is critical to the safety of the paddler. The natural and artificial sounds produced on the water can make it difficult to exchange information through verbal communication. The repetition of information occurred regularly throughout the sessions that I observed.

The following conversation from my transcripts is representative of the type of repetition that occurs on the water:

> CoachS3- “Hey PaddlerB1, hold water, we are going to wait here for a moment until the ferry docks."
>
> PaddlerB1- “What’s that?"
>
> CoachS3- “Hold water"
>
> PaddlerB1- “Okay"
>
> CoachS3- “The ferry is getting close, so we’ll wait until it passes. Take a minute to drink some water."
>
> PaddlerB1- “What did you say?"
>
> CoachS3- “Drink some water." – *PaddlerB1 and CoachS3, Transcripts*

The ferry crossing, a stretch of water traversed during training sessions, contains three car ferries transporting people between a small island and the mainland. Each session, the team attempted to time the passage of the area to avoid having to stop. A nearby gas station and boat rental docks often created enough water turbulence and traffic along with the ferries, that stopping introduced undesired complexity for the team. In the exchange quoted here, the team was unable to get the timing right and forced into a quick stop. The artificial sounds produced by boat traffic made it difficult for PaddlerB1 to hear the command to hold water. In this particular instance, it was critical that the command was followed to prevent additional complications (*e.g.* reversing direction, navigating between other boats, or bumping into docks). Recognizing the need for immediate action, CoachS3 quickly adjusted his informational structure from conversational to concise to reduce the chance of the command being missed a second time. Although productive in situations like the one described here, separating the explanation from the command is counterintuitive for coaches who are accustomed to descriptive dialog with visually impaired paddlers. Yet in moments where precise command interpretation and response is required, the sighted members of the team often worked against their intuition. Gradually, a pattern emerged where, once everyone felt the situation was under control, the team would describe the conditions that lead to the change in goal, as captured in my field notes:

> Typically we only have one ferry to contend with, but today all three were actively moving across our lane at the same time. To make things more complicated, a rental boat was leaving the dock near us and some jet ski's were coming in for fuel. We decided to just take a break early before the final stretch back. CoachS1 told everyone to hold water. As CoachS3 pulled alongside PaddlerB1's canoe, I rested my feet on the outrigger for stability and to keep the canoe from drifting. – *Field notes*

As I kept PaddlerB1's canoe alongside the support boat, CoachS3 provided a detailed description of what was happening for everyone. By limiting verbal instruction to operational commands until those operations have been completed, situational complexity is reduced, creating a safer environment for the entire team. Although complex areas on the water like the ferry crossing generated an urgency for precise verbal communication, I also observed how complex verbal communication disrupted the team in calm, open areas on the water. Wind, waves, canoe sounds, and positioning frequently made it difficult for paddler and coach to hear each other. In my field notes I describe once instance where the inability to hear affected the team:

> The wind today made it difficult to hear each other unless the support boat was alongside the canoe. Everyone had to repeat themselves multiple times. The repetition clearly became a source of frustration, as most communications just ended up with someone saying "never mind". The typical coaching guidance that CoachS3 and CoachS1 offer to the paddlers fell off while we were out past the break. – *Field notes*

Even though the paddling conditions were safe, and command interpretation was not critical, when the ability to comprehend verbal instruction of any type breaks down, both coaches and paddler are negatively affected. For coaches, when accurate verbal communication requires repetition, an increase in attention is required to anticipate changing conditions to offset the risk that a paddler might not hear the first call of a command. When combined with frequent command repetition, coaches experience additional cognitive burden. After one particularly challenging session CoachS3 described the experience as exhausting, as captured in my field notes:

> As we were preparing the launch boat this morning, CoachS3, reflecting on last weeks session commented on how tired he was by the time he got home. "I know PaddlerB1 is doing all the work, but man, when we got home, I crashed. Took a long nap! It's surprisingly stressful out there having to keep track of [the paddlers] steer PaddlerB1 and the launch." Later in the session, I asked CoachS1 and CoachS3 if they get stressed or anxious when the water is busy:
>
> CoachS3 – "Yeah"
>
> CoachS1 – "No. [laughter]"
>
> CoachS3 – "I still get nervous"
>
> CoachS1 – "I mean, if we were in the main channel and boats were coming in everywhere, yeah maybe, but here, no it was good."
>
> – *Field notes*

Although CoachS1's experience as a coach and outrigger paddler outweighs CoachS3's by many years, the responses of both coaches indicate that anxiety accompanies increased situational complexity. For paddlers, the breakdown of verbal instruction demands additional attention from the auditory modality, subsequently disrupting focus on their current goals.

After completing a test session with the CoOP system, PaddlerLV1 noted how the absence of verbal commands improved his focus, from my transcripts:

> Before, I did paddle without [CoOP] and I had people around me, they were always telling me left, right, or check what's ahead of you. Now that I am able to paddle [with CoOP] I don't have to worry about objects on the water where I might crash or hit and I can just focus on technique or whatever I am working on that day whether it's the catch or balancing the canoe. – *PaddlerLV1 Interview*

Here, PaddlerLV1 is describing one advantage to being remotely steered with CoOP. He acknowledges his desire to focus on a particular goal without distraction from navigational commands. When verbal commands that are required (*e.g.*, paddling motion, situational context) are constantly repeated the advantage of the CoOP system's ability to reduce verbal communication is lost as coach and paddler engage in the types of cyclical conversation quoted earlier between PaddlerB1 and CoachS3. Supporting paddler focus, by avoiding unnecessary verbal communication enables paddlers to attend to non-verbal sounds, such as those discussed in Section 4.2.3. Paddlers rely on non-verbal auditory information to build context within their environment. From the sounds of water against the hull of their canoe, which serve as an auditory meter for cadence, to familiar sounds along a particular route that represent landmarks. The practices that paddlers have adopted serve to enrich the paddling experience through deeper immersion in their activity. During a race event in which CoachS2, CoachS1, and CoachS3were present and actively supporting PaddlerB1 and PaddlerLV1 I observed that CoachS3 had remained relatively quiet throughout the event. When asked, his response acknowledged the distraction that too much verbal communication can have on the paddlers, from my field notes:

> At about the halfway point of the race, I noticed [CoachS3], who usually provides [PaddlerB1] with a lot of feedback, pep talk, and contextual information, was

> fairly quiet. I asked him why and he said, "You know, too many cooks in the kitchen. I don't want to overwhelm him." – *Field notes*

With both CoachS1 and CoachS2 providing PaddlerB1 with verbal information, CoachS3 recognized that the addition of his verbal input would only serve to disrupt PaddlerB1's focus. Finding balance between helpful and excessive verbal information can be challenging, particularly when multiple sighted support crew are involved in a paddling activity. Each sighted person makes their own observations about the state of the paddler and paddling conditions, and must decide whether that information is relevant and how it will infringe on paddler focus. While accompanying CoachS3 during early training sessions, I frequently made observations on the water that I shared with PaddlerB1. As our process developed, we recognized how multiple voices added unnecessary complexity to the activity. Over time, CoachS3 adopted the role of communicator, while I relayed my observations to him. Not only did this practice reduce the sources of verbal information, but it allowed CoachS3 to filter what he thought was relevant for the paddler at a given moment.

As I described earlier, verbal information pertaining to the motion of the canoe is critical to ensuring paddler safety. However, when the water is calm with little activity, coaches use the opportunity to both teach and introduce descriptive language to enrich paddler experience. The following excerpt from one training session, captures how coaches take advantage of in-depth instruction and offer situational context when the water is calm:

> CoachS1– That's the thing, we want you to be strong right at the acceleration point.
>
> PaddlerB1– Yeah
>
> CoachS1– Because if you are strong at the back of the stroke you're not gonna be going anywhere.
>
> CoachS1– Alright, let's do another five, catch your breath.

PaddlerB1– What you're saying...this arm here [holds up left arm]?

CoachS1– Yeah.

(loud sounds of water splashing nearby from a hose)

CoachS3– Yep. And [PaddlerB1], what you're hearing, that water, is someone washing their boat.

(canoe and support boat drift closer to moorings)

CoachS1– Okay, yeah we don't want you to get wet (laughter), so let's go...3.2.1.

([PaddlerB1] starts paddling as we follow in the support boat)

CoachS1– There you go, slight bend. Okay, on your other side now, I want that wrist back, slight bend.

PaddlerB1– Okay.

CoachS3– There you go, boom. boom. boom. boom. All your power is right there...boom. (pacing PaddlerB1's stroke)

– *Transcripts*

In this exchange, CoachS3 waits until CoachS1 and PaddlerB1 have completed their discussion about arm position to describe the non-verbal sounds occurring nearby. The spraying sounds from a nearby yacht were loud enough to capture everyone's attention. CoachS3, recognizing the sudden injection of a new sound in an otherwise quiet area, used the opportunity to provide PaddlerB1 with additional context about the situation. Even though this was a passing moment, we were only in audible range of the sounds for a few minutes, these types of descriptions provide paddlers with contextual information about the area. Many of the locations encountered along paddling routes have become so familiar that paddlers recognize where they are by the auditorily unique sounds they produce. For example, paddlers have developed a tradition of shouting “Go Makapo!" when passing under a bridge which sits at approximately one mile from the canoe landing. Coaches and other support crew who take the opportunity to verbally communicate context in this way help paddlers build their own mental models of the surrounding area.

The previous exchange also highlights another way coaches rely on auditory information to teach paddlers. As soon as PaddlerB1 found a paddling rhythm, CoachS3 started to reinforce his motion through verbal messaging. As I described earlier, PaddlerLV1 relies on sounds of water against the canoe as a metric for cadence. Here, CoachS3 verbally provides cadence for PaddlerB1 with sustained repetition using the word "boom." In both cases, auditory information is a significant factor in the development of paddler skill. The more information that a paddler has available, the better equipped they are to accomplish their goals. While the majority of auditory information that coaches provide is aligned towards learning, there is also a desire to ensure that paddlers retain as much of the sighted experience as possible. The most consistent use of auditory information throughout my time with Makapo was the verbal signaling of waves. The following observation from my field notes captures the experience:

> As a series of waves approached us, the coaches signaled to the paddlers to prepare for the "bump." DirectorLV1 and PaddlerLV1 were able to identify the angle of the waves and aligned their canoes perpendicularly as CoachS3 aligned PaddlerB1 using CoOP. As the first wave approached, CoachS3 provided a countdown to its arrival. Everyone started paddling faster in anticipation of the wave in an attempt to stay on it as long as possible. After the waves passed, I asked the paddlers if they could feel their speed increase. Everyone agreed, commenting that the waves were fun to ride as the canoes move considerably. CoachS1 said, "Definitely lots of fun, you can get a bump from even the smallest waves, including wakes from the boats that pass us, but the big ones are the most fun." – *Field notes*

The waves described during this particular session were naturally occurring; however, as CoachS1 points out, paddlers look for bumps from all types currents on the water. DirectorLV1, whose visual acuity is strong enough to identify wakes from passing boats, frequently positioned himself to “ride", or capture momentum, from the wakes of large yachts. For other paddlers who are not able to visually identify waves, coaches position the canoe using CoOP, inform the paddler of an inbound wave, and then verbally indicate the wave’s approach. The subsequent increase is felt by the paddlers who are free to time their paddle stroke to stay on the wave as long as possible.

In this section, I have described the ways verbal and non-verbal auditory cues enrich individual experiences for blind and low vision paddlers. Auditory information plays an important role in supporting paddling activities by providing an immersive experience, supporting skill development, and ensuring that paddlers remain safe. Despite the importance of auditory communication, identifying a consistent and reliable strategy for communicating auditorily remains an ongoing challenge for Makapo. In the next section, I discuss breakdowns in auditory communication.

### Challenges of Auditory Communication

As both a sport and leisure activity, outrigger canoeing with visually impaired paddlers depends on clear auditory communication to be successful. Supporting blind and low-vision paddlers on the water presents numerous challenges that make communication and interpretation of auditory information difficult. In this section, I describe the ways that natural and artificial auditory artifacts disrupt communication between coach and paddler.

The most persistent challenge to reliable auditory communication was proximity between support boat and canoe and ambient sounds. Ambient noises generated by weather conditions, the watercraft used by the team, boats, and operations of other people, fluctuated constantly during each paddling session. The variation in audibility from ambient noise was compounded by the shifting proximity between support boat and paddler, one of several technical issues captured in my field notes:

> One of the challenges that we have with verbal communication is maintaining a consistent distance between the launch boat and canoe. The throttle on the launch boat is sticky, which makes it hard to dial in at a position that keeps pace with the paddler. This forces the boat driver to constantly speed up or slow down to stay within an audible distance with the canoe. – *Field notes*

The undulation that arose from the throttle issue introduced an additional layer of coordination to ensure verbal communication between coach and paddler was successful. Coaches either had to steer the support boat into audible range or project verbal instructions loudly to ensure paddler comprehension. When using CoOP, the preferred position for accurate steering is to align the support boat directly behind the canoe. Steering CoOP from behind the canoe allows the steerperson to more easily assess canoe movement, which changes constantly from water currents and wind, and make micro-adjustments to hold a straight line. Conversely, the front-back alignment prevents paddlers from communicating with coaches unless they turn their heads to the left or right, which often leads to an unbalanced body position than can flip the canoe. However, additional challenges arise when ambient noises on the water require a close proximity, as noted in my field notes:

> As we approached the four mile mark in the race, as our launch boat pulled up alongside PaddlerB1 he flipped the canoe. We approached pretty quickly about a two feet away from the left side of his canoe. The sentiment among the coaches

> was that the boat wake may have surprised him and partially instigated the flip.
> – *Field notes*

The “flip" described in this quote is an edge case; coaches frequently pulled alongside canoes without incident. However, it draws attention to how the tension between proximity, paddler control, and verbal communication leaves coaches with one more variable to factor into their practice. Attempts to ameliorate the effects of ambient sounds and proximity remain largely unsuccessful. Makapo coaches attempted a variety of configurations using two-way radios. The earliest effort involved attaching a radio to the paddler, as captured in my field notes:

> Today the team attempted to use an inexpensive two-way radio system for communication. CoachS3 attached the handset to PaddlerB1's belt, ran the microphone wire along his back, and clipped the microphone to the shoulder strap of his pack. To communicate with the support boat, PaddlerB1 had to reach to his chest and press the talk button attached to the microphone wire. It was working well until we reached the one mile marker, where PaddlerB1's paddling motion eventually dislodged the handset from his belt and it fell into the water. – *Field notes*

The constant motion of the body made it difficult to use traditional two-way radios effectively. Although radios designed for similarly high-motion activities like outrigger paddling do exist, the added expense pushes against Makapo's desire to ensure the activity remain accessible to their members. Subsequent attempts to use two-way radios were slightly more successful, however, coaches found that the availability of a communication interface inferred a need for paddlers to respond. The process of removing one hand from the paddler to activate the microphone disrupted paddler focus. Eventually, the microphone was removed from the paddler's radio entirely and the handset was set in output mode only. Using this configuration, the coaches could communicate one-way with paddlers, removing the expectation of response. While this approach simplified configuration and interaction for both coach and paddler, the process was cumbersome enough that the team eventually reserved use of the radio for races. I asked the team how the radio was working during an interview, from my transcripts:

> CoachS1 – "Yeah, I think for the most part it went well. Although PaddlerLV1 said it was a bit static-y, which made it hard to hear."
>
> DirectorLV1 – "It was pretty hard to hear. Maybe because it was in a bag."
>
> ResearcherS1 – "How was it [radio] attached to the canoe?"
>
> CoachS1 – "We just tied it down in the storage area behind the seat." – *Transcripts*

The radios that the team used were waterproof, but not submersible. To keep the radio dry, CoachS1 sealed it in an water resistant bag for additional protection. The responses from CoachS1 and DirectorLV1 highlight the challenges faced when introducing technological solutions to a water-based activity like outrigger paddling. Although radio communication addressed issues of proximity and ambient noise, obstacles like clarity and reliability brought new challenges. As DirectorLV1noted, the bag may have contributed to the audibility issues with the radio, but its location behind the paddler was likely suboptimal. As I described in section 4.2.4, attempts to communicate verbally from behind the paddler often went unheard as the ambient noises in front of the paddler dominated their auditory channel, an effect that may have disrupted audibility of the radio as well. Certainly, a higher quality waterproof radio combined with a custom point of attachment in front of the paddler would alleviate some of the issues I have described. However, even without audibility issues, verbal communication can be challenging. The following transcripts are from two different sessions during which two blind paddlers, PaddlerB1 and PaddlerB2 were required to steer themselves by interpreting verbal instructions.

Transcript of CoachS3 instructing PaddlerB1 on how to steer the canoe[3]:

CoachS3 – “tap left, that’s it"

CoachS3 – “tap right, find center"

CoachS3 – “there you go, keep working it."

CoachS3 – “tap right, that’s it...find center...tap left [PaddlerB1], find center. Alright."

CoachS3 – “tap left"

CoachS3 – “tap right...and tap left, you’re kind of slaloming back and forth, tap right. Right there."

CoachS3 – “tap left, there you go, perfect, keep that line"

CoachS3 – “alright, tap right, alright right there, hold that line."

CoachS3 – “tap right"

CoachS3 – “tap left, there you go"

CoachS3 – [louder] “tap right, there you go, right there, hold that line"

CoachS3 – [louder] “tap left...there you go...keep working it, getting near the bridge"

CoachS3 – “tap right, and hold it there, tap left, remember to find the center after you tap"

PaddlerB1 – “Okay"

In calm conditions on the water, pedals only need a slight adjustment to keep the canoe moving in a straight line. If the pedal adjustments are small and quick enough, the canoe will “hold a line," as CoachS3 instructs. In the instance presented above, PaddlerB1 was

[3]Full transcript can be found in Appendix A.2

oversteering by either pressing the pedal all the way down or holding it in a turn position too long after each command, causing the canoe to move in an 's' like pattern. Pedal motion is fluid, so PaddlerB1's only ability to differentiate between small and large movements is proprioceptively, by assessing the relative position change between each of his feet. Furthermore, in the canoes used by Makapo, the pedals are not self-centering (some canoe models do offer this feature), requiring the paddler to manually return the pedals to center after each turn. The next transcript is from a similar experience shared by CoachS1 and PaddlerB2.

Transcript of CoachS1 instructing PaddlerB2 how to steer the canoe[4]:

CoachS1 – "Okay you want to press to your right side and paddle"

CoachS1 – "Hard right"

CoachS1 – "Paddle straight"

CoachS1 – "Keep pushing to your right"

CoachS1 – "Hard right"

CoachS1 – "Hard right, keep going to your right"

CoachS1 – "Stay on the right"

CoachS1 – "Stay on the right"

CoachS1 – "Stay on the right"

CoachS1 – "Okay even them out a little bit"

CoachS1 – "Straight ahead, little bit to your left"

CoachS1 – "Okay, little right, good"

CoachS1 – "Gonna be that way the whole time"

( 30 seconds pass)

[4]Full transcript can be found in Appendix A.1

CoachS1 – "Okay little right, just a hair to the left"

During both of the sessions captured here, CoachS1 and CoachS3 were unexpectedly required to provide *ad hoc* verbal instruction to the paddlers after the CoOP system had stopped working. Both sessions reveal how frequently the rudder of the outrigger canoe needs to be adjusted to maintain a straight path. The variation in language that both coaches use, and the resulting outcomes from each paddler, reflect the challenges of using informal language in an interdependent activity. Both paddlers had little to no concept of how the commands "tap", "hard" press, "hold that line", or "straight ahead" translated to pedal movement. For CoachS3, the effect was a slaloming motion as PaddlerB1 oversteered the canoe with each press of the pedal. Whereas CoachS1 and PaddlerB2 were competing against a strong inbound tidal current that kept the canoe moving to the left, demanding that PaddlerB2 keep the pedals in an offset position. The selection and use of metaphors to describe the types of actions that a paddler should take depend on the individual experiences of each paddler as well as the coaches interpretation of the actions the paddler is taking. CoachS1 eventually adjusted his language to be more descriptive, from later in the session:

CoachS1 – "Okay try not to tap too hard so just little bits at a time"

CoachS1 – "Okay left, left"

CoachS1 – "Okay,right"

CoachS1 – "Little right"

CoachS1 – "Small tap left"

CoachS1 – "Small tap right"

...

CoachS1 – "Tap left"

CoachS1 – "Small tap right"

CoachS1 – "Small tap left, There you go, left a bit"

...

CoachS1 – "Okay, you are about five feet from shore, slow down"

CoachS1 – "Stop stop stop stop stop stop stop STOP STOP STOP"

The previous quote exemplifies the evolution of CoachS1's verbal commands over the course of the fifteen minutes that PaddlerB2 was in control of the canoe. CoachS1 recognized the need to communicate the different amounts of pressure that PaddlerB2 should apply to the pedals. Once PaddlerB2 realized that the canoe responded differently to variations in pressure on the pedal, he was able to reduce the amount of oversteer that was occurring and keep the boat along a straighter path, which he described in a post session interview:

> It was odd, when [CoOP] was taken off remote, [the canoe] felt like it was hard to push the pedals. I had to push pretty hard and then I felt a click in the back, and then the peddles moved easier. Then I also understood CoachS1, at first I wasn't understanding when he said tap, and now I know when he says small tap left you just tap the pedal and then take your foot off. *PaddlerB2 – Transcripts*

Most of the issues that both PaddlerB2and PaddlerB1 experienced occurred from a lack of preparation due to the unanticipated need to self steer, which CoachS1 later acknowledged:

> I think that's the other thing, we need to write up the dialogue and go over it with [the paddlers] so they know small tap, hold, etc. So maybe I will write up a thing on what the different commands are. *CoachS1 – Transcripts*

Here, CoachS1 acknowledges how the gap in preparation of verbal commands in a self steering condition added undesirable complexity to the situation. CoachS1's interest in formalizing coach—paddler dialogue for conditions when verbal communication is required, highlights his desire to ensure a positive experience for Makapo paddlers. The breakdowns that coaches and paddlers experienced, draw attention to the need for a more structured approach to mixed-ability outrigger paddling. Implementing a formal training session where paddlers learn how commands map to specific physical movements, as CoachS1 suggests, could mitigate the obstacles described in these quotes lead to improved experiences for everyone involved. CoachS1's recognition of the need for a coaching vocabulary came at the end of our multi-year engagement together, after more than fifty hours of practice sessions with CoOP. Although the breakdown of the CoOP system brought the need for a common vocabulary into focus, the long-term, sustained use of the system provided CoachS1 with the knowledge to make it happen.

As I have described in this section, audibility on the water is inconsistent at best. Technological workarounds hold promise for supporting reliable communication, but low-cost multi-purpose systems introduce new obstacles. By offsetting a significant portion of verbal communication to a sighted steerperson, CoOP does free the auditory channel to focus on other auditory artifacts. However, an over-dependence on CoOP can lead to conditions that neither paddler or steerperson are prepared to handle. These instances also suggest that there may be an opportunity to rethink how CoOP is used. The conditions that led to CoOP's failure for both of the sessions presented here were vastly different. For CoachS3 and PaddlerB1, a lack of adequate adhesion when installing CoOP caused it to fall off the canoe. For CoachS1 and PaddlerB2, a subsequent investigation revealed that the control receiver pins used for the motor connection were corroded. While the causes of these failures can be easily remedied, by experiencing failure during live training conditions, we discovered edge conditions that the design team had yet to consider. Once a formal dialogue is operationalized by both coach and paddler, there may be conditions in which CoOP could be disconnected, transferring full control of the canoe to the paddler. At this level of function, CoOP's role is transformed from a shared-control device to a shared-support device, aligning more closely with the types of patterns employed while using a white cane (Hersh, 2015; Ball & Nicolle, 2015) or guide dog (Lloyd, Budge, Stafford, & La Grow, 2009).

While I framed the two sessions during which paddlers steered by verbal command as system breakdowns, it's important to note that neither paddler took issue with having to steer themselves. For PaddlerB1, who was focused on training, the constant oversteering was counterproductive, but not unejoyable. During a post-session interview, PaddlerB2 expressed disappointment that the system failed, stating "I wish I had more time out." A structured, systematic approach, as CoachS1 suggests, could resolve many of the issues I have presented. Rather than rely on CoOP for all forms of blind and low-vision one-person outrigger paddling, a pre-established communication protocol would enable paddlers and coaches to *choose* when to use CoOP.

## 4.3 Conclusion

In this chapter, I presented the results of a multi-year public co-design project with sighted and visually impaired outrigger canoe paddlers. Through the co-design and development of a shared control steering system called CoOP, I along with my design team, empowered blind paddlers to paddle alone in one-person canoes. I identified four major themes that can inform how designers and researchers think about assistive technology design. First, public-facing co-design with sighted and visually impaired communities brings awareness of the disabled experience to the surrounding community. Second, leveraging existing economies and DIY practices can support the creation of functional novel assistive technologies. Third, the physicality and interdependence of real-world activity presents challenges and opportunities for multi-modal assistive technology design. This interdependent approach would never have been achieved without the insights of a mixed-ability, public-facing, co-design team. Fourth, the auditory channel is responsible for interpreting a broad range of information that must be factored into assistive technology design. Taken together, these results remind us that not only must we consider assistive technologies in light of context and abilities of primary users, but that interdependent users—made explicitly here but implicit in so many other assistive technologies—have important perspectives as well.

# Chapter 5

# Discussion

In this dissertation, I consider how audio-only assistive technologies augmented with additional modalities can reduce dependence on audio, improving the experience of activities for blind and low vision users. I examined how people with visual impairments interpret and use auditory forms of information across two distinct learning contexts: computer skills training and outrigger paddling. In absence of the visual cues that drive the learning process within these contexts, the information required to build knowledge is translated to speech and non-speech sounds. When placed alongside the natural and artificial noises produced in everyday life, these sounds are unremitting. They fill spaces in equally useful and useless patterns, leaving the individual to mediate the relevance of each sound they hear, a process further confounded by the nature of sound. From its linear structure to its ephemeral existence, auditory information is unquestionably different than the information humans process visually. Adapting sound to communicate visual information in assistive technologies must be carried out with a contextual understanding of the auditory environment in which it is designed to serve.

My motivation to help shape our scholarly understanding of the relationship between sound and assistive technology flows from the observations I made working alongside the visually impaired community over the course of my doctoral studies. I observed how, in the absence of sight, the auditory channel incurs an increased responsibility for navigating the world. Designing assistive technologies to render information through sound compounds the amount of work the auditory channel must perform. The auditory channel can not be expected to simply transition to such a dualistic role without first understanding the context, environment, and activities that drive the need for auditory information. Sound, and the information it produces, is far richer than we might imagine, as LaBelle, in his exploration of sound in culture and everyday life explains:

> From one body to the other, a thread is made that stitches the two together in a temporal instant, while remaining loose, slack, to unfurl back into the general humdrum of place. Sound might be heard to say, *This is our moment*...In the movement of sound, the making of an exchange is enacted; a place is generated by the temporality of the auditory. *This is our moment* is also immediately, *This is our place. – (LaBelle, 2010)*

In the passage above LaBelle, reflecting on an encounter in which he observes a father explaining sound to his son, posits how sound, through the creation of knowledge and experience, offers a lens to examine our "contemporary condition" (LaBelle, 2010). Sound shapes our world, momentarily, leaving a vocabulary for social exchange as it passes. This ephemerality, according to LaBelle, is a feature through which all of sounds incarnations, bind us to a moment in time. He continues:

> [Sound] exists as a network that teaches us how to belong, to find place, as well how not to belong, to drift...Auditory knowledge is non-dualistic. It is based on empathy *and* divergence, allowing for careful understanding and deep involve-

> ment in the present while connecting to the dynamics of mediation, displacement, and virtuality.– *(LaBelle, 2010)*

For LaBelle, the world around us can be understood through the acoustic experiences it reveals. Listening, therefore, becomes an indispensable action for interpreting and engaging with the sounds that we encounter. How then might we, as designers of technologies that assist people who are visually impaired, identify and respect the boundaries of antecedent forms of auditory knowledge in our work? From LaBelle's perspective, sound is more than just the temporal information it produces. The ways in which sound intervenes in an activity, shaping the course it might take; moves from place to place, building an auditory landscape; and enables the imagination to give form to what is unseen; suggest how a framework that accommodates the nuances of sound can reshape our approach to the design of nonvisual assistive technologies.

To understand how the sounds that are produced within auditory contexts influence the actions of the individual, in Chapter 1, I asked the questions: *How can tangible and mixed ability computational systems be designed to reduce audio and interactions for blind and low vision people in everyday activities?* and *do assistive technologies that are less reliant on auditory input improve blind and low vision user interactions?* By augmenting traditional sonic outputs through computationally driven interventions, I examined how removing certain types of auditory information can improve experiences for blind and low vision individuals. By working alongside the visually impaired community, I identified design strategies for reducing the dependency on audio in nonvisual assistive technologies. In this chapter, I demonstrate how the results presented in chapters 3 and 4 answer these questions.

## 5.1 Towards a Multimodal Activity Centered Interaction Space

The work I presented in chapters 3 and 4, explores two approaches to reducing the amount of auditory information required to complete an activity in absence of visual information. My work on the Tangible Desktop suggests that introducing additional modalities for interpreting information holds promise for improving the enjoyment and efficiency of general computing tasks. The multimodal approach to nonvisual computing employed by the Tangible Desktop system, demonstrated one way to effectively compliment audio-only output in assistive technology systems. By creating physical representations of contextual auditory phenomena like computer programs or document structure, the Tangible Desktop simplified the stream of auditory information produced by a screen reader, freeing the listener to focus on other sources of auditory knowledge. In this section, I discuss how the Tangible Desktop contributes to scholarly thought on multimodal nonvisual computer interaction. I then present design considerations for the application of an activity centered approach to multimodal interfaces.

### 5.1.1 Identifying New Modalities

As I described in chapter 2, scholars have explored multimodal and tangible forms of computer interaction for much of HCI's history. Individually, concepts such as positional constraints (Poll & Waterham, 1995), haptics (L. Brown et al., 2005; Xu et al., 2011), and physical icons (Ishii & Ullmer, 1997) have been shown to be meaningful for nonvisual and sighted users alike. Despite the positive outcomes that continue to emerge from multimodal research, we have yet to see a shift towards incorporating these interactions into the day-to-day activities of people who are visually impaired. There are numerous explanations for the disparity between utility and adoption, from market share to ecosystem, moving complex multimodal systems from research to production is a daunting task. Additionally, the widespread adoption of smartphones that incorporate increasingly powerful use of text-to-speech and vibrotactile feedback (Kane, Jayant, Wobbrock, & Ladner, 2009; Rodrigues, Montague, Nicolau, & Guerreiro, 2015), provide a powerful platform for nonvisual computational tasks (*e.g.*, (Dim & Ren, 2014; Wang, Yu, Yang, He, & Shi, 2019; Maiero et al., 2019) as recent examples). What the Tangible Desktop and KInD offer then, is not a prescriptive approach for designing multimodal interfaces, but evidence of the importance of understanding the role of sound in the auditory channel.

Through my field work at Empowertech, I identified how the structural information required to contextualize auditory knowledge was largely inconsequential to a given task. For example, a person browsing a web page needs to know whether information is a heading, actionable (*e.g.*, hyperlink or button), or content; however, once they receive that information it serves no further purpose. Communicating this information auditorily through text-to-speech, therefore, is unnecessarily cumbersome. By transferring this contextual information to the haptic channel, the Tangible Desktop isolates the knowledge a person seeks to the auditory channel. The advantages to this approach, which I describe in chapter 3 section 3.3.1, serve to free the auditory channel for a range of new activities. A person exploring a document, for example, might rely on their physical channel to locate a specific section while simultaneously engaged in a conversation with a peer. Similarly, the absence of contextual information in the audio stream reduces the amount of auditory information a person must process, thereby reducing the time to listen to content.

The improved task times I observed with participants in my experimental study of the Tangible Desktop, indicate that a system designed to utilize multiple modalities is beneficial for basic computing tasks. One concern with a tangible approach is the risk of increasing the demand on cognitive load by moving recognition from the auditory channel to the tactile. When scaled to an everyday working environment with a larger set of identifiers, the burden on recall could outweigh the advantages gained by direct access. The KInD system offers one potential alternative by re-framing the desktop interaction paradigm around activity. In the next section, I reflect on how core concepts of activity theory can inform how we think about designing computational systems for nonvisual use that fundamentally rethink how computational information is structured and communicated.

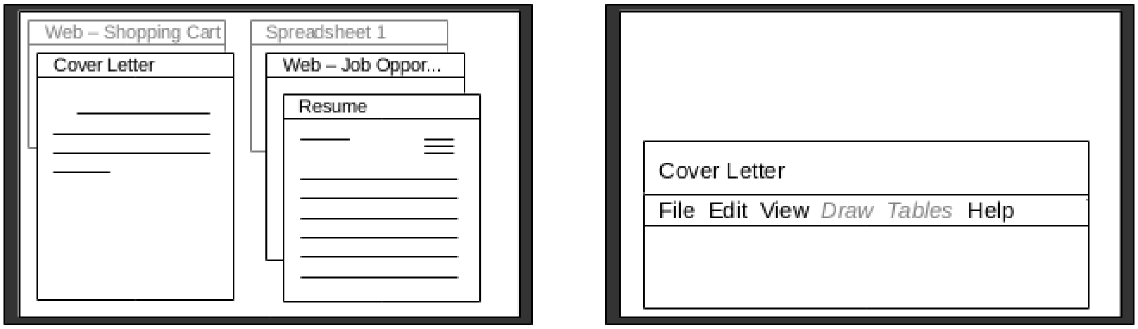
### 5.1.2 Applying an Activity Theory Frame to Support Multiple Modalities

The operationalizing loop described in section 3.4.3, the idea that we cease to "see" the tools we are using as they become a natural part of the interaction, frames much of what I view as the end goal for accessible tools. Current audio-first assistive technologies, such as screen readers, assert themselves aggressively into the workflow of a nonvisual computer user. The white cane, as a counter example, is a dramatic alternative that is quite literally perceived as a physical extension of the blind user's arm (Chebat et al., 2018). Framing nonvisual interaction by activity provides a mechanism by which we can raise the quality of interaction with nonvisual digital objects to that of the cane, and operationalize interactions for blind users.

Prior research (Nardi, 1996a; Kaptelinin & Nardi, 2006) notes that unconscious operation arises through the unique properties of both individual and interaction. Thus, attempting to design for unconscious operation directly may be somewhat impractical. What I offer here, then, is not a prescribed set of requirements or implications for design. Rather, I propose multiple pathways to operational actions such that designers who restructure their interactions around the notion of activity can improve the potential for unconscious operations to emerge with time, particularly for nonvisual computer users for whom traditional approaches are noisy and interfering. Specifically, nonvisual systems should be redesigned from the ground up with a consideration for the fundamental goals of visually impaired users rather than act as translators for the tasks sighted users might complete to accomplish those same goals. Second, a consideration for activity must not end with the initial goal but must be allowed to adapt and change over time. Finally, in addition to considering that blind users might engage different tasks than sighted users, an activity-centered perspective to assistive technology design explicitly considers the way these tasks are carried out. Next, I describe the considerations designers must engage to merge nonvisual technologies with the more natural orientation of human activity.

### Tasks Vary by Activity, Interaction, and Ability

In a computational setting similar to those I presented in chapter 3, an activity-centered model might support attention flow as needed to the tools and tasks that make up the overall activity. Like the carpenter's work bench, when hammering a nail, the unneeded saw rests out of sight and mind. In a desktop computing environment, sighted users can ignore icons, windows, and content that do not relate to the task at hand. For the nonvisual user, however, screen translation tools do not filter for contextually relevant information, forcing unnecessary content into the processing stream. Filtering out unwanted information, to the extent it is possible, requires workarounds such as increased rate of speech and keyboard shortcuts. The problems that arise from these workarounds are well established (Lazar et al., 2006, 2007; Frick, Gower, Kempen, & Wolff, 2007; Bell & Mino, 2015) as well as attempts at resolving them (Vigo & Harper, 2013a; A. Brown & Harper, 2013; Plessers et al., 2005). The issue an orientation towards activity highlights, however, goes beyond these workarounds and their challenges. Rather, the problem rests in the inability of the system to truly understand the activity, and therefore its sub-tasks. The system adheres to an interaction model that does not match user needs, the driving force behind human activity (Kaptelinin & Nardi, 2006).




(a) Activity Level

(b) Application Level

Figure 5.1: Use of activity hierarchy for nonvisual interfaces. (a) A typical desktop environment in which multiple applications are open. Since the user is currently focused on writing a cover letter, unrelated applications (*i.e.,* Shopping and Spreadsheet) are not included in the audio space. (b) The system is aware of the requirements to create a resume cover letter and automatically removes unnecessary application features (*i.e.,* Draw and Tables) from the audio space.

Existing computer hardware already has the power to track, model, and act on far more about the people using it than typically gets employed in practice. Even in much simpler systems, such as the Victor Reader which I described in chapter 3, a more appropriate interaction model can present a rapid path towards unconscious operations. The Victor Reader works, in large part, because its task is not something that sighted computer users ever do. It was built from the beginning for a different interaction model, rather than retrofitted on top of an interaction model made for people with sight. Input techniques and their associated hardware for more fully featured systems should similarly be redesigned with a consideration for the specific needs, then tasks, then interactions of visually impaired users.

When interaction is modeled on need, comprehension is transferred from the user to the system. Intuitively, a user expects only certain features to accomplish a particular task. Rather than present all features equally, a system or tool should present only the features required to complete the task. Using the hierarchical structure of activity, I illustrate the concept of need through the following scenario:

**At the level of activity** A user needs to write a cover letter for a resume. The nonvisual activity centered system recognizes this need as one that only requires a subset of the applications available on the system (Figure 5.1 (a)). For example, while working on a resume a user might require a word processor to write the letter, a copy of the resume, and a web browser to reference a job application. While the activity is active, from the perspective of the user, these three tools are the only computational entities available in the system.

**At the level of application** The word processor that is being used to draft the cover letter, recognizes the need and limits its feature set to only interactions suitable for letter writing (Figure 5.1 (b)). From the perspective of the user, features such as tables, drawings, images, and macros do not exist.

At both levels, the audio stream is isolated from unnecessary and unsolicited entities running on the system, thus narrowing the presentation of information to match user expectation. As the scenario depicted in figure 5.1 highlights, isolation does not remove entities from the environment, rather, they are simply not communicated according to the context of the current activity. Just as the sighted user chooses to ignore unrelated visual information, the isolated audio stream enables nonvisual users to ignore unrelated audio information. Likewise, as the sighted user can quickly shift attention to previously ignored entities, a nonvisual user can expand outside of the isolated context to support shifting needs. We see interaction models built upon user need as an opportunity for designers to remove the obstacles that delay unconscious operation. The natural binding between need and expectation enables greater flexibility in how computational entities are presented nonvisually.

### Make Change a First Class Citizen

Activities are shaped by "virtue of their differing objects (Nardi, 1996b)." When applied to computational systems, the unique footprint that objects give to activity is commonly used as the basis for modeling activity centered interactions (Voida & Mynatt, 2009; Bardram & E., 2005; Kaptelinin, 2003). An activity-centered approach then emphasizes the ability to pause, port, and resume activities alongside the ability to understand what it is that the system is doing on behalf of the user. This emphasis requires that the user know what has changed in the interface while working on an activity, between moments of pause, and across multiple platforms. Unfortunately, as I observed in my fieldwork and as many researchers have noted (*e.g.*, (S. A. Brewster, Wright, & Edwards, 1993; Ratanasit & Moore, 2005; M. S. Baldwin et al., 2017; E. Mynatt, 1997; Guerreiro & Gonçalves, 2014)), audio is the primary mode of interaction for nonvisual computing users, and it is ephemeral and sequential and for the most part, individual and non-portable.

For most people, the visuospatial sketchpad is more capable of retaining extensive information than the phonological loop (Baddeley, 1992). This increased retention enables users to quickly scan the computing context visually to gain a firm understanding of what has changed in the interface. For nonvisual users, if the interface is complex, comparing it to the prior version may require extensive listening, well beyond the capabilities of the phonological loop, and requiring a re–listening to the audio simply to orient oneself to the current system state.

Therefore, reconstructing the task to integrate change into the set of base interactions that a nonvisual user has available introduces a path for simplyfying change management. Orienting a system towards activity reduces the complexity of managing change in an interaction model. The activity is, or at least could be, responsible for the state of all objects in use for task completion. Placing this responsibility on the activity removes it from the user, freeing them from having to manage system level tasks (*e.g.,* locate, open, save). Furthermore, by assigning the role of object management to the system, a record of changes within the activity is captured, effectively generating a timeline of interaction events that support change. For the nonvisual user, change management represents a significant reduction in the amount of auditory information that must be processed during task completion.

### Design for Goals Not Applications

As noted above, much of desktop computing for blind and low vision users is accomplished by translating the activities, tasks, and interactions of sighted users into a nonvisual paradigm. One of the primary limitations of this approach is the need to work across applications and the limited ability to translate visual information effectively in the multi-application computing environment. Because a nonvisual user might carry out a particular activity using a drastically different set of tasks than a sighted user might use, a nonvisual user is likely to make use of applications in different ways and possibly use different applications. Thus, any computing system built on the notion of activity would need to consider differences in task decomposition and its associated applications.

To speed up working with nonvisual computing systems, many advanced users memorize a variety of commands. Efficient computer use often only occurs upon mastering recall of dozens of commands. Even after reaching this level of mastery, however, users still have work to do any time a new application needs to be accessed. Commands can have similar actions across a variety of applications, different actions even for very similar applications, or simply be repurposed to meet the changing requirements of a single application. An activity centered approach would ensure that similar interactions map to similar tasks and have similar outcomes. A strict visual to auditory translation of existing interfaces cannot achieve this, but transformation of actions can, and designers of accessible systems should work to this ultimate end goal.

Activity serves as a natural extension for computer interaction by closely aligning with user expectation to support unconscious operations. An activity centered approach introduces ways for designers to address many of the known barriers in nonvisual computing, such as excessive audio, unexpected behaviors, and command memorization (M. S. Baldwin et al., 2017). Solving these challenges alone does not completely fix the problems associated with nonvisual computing, such as accessibility compliance (Richards et al., 2012; Mankoff et al., 2005). Computing is ever more complex, requiring rethinking the entire interaction and architecture from the ground up for a truly accessible experience. Rather than adapt the interaction and tasks common to sighted users, activity theory suggests a different way of thinking about design for accessibility, one that provides the fundamental restructuring I have articulated here.

## 5.2 Co-Design by Activity

Organizing the visual to nonvisual translation of information around activity presents myriad ways to support operationalizing user interaction. Designing to support the individual needs and motives of the user introduces more natural ways to present computational information. In the last section I demonstrated how principles of activity theory can be used to rethink the design of nonvisual desktop computing systems. Yet, the utility of activity theory extends beyond the boundaries of the computer operating system (Clemmensen et al., 2016). In this section, I discuss how activity theory can be applied to assistive technology design more broadly. I start by reflecting on the public-facing, co-design work I completed with Makapo. I describe how public input, reciprocity with partners, and open iterative design, positively contributed to the development of CoOP. I conclude with a discussion on how active participation and emphasis on user activity supports the design process.

### 5.2.1 Public Co-Design

The public-facing, co-design process that I carried out alongside my community partners revealed insights that guided the iterative process, brought awareness to the blind paddling community, and set a path for long term viability. Individually, each of these practices contribute to the types of positive outcomes that I experienced in my work with Makapo. My results highlight the ways that public co-design with mixed-ability groups can support the development of assistive technology. In particular, the publicity and the physicality of the public design work contributed to improved awareness of blind athletes, as well as the eventual inclusion of people with a variety of abilities in activities surrounding design and use of the resultant technology (*i.e.,* CoOP). People began to understand what we were doing, why it was potentially relevant to them, the abilities of the blind paddlers sharing space with sighted paddlers, and that they could contribute to our work with Makapo. In this section, I reflect on what contributed to the success of this project and present strategies for conducting public co-design to build assistive technology.

#### Support Public Input

While mixed-fidelity prototypes serve to invite critique, public input in the design process can supply crucial design considerations for and raise community awareness of the disability community. Technology is frequently described as a force for inclusion, enabling people with disabilities to participate in activities that might otherwise be unavailable to them. Similarly, a significant effort has been put forth to formulate methodological approaches that include people with disability in the design process (Morrison et al., 2017; Shinohara et al., 2018; Wobbrock et al., 2011; Mankoff et al., 2010; Meissner et al., 2017). What I saw in this work, however, is another potential future, one in which people with disabilities and others with different abilities can participate together, not only in the design and use of assistive

Figure 5.2: A typical day on the wash deck. The Makapo practice team installs CoOP, an activity that frequently piqued the interest of nearby paddlers.

technologies but also in broader contexts (*e.g.,* in sports or recreation).

As Bennett et al. argued, an interdependent perspective "considers everyone and everything in an interaction to be mutually reliant" (Bennett et al., 2018). Design with a mixed-ability group in a public space can serve to affirm to the broader community that people with disabilities are not simply "recipients", as Bennett et al. point out, but designers, facilitators, and participants in the co-design process. We identified mixed-ability challenges *in situ* rather than through a structured evaluation (see Figure 5.2), enabling us to refocus our design requirements after each session. The introduction of a tactile switch to CoOP, inspired by a conversation about the tactile qualities of Velcro (see Section 4.2.1), exemplifies the ways that mixed-ability co-design can support assistive technology design. As a sighted research team designing for a blind and low vision community, providing our end users with prototypes early and often directly shaped our decisions.

Although our process grew organically from the conditions in which we conducted our work, our inclusive design practice, *ad hoc* iterative process, and public dialogue align closely with Teal and French's Designed Engagement (Teal & French, 2016). Similar to their "pop-up" engagement approach, many of our design insights were drawn spontaneously from the public, rather than through formal recruitment methods employed by traditional participatory design methods. However, our work diverges from Designed Engagement in two distinct ways that inform assistive technology design practice. First, our design team included members from the community that our work was designed to serve, integrating the co-design principles articulated by Sanders and Stappers into our public engagement. Second, our work remained situated within a single location for an extended period of time, establishing a consistent presence for the community to both actively and passively observe. Collectively, these practices contributed to the increased public attentiveness to inclusivity and community interest I described in Chapter 4.

### Facilitate Rapport and Trust Through Reciprocity

Rapport and trust can be crucial for obtaining and maintaining access to a community or research site. While values-centric participatory design approaches like co-design can build rapport and trust between researcher and community partners through methods like reciprocity (Le Dantec & Fox, 2015; Gregory, 2003), explicit actions to drive trust are less mature, and far from solved. Next, I describe two approaches that were pivotal for our relationship with our community partner—but acknowledge that they may not work for all community partners and contexts.

First, an ongoing issue with participatory design stems from concerns from community partners that researchers are just "mining" the community for data (Le Dantec & Fox, 2015). DirectorLV1's depiction of Makapo's preliminary efforts in Section 4.2.3 expressed their caution towards researchers pursuing designs that align closer to research than community goals.

Although I was also initially imagining a different research path, I was able to build rapport by shifting my research trajectory to align with Makapo's immediate need for a training tool that could be used without researcher supervision. However, mere verbal agreements did not assuage DirectorLV1's fears of misaligned goals. It was not until the first short-lived test on the water that I was viewed as taking Makapo's goals seriously. After which, DirectorLV1 began investing more of their resources and enthusiasm into this project. Using mixed-fidelity rapid prototyping—although not robust—served to demonstrate to Makapo that our goals were aligned.

Second, to maintain our rapport with Makapo, I wanted to ensure that they didn't feel like I was developing something that would be unusable after our partnership ended. Managing the transition from research to application of technological systems can pose challenges for co-design projects. As Hayes explained, it is not enough to simply leave an intervention behind, researchers must ensure that partner organizations have the resources to utilize and maintain deployments (Hayes, 2011). As I described in Chapter 4, the DIY approach to design and co-design practice, assured that CoOP could be maintained by our community partners at the end of the project. For example, I frequently disassembled the prototypes, described each component, and shared cost estimates and acquisition sources with community members. This process helped demystify CoOP, which instilled Makapo with the confidence that they would be able to reproduce CoOP independent of my involvement.

### Mixed-Fidelity Prototypes

Our practice of placing early iterations of CoOP in view of the public served to generate community interest. Many of our sessions displayed incomplete or semi-functional prototypes that were optimized to evaluate specific features. As is common with traditional rapid prototyping design iterations, I was focused on evaluating early and often before committing to high quality output, which provided the knock-on effect of giving both blind and sighted paddlers—involved with Makapo and not—the ability to critique and respond to the designs.

In our first evaluation session, I was unprepared for the level of interest I received from the public, leaving concerns about how early prototypes reflected my abilities as a designer and researcher. As a researcher stepping into a new community, I worried that presenting non-functioning, incomplete work would undermine community confidence. However, an analysis of my preliminary session revealed how public input contributed to the design process. As the work progressed, I felt more comfortable adopting a strategy of showcasing the rapid prototyping approaches to the public. I openly accepted and attributed guidance on design and engineering decisions in future sessions.

The mixed-fidelity nature of the prototypes helped to manage expectations, passively—but explicitly—inviting commentary on components that were being tested (*e.g.,* does this make sense at all vs can it last hours). In traditional user-centered design practices, the researcher typically embraces a leadership role within the project (Sanders & Stappers, 2008), but my use of mixed-fidelity prototypes empowered community partners to share ownership and opinions of various aspects of the project.

### 5.2.2 Designing by Activity and Active Participation

I return to activity theory as a guiding framework for its characterization of how community and mediating tools move the individual (subject) towards an outcome (object(s)) by supporting the completion of goals. Assistive technologies are, unmistakably, representative of the mediating tools that modern interpretations of activity theory build upon. Thus, it holds that applying activity theory principles to systems of activity where an assistive technology is employed can be a natural and beneficial strategy.

The development of the CoOP system was guided by a desire to support outrigger canoeing skill development. By adopting an orientation towards activity, motivating factors present in traditional HCI based assistive technology research (*e.g.*, novel interactions and performance improvements) were de-emphasized in service of the pursuit of activity. This is not to say that traditional assistive technology research is absent value, far from it, as these types of endeavors certainly bring valuable insights to light. Rather, I draw attention to how research attuned to individual activities can further answer critical questions about the conceptualization, design, and implementation of assistive technology.

As Kaptelinin and Nardi encourage, an emphasis on activity broadens technological research to include not just operators and consumers of a particular interaction space, but "developing human beings who create meaning in their lives through acting with technology (Kaptelinin & Nardi, 2006). In my work, I have found this perspective relevant as a non-disabled researcher, unable to share the lived experiences of the community that I serve. During CoOP's multi-year deployment, although I explored ways to make the system more robust (*e.g.*, resolving degradation of electrical components from salt water), these were largely procedural and in service of keeping the system operational. Yet, by actively immersing myself in the activities carried out by my community partners, I learned how the technology I co-created brought meaning to those who used it, bringing forth new questions that shaped future design. While working on CoOP, I learned how to paddle an outrigger canoe. Activity theory argues that tool use is most effective when the individual can engage in unconscious operation. Sharing in the types of interactions a tool is expected to support brings designers of assistive technology closer to understanding the obstacles and constraints that affect operation. For example, by learning to paddle, I identified through experience the types of operations required to paddle effectively. Recognizing how these operations (*e.g.*, balance, body position, stroke) interacted with each other over the course of the activity influenced how design aspects of CoOP were implemented, specifically in support of unconscious operation. I view this form of active participation as an essential process for unmasking design constraints, building rapport, and establishing trust in a way that traditional participatory and in-the-wild practices might not reveal.

Gay and Hembrooke, writing on the transformative power of mediating tools, describe how tools "*disclose behaviors and social phenomena that have remained hidden and unexamined, even unimagined, because no technologies existed to reveal them.*" (Gay & Hembrooke, 2004). While this quote holds true for the study of all mediating tools, regardless of theoretical perspective, Gay and Hembrooke are referring here to tool use from the viewpoint of activity theory. Gay and Hembrooke view systems of activity as fluid, "characterized by ambiguity and change" (Gay & Hembrooke, 2004). The measure of activity before, during, and after technological design work reveals the "unexamined" and "unimagined" through observation of changes across an activity system. As an assistive technology intervention, the CoOP system did not intrinsically offer new insights for interaction, rather, it reappropriated existing components for a different purpose. Yet, as a mediating tool in support of activity, an entirely novel space for mixed-ability engagement was created. Within that space, I examined how CoOP transformed the coach-paddler relationship, affected paddler agency, and shifted social perceptions towards the visually impaired community. These observations emerged from changes that occurred across the communities within which CoOP was used, over time. The CoOP system accomplished its primary goal by the end of the original ten month design and evaluation period, providing Makapo with a tool to improve paddler skill. However, through prolonged use across leisure, training, and competitive activities, CoOP challenged the way I, as well as my co-researchers and community partners, imagined its purpose. The breakdowns that occurred with the CoOP system that I describe in section 4.1 of chapter 4, for example, changed the relationship between paddler and coach. For the paddler, although the overall activity (outrigger canoe training) remained the same, the loss of mediating tool (CoOP) transformed the objects required to reach the desired outcome from balance and stroke to balance, stroke, and steer. This change highlights a point of contrast worthy of deeper exploration: How does the additional object (steering) and the actions required to support it affect the paddler/coach relationship? As described in my results, changing environmental conditions can drastically influence verbal communication, how might changes in audibility determine the use of CoOP? Are there situations in which paddler steering control is desired? These types of questions become possible when designing at the level of activity.

At several points in this dissertation I have presented the work of Bennett et al. (Bennett et al., 2018) in support of the mixed-ability shared control design of the CoOP system. The interdependent frame that Bennett et al. introduce to HCI, thoughtfully persuades researchers and designers to consider the types of natural, supportive interactions that exist between disabled and non-disabled individuals, arguing that in many cases these interactions are desirable and should not be replaced through technological intervention (Bennett et al., 2018). Rather, Bennett et al. encourage their readers to apply an interdependent frame to assistive technology design so that these mixed-ability interactions might be supported. Such an expansion, from an individual to individuals, gives rise to a community. Designing for a community of mixed-ability people sharing an assistive technology, brings with it a myriad of nuanced design considerations in need of empirical investigation. Activity theory lends itself well to the many types of perspectives that inevitably arise when designing to concurrently support differing abilities by acknowledging the relationships that exist between an individual, mediating tool, and community in an activity system.

The concept of community, introduced by Yrjö Engeström alongside *rules* and the *division of labor*, broadened the conceptual model of activity theory to include the "societal and collaborative nature" of individual actions in an activity (Engeström et al., 1999). Engeström recognized how activities carried out by a subject give rise to a collaboration, and the collaboration inevitably leads to a community (*e.g.*, a bricklayer who depends on a cement maker and brick maker to build a wall leads to a masons union). Through the application of Bennett et al.'s interdependent frame, communities, both new and existing, will arise or reconfigure to support assistive technology designed for interdependence. In Engeström's activity system, assistive technology designers might view these communities as external structures of support, ensuring that the subject completes their goals. For example, the community that formed around CoOP served to support the goals of Makapo paddlers, ensuring that they successfully completed a paddling activity. However, what is absent from this view are the intangible benefits of community cooperation, which I argue are worth considering as more than supporting structures in an activity.

While building the CoOP system, I viewed the paddling community as the social and collaborative foundation through which a paddler completes their goals. But in learning how to use CoOP, the community transformed. Though still a social collective centered around outrigger paddling, new roles emerged (*e.g.*, steerperson, guide, launch and landing coordinator), new rules were defined (*e.g.*, verbal command and coaching language, support crew coordination), and new designs were started in support of this evolving community. The experiences, practices, stories, and artifacts that led to this emergent community, inform how members collaborate and improve their activities (Wenger, 1999). Viewed from this perspective, a community is more than a support structure for the subject working towards an object, it becomes the object of an activity.

Recent scholarly efforts, noting weaknesses in Engeström's conceptualization of community (a concern confirmed by Engeström himself (Engeström, 2009)) argue in support of the view that community is an object/outcome, not simply a parameter of the subject's activity (Taylor, 2009). This as an important distinction for the application of activity theory to interdependent assistive technology design. Measuring the outcome of an activity by the types of communities that it might form, feeds the design directions of the assistive technology. With CoOP, I recognized the importance of community early in the design process. For example, the unsolicited feedback that the design team received during CoOP's development demonstrated how public engagement could lead to community structures that might not otherwise exist. As CoOP moved from intervention to functional system, this insight motivated discussions with Makapo about ways CoOP might facilitate paddler agency, leading to the design of interactions with CoOP that could be performed without sight (see section 4.2.1). We envisioned a scenario where a blind paddler, through ownership of CoOP, could arrive at a paddling facility and find a paddling partner—another paddler willing to attach the steering portion of CoOP to their canoe to navigate the blind paddler. We viewed this is an outcome; a mixed-ability cooperative paddling community where blind and sighted paddlers could spontaneously decide to paddle together.

Interdependence is intertwined with awareness, issues of stigma, and public support of disability. Applying an activity theory model to interdependent assistive technology enables designers to look beyond the design constraints of a mixed-ability dyad and towards the broader community as a source for design insight.

# Chapter 6

# Conclusion

In this dissertation, I examined how people with blindness access and act upon sound through assistive technology. I demonstrated how the auditory modality is a rich source of knowledge about the world in both its strengths and limitations. And I presented strategies for building a deeper understanding of auditory knowledge to inform assistive technology design. Grounded through long-term field studies, my work points to a design space that has yet to be fully explored.

The unique qualities of sound, its linear structure, ephemerality, movement, virtuality, and familiarity are characteristics that when understood from within the contexts in which they occur present radical new ways to think about the design of assistive technology. The observation of screen reader use in a visually impaired computer training class, as one example from my work, highlighted how disruptive the text-to-speech output of a screen reader can be to the local soundscape. The ephemerality of sound furthers the screen reader's disruptive nature by forcing a model of repetition. Expanding beyond the audio based model of computer interaction to tangible representations of visual information, therefore, is a far more suitable structure for certain types of auditory information. Similarly, the observation of auditory knowledge creation during an outrigger canoe paddling activity, as another example from my work, highlights how the displacement and familiarity of sound provide anchor points that tether a paddler to a place in space and time. By adopting a practice of shared control with a sighted partner, the blind paddler's auditory channel is unburdened from verbal navigation commands, and thus free to create these anchor points.

These examples, though dramatically different, demonstrate how attending to context and sound reveal critical insights that can be used to guide the design of assistive technology. How then can designers interested in reducing dependence on audio in assistive technologies bring these insights to light? My work examines one path, but there are certainly others worth exploring. By applying an activity theory lens to the interests of my research communities, I framed the assistive technologies that I created around the needs, goals, and motives of community members. This change in perspective, from application to activity, is a distinction that proved beneficial within the contexts of my research. Further work is needed to understand how activity centered design can be applied across multiple assistive technology contexts. Additionally, my results would never have been achieved without the insights drawn from the extensive time spent with my research partners. In particular, the mixed-ability, public-facing, co-design team that contributed to the development of the CoOP system, collectively brought that work to light.

Taken together, this work reminds us that not only must we consider assistive technologies in light of context and abilities of primary users, but that interdependent users—made explicitly here but implicit in so many other assistive technologies—have important perspectives as well. In some ways, assistive technologies have often considered the allies and the co-users. I found in public-facing interactions, however, that other participants—the bystanders, the curious, the nearby—became essential parts of the design process and ultimately product development and testing as well. By bringing design out of the shadows—and in the case of my work with Makapo quite literally into the bright sunshine of an ocean dock—assistive technology researchers and designers can bring work from other contexts into the light as well.

# References

Abdolrahmani, A., & Kuber, R. (2016). Should i trust it when i cannot see it?: Credibility assessment for blind web users. In *Proceedings of the 18th international acm sigaccess conference on computers and accessibility* (pp. 191–199). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2982142.2982173 doi: 10.1145/2982142.2982173

Abner, G., & Lahm, E. (2002). Implementation of assistive technology with students who are visually impaired: Teachers' readiness. *Journal of Visual Impairment & Blindness (JVIB)*.

Abras, C., Maloney-Krichmar, D., Preece, J., et al. (2004). User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, *37*(4), 445–456.

Administration, U. G. S. (2020). *Web Content Accessibility Guidelines (WCAG).* Retrieved 2020-08-01, from https://www.section508.gov/

Alma, M. A., Van der Mei, S. F., Groothoff, J. W., & Suurmeijer, T. P. (2012). Determinants of social participation of visually impaired older adults. *Quality of Life Research*, *21*(1), 87–97.

Andrews, C. (2014). Accessible participatory design: engaging and including visually impaired participants. In *Inclusive designing* (pp. 201–210). Springer.

Ängeslevä, J., Oakley, I., Hughes, S., & O'Modhrain, S. (2003). Body mnemonics portable device interaction design concept. , *3*, 2–5.

Asakawa, C., Takagi, H., Ino, S., & Ifukube, T. (2002). Auditory and tactile interfaces for representing the visual effects on the web. *Proceedings of the fifth . . .* .

Avila Soto, M., Funk, M., Hoppe, M., Boldt, R., Wolf, K., & Henze, N. (2017). Dronenavigator: Using leashed and free-floating quadcopters to navigate visually impaired travelers. In *Proceedings of the 19th international acm sigaccess conference on computers and accessibility* (pp. 300–304).

Azenkot, S., Feng, C., & Cakmak, M. (2016). Enabling building service robots to guide blind people a participatory design approach. In *2016 11th acm/ieee international conference on human-robot interaction (hri)* (pp. 3–10).

Baddeley, A. (1992). Working memory and conscious awareness. In *Theories of memory* (pp. 11–20). Lawrence Erlbaum Associates.

Baldwin, C. L. (2016). *Auditory cognition and human performance: Research and applica-*

*tions*. CRC Press.

Baldwin, M. S., Hayes, G. R., Haimson, O. L., Mankoff, J., & Hudson, S. E. (2017). The tangible desktop: A multimodal approach to nonvisual computing. *ACM Transactions on Accessible Computing (TACCESS)*, *10*(3), 9.

Baldwin, M. S., Hirano, S. H., Mankoff, J., & Hayes, G. R. (2019, November). Design in the public square: Supporting assistive technology design through public mixed-ability cooperation. *Proc. ACM Hum.-Comput. Interact.*, *3*(CSCW). Retrieved from `https://doi.org/10.1145/3359257` doi: 10.1145/3359257

Baldwin, M. S., Mankoff, J., Nardi, B., & Hayes, G. R. (Submitted). An activity centered approach to nonvisual computer interaction. *ACM Transactions on Computer-Human Interaction*.

Ball, E. M., & Nicolle, C. A. (2015). Changing what it means to be âĂIJnormalâĂİ: A grounded theory study of the mobility choices of people who are blind or have low vision. *Journal of Visual Impairment & Blindness*, *109*(4), 291–301.

Bardram, J. E., Bunde-Pedersen, J., & Soegaard, M. (n.d.). Support for Activity-Based Computing in a Personal Computing Operating System..

Bardram, J. E., & E. (2005, sep). Activity-based computing: support for mobility and collaboration in ubiquitous computing. *Personal and Ubiquitous Computing*, *9*(5), 312–322. doi: 10.1007/s00779-004-0335-2

Bardzell, S., Rosner, D. K., & Bardzell, J. (2012). Crafting quality in design: Integrity, creativity, and public sensibility. In *Proceedings of the designing interactive systems conference* (pp. 11–20). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2317956.2317959` doi: 10.1145/2317956.2317959

Barr, J. J., & Bracchitta, K. (2015). Attitudes toward individuals with disabilities: The effects of contact with different disability types. *Current Psychology*, *34*(2), 223–238.

Bau, O., Poupyrev, Israr, a., & Harrison. (2010). TeslaTouch: electrovibration for touch surfaces. *Proceedings of UIST 2010*, 283–292. doi: 10.1145/1866029.1866074

Bell, E. C., & Mino, N. M. (2015). Employment Outcomes for Blind and Visually Impaired Adults.

Bennett, C. L., Brady, E., & Branham, S. M. (2018). Interdependence as a frame for assistive technology research and design. In *Proceedings of the 20th international acm sigaccess conference on computers and accessibility* (pp. 161–173). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3234695.3236348` doi: 10.1145/3234695.3236348

Bigham, J. P., Jayant, C., Ji, H., Little, G., Miller, A., Miller, R. C., ... others (2010). Vizwiz: nearly real-time answers to visual questions. In *Proceedings of the 23nd annual acm symposium on user interface software and technology* (pp. 333–342).

Billah, S. M., Ashok, V., Porter, D. E., & Ramakrishnan, I. (2017). Ubiquitous accessibility for people with visual impairments: Are we there yet? In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 5862–5868). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3025453.3025731` doi: 10.1145/3025453.3025731

Blattner, M. (1989). Earcons and icons: Their structure and common design principles. *HumanâĂŞComputer ...*. Retrieved from `http://www.tandfonline.com/doi/abs/10.1207/s15327051hci0401{_}1`

Bødker, S. (1990). *Through the interface: A human activity approach to user interface design*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc.

Bødker, S. (1991). Through the interface-a human activity approach to user interface design. *DAIMI Report Series*, *16*(224).

Bødker, S., & Andersen, P. B. (2005). Complex mediation. *Human-computer interaction*, *20*(4), 353–402.

Borodin, Y., Bigham, J. P., Dausch, G., & Ramakrishnan, I. V. (2010). More than meets the eye: a survey of screen-reader browsing strategies. In (p. 13). ACM.

Boyd, L. (1990). The Graphical User Interface Crisis: Danger and Opportunity. Retrieved from `http://eric.ed.gov/?id=ED333687`

Branham, S. M., & Kane, S. K. (2015). The invisible work of accessibility: how blind employees manage accessibility in mixed-ability workplaces. In *Proceedings of the 17th international acm sigaccess conference on computers & accessibility* (pp. 163–171).

Brewster, S., & Brown, L. M. (2004). Tactons: Structured Tactile Messages for Non-visual Information Display. In (pp. 15–23). Darlinghurst, Australia, Australia: Australian Computer Society, Inc. Retrieved from `http://dl.acm.org/citation.cfm?id=976310.976313http://dl.acm.org/ft{_}gateway.cfm?id=976313{&}type=pdf`

Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*(April), 222–227. Retrieved from `http://dx.doi.org/10.1145/169059.169179` doi: http://doi.acm.org/10.1145/169059.169179

Brown, A., & Harper, S. (2013). Dynamic injection of WAI-ARIA into web content. *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility - W4A '13*, 1. Retrieved from `http://www.scopus.com/inward/record.url?eid=2-s2.0-84879810872{&}partnerID=tZOtx3y1` doi: 10.1145/2461121.2461141

Brown, L., Brewster, S., & Purchase, H. (2005). A First Investigation into the Effectiveness of Tactons. In *First joint eurohaptics conference and symposium on haptic interfaces for virtual environment and teleoperator systems* (pp. 167–176). IEEE. Retrieved from `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1406930` doi: 10.1109/WHC.2005.6

Challis, B. P., & Edwards, A. D. (2000). Design principles for tactile interaction. In *International workshop on haptic human-computer interaction* (pp. 17–24).

Chebat, D.-R., Harrar, V., Kupers, R., Maidenbaum, S., Amedi, A., & Ptito, M. (2018). Sensory substitution and the neural correlates of navigation in blindness. In *Mobility of visually impaired people* (pp. 167–200). Springer.

Cimarolli, V. R. (2002). The impact of perceived overprotection on adjustment to age-related vision loss.

Cimarolli, V. R., Reinhardt, J. P., & Horowitz, A. (2006). Perceived overprotection: support gone bad? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *61*(1), S18–S23.

Clemmensen, T., Kaptelinin, V., & Nardi, B. (2016). Making hci theory work: an analysis of the use of activity theory in hci research. *Behaviour & Information Technology*, *35*(8), 608–627.

Collignon, O., & De Volder, A. G. (2009). Further evidence that congenitally blind participants react faster to auditory and tactile spatial targets. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *63*(4), 287.

Cornet, V., Voida, S., & Holden, R. J. (2018). Activity theory analysis of heart failure self-care. *Mind, Culture, and Activity*, *25*(1), 22–39.

Cortes, U., Barrue, C., Martinez, A. B., Urdiales, C., Campana, F., Annicchiarico, R., & Caltagirone, C. (2010). Assistive technologies for the new generation of senior citizens: the share-it approach. *International Journal of Computers in Healthcare*, *1*(1), 35–65.

Crispien, K., Fellbaum, K., Savidis, A., & Stephanidis, C. (1996). A 3d-auditory environment for hierarchical navigation in non-visual interaction..

Dabbish, L., Mark, G., & González, V. M. (2011). Why do i keep interrupting myself?: Environment, habit and self-interruption. , 3127–3130. Retrieved from `http://doi.acm.org/10.1145/1978942.1979405` doi: 10.1145/1978942.1979405

Dakopoulos, D., & Bourbakis, N. G. (2009). Wearable obstacle avoidance electronic travel aids for blind: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(1), 25–35.

Dakopoulos, D., & Bourbakis, N. G. (2010). Wearable obstacle avoidance electronic travel aids for blind: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(1), 25–35.

DeWalt, K. M., DeWalt, B. R., & Wayland, C. B. (1998). Participant observation. *Handbook of methods in cultural anthropology*, 259–300.

Dim, N. K., & Ren, X. (2014). Designing motion gesture interfaces in mobile phones for blind people. *Journal of Computer Science and technology*, *29*(5), 812–824.

Donker, H., Klante, P., & Gorny, P. (2002). The Design of Auditory User Interfaces for Blind Users. *Proceedings of the second Nordic conference on Human-computer interaction*, 149–155. doi: 10.1145/572020.572038

Ducasse, J., Macé, M. J., Serrano, M., & Jouffrais, C. (2016). Tangible reels: construction and exploration of tangible maps by visually impaired users. In *Proceedings of the 2016 chi conference on human factors in computing systems* (pp. 2186–2197).

Edwards, W. K., Mynatt, E. D., & Stockton, K. (1994). Providing access to graphical user interfacesâĂŤnot graphical screens. In (pp. 47–54). ACM. Retrieved from `http://dl.acm.org/citation.cfm?id=191041https://smartech.gatech.edu/bitstream/handle/1853/3579/95-35.pdf`

Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic fieldnotes*. University of Chicago Press.

Engeström, Y. (2009). The future of activity theory: A rough draft. *Learning and expanding with activity theory*, 303–328.

Engeström, Y., et al. (1999). Activity theory and individual and social transformation. *Perspectives on activity theory*, *19*(38), 19–30.

Fiannaca, A., Apostolopoulous, I., & Folmer, E. (2014). Headlock: a wearable navigation aid that helps blind cane users traverse large open spaces. In *Proceedings of the 16th international acm sigaccess conference on computers & accessibility* (pp. 19–26).

Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, *8*(5), 347–358. doi: 10.1007/s00779-004-0297-4

Folmer, E., & Morelli, T. (2012). Spatial gestures using a tactile-proprioceptive display. In (pp. 139–142). ACM. Retrieved from `http://dl.acm.org/`

citation.cfm?id=2148161http://www.tiii.be/wp-content/uploads/2012/03/139-102-folmer.done{_}.pdf

Foulke, E., & Sticht, T. (1969). Review of research on the intelligibility and comprehension of accelerated speech. *Psychological bulletin*.

Frick, K. D., Gower, E. W., Kempen, J. H., & Wolff, J. L. (2007). Economic Impact of Visual Impairment and Blindness in the United States. *Archives of ophthalmology*, *125*(4), 544. doi: 10.1001/archopht.125.4.544

Gaver, W. W. (1989, March). The sonicfinder: An interface that uses auditory icons. *Hum.-Comput. Interact.*, *4*(1), 67–94. Retrieved from http://dx.doi.org/10.1207/s15327051hci0401_3 doi: 10.1207/s15327051hci0401_3

Gay, G., & Hembrooke, H. (2004). *Activity-centered design: An ecological approach to designing smart tools and usable systems.* Mit Press.

González, V. M., & Mark, G. (2004). "constant, constant, multi-tasking craziness": Managing multiple working spheres. , 113–120. Retrieved from http://doi.acm.org/10.1145/985692.985707 doi: 10.1145/985692.985707

Google. (2019). *Google drive: Free cloud storage for personal use.* Retrieved 2019-03-25, from https://www.google.com/drive

Gregory, J. (2003). Scandinavian approaches to participatory design. *International Journal of Engineering Education*, *19*(1), 62–74.

Grudin, J. (2008). A moving target: The evolution of hci. *The human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications*, 1–24.

Guerreiro, J., & Gonçalves, D. (2014). Text-to-speeches: Evaluating the perception of concurrent speech by blind people. *ASSETS14 - Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility*, 169–176. doi: 10.1145/2661334.2661367

Guerreiro, J., Rodrigues, A., Montague, K., Guerreiro, T., Nicolau, H., & Gonçalves, D. (2015). Tablets get physical: non-visual text entry on tablet devices. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 39–42).

Günther, S., Müller, F., Funk, M., Kirchner, J., Dezfuli, N., & Mühlhäuser, M. (2018). Tactileglove: Assistive spatial guidance in 3d space through vibrotactile navigation. In *Proceedings of the 11th pervasive technologies related to assistive environments conference* (pp. 273–280).

Gupta, S., Morris, D., Patel, S. N., & Tan, D. (2013). AirWave: Non-Contact Haptic Feedback Using Air Vortex Rings. *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13*, 419–428. doi: 10.1145/2493432.2493463

Hamidi, F., Baljko, M., Kunic, T., & Feraday, R. (2014). Do-it-yourself (diy) assistive technology: a communication board case study. In *International conference on computers for handicapped persons* (pp. 287–294).

Hayes, G. R. (2011, August). The relationship of action research to human-computer interaction. *ACM Trans. Comput.-Hum. Interact.*, *18*(3), 15:1–15:20. Retrieved from http://doi.acm.org/10.1145/1993060.1993065 doi: 10.1145/1993060.1993065

Hayes, G. R. (2014). Knowing by doing: action research as an approach to hci. In *Ways of knowing in hci* (pp. 49–68). Springer.

Hersh, M. (2015). Cane use and late onset visual impairment. *Technology and Disability*, *27*(3), 103–116.

Hötting, K., & Röder, B. (2009). Auditory and auditory-tactile processing in congenitally blind humans. *Hearing research*, *258*(1-2), 165–174.

Hugdahl, K., Ek, M., Takio, F., Rintee, T., Tuomainen, J., Haarala, C., & Hämäläinen, H. (2004). Blind individuals show enhanced perceptual and attentional sensitivity for identification of speech sounds. *Cognitive brain research*, *19*(1), 28–32.

Hull, J. (1992). *Touching the rock: An experience of blindness.* Vintage Books. Retrieved from `https://books.google.com/books?id=FcKGDwAAQBAJ`

Hurst, A., & Tobias, J. (2011). Empowering individuals with do-it-yourself assistive technology. In *The proceedings of the 13th international acm sigaccess conference on computers and accessibility* (pp. 11–18).

Igarashi, T., Moscovich, T., & Hughes, J. F. (2005). As-rigid-as-possible shape manipulation. In *Acm siggraph 2005 papers* (pp. 1134–1141). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/1186822.1073323` doi: 10.1145/1186822.1073323

Inc.", G. (2019). *GoPro suction cup camera mount.* Retrieved 2019-03-25, from `https://shop.gopro.com/mounts/suction-cup/AUCMT-302.html`

Ishii, H., & Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the acm sigchi conference on human factors in computing systems* (pp. 234–241). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/258549.258715` doi: 10.1145/258549.258715

Jansen, Y., Karrer, T., & Borchers, J. (2010). MudPad : Tactile Feedback and Haptic Texture Overlay for Touch Surfaces. *Proc. of ACM Int.Conf. on ITS*, 11–14. doi: 10.1145/1936652.1936655

Jayant, C., Ji, H., White, S., & Bigham, J. P. (2011). Supporting blind photography. In *The proceedings of the 13th international acm sigaccess conference on computers and accessibility* (pp. 203–210). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2049536.2049573` doi: 10.1145/2049536.2049573

Kane, S. K., Hurst, A., Buehler, E., Carrington, P., & Williams, M. A. (2014). Collaboratively designing assistive technology. *interactions*, *21*(2), 78–81.

Kane, S. K., Jayant, C., Wobbrock, J. O., & Ladner, R. E. (2009). Freedom to roam: a study of mobile device adoption and accessibility for people with visual and motor disabilities. In *Proceedings of the 11th international acm sigaccess conference on computers and accessibility* (pp. 115–122).

Kaptelinin, V. (2003). Umea: Translating interaction histories into project contexts. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 353–360). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/642611.642673` doi: 10.1145/642611.642673

Kaptelinin, V., & Nardi, B. A. (2006). *Acting with technology: Activity theory and interaction design.* MIT press.

Kaptelinin, V., Nardi, B. A., & Macaulay, C. (1999). Methods & tools: The activity checklist: a tool for representing the âĂIJspaceâĂİ of context. *interactions*, *6*(4), 27–39.

Keith, J. M., Bennetto, L., & Rogge, R. D. (2015). The relationship between contact and attitudes: Reducing prejudice toward individuals with intellectual and developmental

disabilities. *Research in developmental disabilities*, *47*, 14–26.
Klatzky, R. L., & Lederman, S. J. (2003). Touch. *Handbook of psychology*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/0471264385.wei0406/full
Klein, S., Adams, G., Dickin, F., & Simske, S. (2013). 3D Printing: When and where does it make sense? In (Vol. 2013, pp. 5–8). Society for Imaging Science and Technology.
Köhlmann, W., Zinke, F., Schiewe, M., & Jürgensen, H. (2010). User-interface filter for two-dimensional haptic interaction. In *International conference on computers for handicapped persons* (pp. 498–505).
Könings, K. D., Seidel, T., & van Merriënboer, J. J. (2014). Participatory design of learning environments: integrating perspectives of students, teachers, and designers. *Instructional Science*, *42*(1), 1–9.
Koutsoklenis, A., & Papadopoulos, K. (2011). Auditory cues used for wayfinding in urban environments by individuals with visual impairments. *Journal of Visual Impairment & Blindness*, *105*(10), 703–714.
Kuber, R., Yu, W., & O'Modhrain, M. S. (2010). Tactile Web Browsing for Blind Users. *Haptic and Audio Interaction Design*, *6306*, 75–84.
Kurniawan, S. H., & Sutcliffe, A. (2002, Jul). Mental models of blind users in the windows environment. In *Computers helping people with special needs* (p. 568âĂŞ574). Springer, Berlin, Heidelberg. Retrieved from https://link.springer.com/chapter/10.1007/3-540-45491-8_109 doi: 10.1007/3-540-45491-8_109
Kuznetsov, S., Dey, A. K., & Hudson, S. E. (2009). The effectiveness of haptic cues as an assistive technology for human memory. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5538 LNCS*, 168–175. doi: 10.1007/978-3-642-01516-8_12
LaBelle, B. (2010). *Acoustic territories: Sound culture and everyday life*. Bloomsbury Publishing USA.
Lacey, G., & MacNamara, S. (2000). Context-aware shared control of a robot mobility aid for the elderly blind. *The International Journal of Robotics Research*, *19*(11), 1054–1065.
Lazar, J., Allen, A., Kleinman, J., & Malarkey, C. (2007, may). What Frustrates Screen Reader Users on the Web: A Study of 100 Blind Users. *International Journal of Human-Computer Interaction*, *22*(3), 247–269. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/10447310709336964 doi: 10.1080/10447310709336964
Lazar, J., Feng, J., & Allen, A. (2006). Determining the Impact of Computer Frustration on the Mood of Blind Users Browsing the Web. *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*, 149–156. doi: 10.1145/1168987.1169013
Le Dantec, C. A., & Fox, S. (2015). Strangers at the gate: Gaining access, building rapport, and co-constructing community-based research. In *Proceedings of the 18th acm conference on computer supported cooperative work & social computing* (pp. 1348–1358). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2675133.2675147 doi: 10.1145/2675133.2675147
Leont'ev, A. N. (1974, Dec). The problem of activity in psychology. *Soviet Psychology*, *13*(2), 4âĂŞ33. doi: 10.2753/RPO1061-040513024
Leont'ev, A. N. (1978). Activity, consciousness, and personality.

Leontiev, A. N. (1981). Problems of the development of the mind.

Lessard, N., Paré, M., Lepore, F., & Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, *395*(6699), 278.

Levine, S. P., Bell, D. A., & Koren, Y. (1994). Navchair: An example of a shared-control system for assistive technologies. In *International conference on computers for handicapped persons* (pp. 136–143).

Lloyd, J. K., Budge, R. C., Stafford, K. J., & La Grow, S. J. (2009). A focus group discussion on using guide dogs. *International Journal of Orientation & Mobility*, *2*, 52–64.

Maiero, J., Eibich, D., Kruijff, E., Hinkenjann, A., Stuerzlinger, W., Benko, H., & Ghinea, G. (2019). Back-of-device force feedback improves touchscreen interaction for mobile devices. *IEEE transactions on haptics*, *12*(4), 483–496.

Manduchi, R., & Kurniawan, S. (2018). *Assistive technology for blindness and low vision*. CRC Press.

Mankoff, J., Fait, H., & Tran, T. (2005). Is your web page accessible?: a comparative study of methods for assessing web page accessibility for the blind. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 41–50).

Mankoff, J., Hayes, G. R., & Kasnitz, D. (2010). Disability studies as a source of critical inquiry for the field of assistive technology. In *Proceedings of the 12th international acm sigaccess conference on computers and accessibility* (pp. 3–10).

Manshad, M. S., Pontelli, E., & Manshad, S. J. (2012). Trackable Interactive Multimodal Manipulatives :. (Figure 2), 664–671.

Mark, G., Iqbal, S. T., Czerwinski, M., & Johns, P. (2014). Bored mondays and focused afternoons: The rhythm of attention and online activity in the workplace. , 3025–3034. Retrieved from `http://doi.acm.org/10.1145/2556288.2557204` doi: 10.1145/2556288.2557204

McDaniel, T., Bala, S., Rosenthal, J., Tadayon, R., Tadayon, A., & Panchanathan, S. (2014). Affective haptics for enhancing access to social interactions for individuals who are blind. In *International conference on universal access in human-computer interaction* (pp. 419–429).

McGookin, D., Robertson, E., & Brewster, S. (2010). Clutching at straws: using tangible interaction to provide non-visual access to graphs. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1715–1724).

Meissner, J. L., Vines, J., McLaughlin, J., Nappey, T., Maksimova, J., & Wright, P. (2017). Do-it-yourself empowerment as experienced by novice makers with disabilities. In *Proceedings of the 2017 conference on designing interactive systems* (pp. 1053–1065). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3064663.3064674` doi: 10.1145/3064663.3064674

Menelas, B.-A. J., Picinali, L., Bourdot, P., & Katz, B. F. (2014). Non-visual identification, localization, and selection of entities of interest in a 3d environment. *Journal on Multimodal User Interfaces*, *8*(3), 243–256.

Morrison, C., Cutrell, E., Dhareshwar, A., Doherty, K., Thieme, A., & Taylor, A. (2017). Imagining artificial intelligence applications with people with visual disabilities using tactile ideation. In *Proceedings of the 19th international acm sigaccess conference on computers and accessibility* (pp. 81–90). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3132525.3132530` doi: 10.1145/3132525.3132530

Muehlbradt, A., Koushik, V., & Kane, S. K. (2017). Goby: A wearable swimming aid for blind athletes. In *Proceedings of the 19th international acm sigaccess conference on computers and accessibility* (pp. 377–378).

Muller, M. J. (2003). Participatory design: the third space in hci. *Human-computer interaction: Development process*, *4235*(2003), 165–185.

Mulloy, A. M., Gevarter, C., Hopkins, M., Sutherland, K. S., & Ramdoss, S. T. (2014). Assistive technology for students with visual impairments and blindness. In *Assistive technologies for people with diverse abilities* (pp. 113–156). Springer.

Mynatt, E. (1997). Transforming Graphical Interfaces Into Auditory Interfaces for Blind Users. *Human-Computer Interaction*, *12*(1), 7–45. doi: 10.1207/s15327051hci1201&2_2

Mynatt, E. D., & Weber, G. (1994). Nonvisual Presentation of Graphical User Interfaces: Contrasting Two Approaches. In (pp. 166–172). New York, NY, USA: ACM. doi: 10.1145/191666.191732

Nardi, B. A. (1996a). *Context and consciousness: activity theory and human-computer interaction.* mit Press.

Nardi, B. A. (1996b). Studying context: A comparison of activity theory, situated action models, and distributed cognition. *Context and consciousness: Activity theory and human-computer interaction*, *69102*.

Nguyen, D. H., Marcu, G., Hayes, G. R., Truong, K. N., Scott, J., Langheinrich, M., & Roduner, C. (2009). Encountering sensecam: personal recording technologies in everyday life. In *Proceedings of the 11th international conference on ubiquitous computing* (pp. 165–174).

Nielsen, J. (1994). Heuristic evaluation. *Usability Inspection Methods*, *17*(1), 25–62.

Niemeyer, W., & Starlinger, I. (1981). Do the blind hear better? investigations on auditory processing in congenital or early acquired blindness ii. central functions. *Audiology*, *20*(6), 510–515.

Ochs, E., Kremer-Sadlik, T., Solomon, O., & Sirota, K. G. (2001). Inclusion as social practice: Views of children with autism. *Social Development*, *10*(3), 399–419.

of 1990, A. W. D. A. (1990). *Pub. l. no. 101-336, 104 stat. 328.*

Papadopoulos, K., Papakonstantinou, D., Koutsoklenis, A., Koustriava, E., & Kouderi, V. (2015). Social support, social networks, and happiness of individuals with visual impairments. *Rehabilitation Counseling Bulletin*, *58*(4), 240–249.

Peres, S. C., Best, V., Brock, D., Frauenberger, C., Hermann, T., Neuhoff, J. G., … Stockman, A. (2008). Auditory interfaces. *HCI beyond the GUI: design for haptic, speech, olfactory, and other nontraditional interfaces*, 147–195.

Plessers, P., Casteleyn, S., Yesilada, Y., Troyer, O. D., Stevens, R., Harper, S., & Goble, C. (2005). Accessibility : A Web Engineering Approach. *Human Factors*, *05*, 353–362. Retrieved from `http://portal.acm.org/citation.cfm?doid=1060745.1060799` doi: 10.1145/1060745.1060799

Poll, L., & Waterham, R. (1995, mar). Graphical user interfaces and visually disabled users. *IEEE Transactions on Rehabilitation Engineering*, *3*(1), 65–69. Retrieved from `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=372894` doi: 10.1109/86.372894

Potluri, V., Vaithilingam, P., Iyengar, S., Vidya, Y., Swaminathan, M., & Srinivasa, G. (2018). Codetalk: Improving programming environment accessibility for visually impaired developers. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 618:1–618:11). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3173574.3174192` doi: 10.1145/3173574.3174192

Prescher, D., Weber, G., & Spindler, M. (2010). A tactile windowing system for blind users. *Assets'10*(August 2015), 91–98. doi: 10.1145/1878803.1878821

Profita, H., Albaghli, R., Findlater, L., Jaeger, P., & Kane, S. K. (2016). The at effect: how disability affects the perceived social acceptability of head-mounted display use. In *proceedings of the 2016 chi conference on human factors in computing systems* (pp. 4884–4895).

Profita, H. P., Stangl, A., Matuszewska, L., Sky, S., Kushalnagar, R., & Kane, S. K. (2018, September). "wear it loud": How and why hearing aid and cochlear implant users customize their devices. *ACM Trans. Access. Comput.*, *11*(3), 13:1–13:32. Retrieved from `http://doi.acm.org/10.1145/3214382` doi: 10.1145/3214382

Raman, T. V. (1996). EmacspeakâĂŤa speech interface. In (pp. 66–71). ACM. Retrieved from `http://dl.acm.org/citation.cfm?id=238405`

Ratanasit, D., & Moore, M. M. (2005). Review of Approaches. , *99*(2), 1–38.

Rattenbury, T., & Canny, J. (2007). CAAD. In *Proceedings of the sigchi conference on human factors in computing systems - chi '07* (p. 687). New York, New York, USA: ACM Press. Retrieved from `http://portal.acm.org/citation.cfm?doid=1240624.1240731` doi: 10.1145/1240624.1240731

Rector, K., Bennett, C. L., & Kientz, J. A. (2013). Eyes-free yoga: An exergame using depth cameras for blind & low vision exercise. In *Proceedings of the 15th international acm sigaccess conference on computers and accessibility* (pp. 12:1–12:8). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2513383.2513392` doi: 10.1145/2513383.2513392

Rector, K., Milne, L., Ladner, R. E., Friedman, B., & Kientz, J. A. (2015). Exploring the opportunities and challenges with exercise technologies for people who are blind or low-vision. In *Proceedings of the 17th international acm sigaccess conference on computers & accessibility* (pp. 203–214). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2700648.2809846` doi: 10.1145/2700648.2809846

Richards, J. T., Montague, K., & Hanson, V. L. (2012). Web accessibility as a side effect. In *Proceedings of the 14th international acm sigaccess conference on computers and accessibility* (pp. 79–86).

Ripat, J., & Woodgate, R. (2011). The intersection of culture, disability and assistive technology. *Disability and Rehabilitation: Assistive Technology*, *6*(2), 87-96. Retrieved from `https://doi.org/10.3109/17483107.2010.507859` (PMID: 20698763) doi: 10.3109/17483107.2010.507859

Röder, B., Rösler, F., & Neville, H. J. (2001). Auditory memory in congenitally blind adults: A behavioral-electrophysiological investigation. *Cognitive Brain Research*, *11*(2), 289–303. doi: 10.1016/S0926-6410(01)00002-7

Röder, B., Teder-SaÈlejaÈrvi, W., Sterr, A., Rösler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, *400*(6740), 162–166.

Rodrigues, A., Montague, K., Nicolau, H., & Guerreiro, T. (2015). Getting smartphones to talkback: Understanding the smartphone adoption process of blind users. In *Proceedings of the 17th international acm sigaccess conference on computers & accessibility* (pp. 23–32).

Saha, M., Fiannaca, A. J., Kneisel, M., Cutrell, E., & Morris, M. R. (2019). Closing the gap: Designing for the last-few-meters wayfinding problem for people with visual impairments. In *The 21st international acm sigaccess conference on computers and accessibility* (pp. 222–235). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3308561.3353776` doi: 10.1145/3308561.3353776

Sanders, E. B.-N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign*, *4*(1), 5-18. Retrieved from `https://doi.org/10.1080/15710880701875068` doi: 10.1080/15710880701875068

Sapp, W., & Hatlen, P. (2010). The expanded core curriculum: Where we have been, where we are going, and how we can get there. *Journal of Visual Impairment & Blindness*.

Savage, V., Zhang, X., & Hartmann, B. (2012). Midas: fabricating custom capacitive touch sensors to prototype interactive objects. In (pp. 579–588). ACM. Retrieved from `http://dl.acm.org/citation.cfm?id=2380189http://bid.berkeley.edu/files/papers/savage-midas-uist2012.pdf`

Savidis, A., Stephanidis, C., Korte, A., Crispien, K., & Fellbaum, K. (1996). A generic direct-manipulation 3d-auditory environment for hierarchical navigation in non-visual interaction. In *Proceedings of the second annual acm conference on assistive technologies* (pp. 117–123). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/228347.228366` doi: 10.1145/228347.228366

Schmandt, C. (1998). Audio hallway: a virtual acoustic environment for browsing. In *Proceedings of the 11th annual acm symposium on user interface software and technology* (pp. 163–170).

Schneider, B., Wallace, J. M., Blikstein, P., & Pea, R. (2013). Preparing for future learning with a tangible user interface: the case of neuroscience. *Learning Technologies, IEEE Transactions on*, *6*(2), 117–129. Retrieved from `http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=6477032http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true{&}arnumber=6477032`

Scior, K. (2011). Public awareness, attitudes and beliefs regarding intellectual disability: A systematic review. *Research in developmental disabilities*, *32*(6), 2164–2182.

Shaer, O., & Hornecker, E. (2010). Tangible user interfaces: past, present, and future directions. *Foundations and Trends in Human-Computer Interaction*, *3*(1âĂŞ2), 1–137. Retrieved from `http://dl.acm.org/citation.cfm?id=1755234` doi: 10.1561/1100000026

Shinohara, K., Bennett, C. L., Pratt, W., & Wobbrock, J. O. (2018, March). Tenets for social accessibility: Towards humanizing disabled people in design. *ACM Trans. Access. Comput.*, *11*(1), 6:1–6:31. Retrieved from `http://doi.acm.org/10.1145/3178855` doi: 10.1145/3178855

Shinohara, K., & Wobbrock, J. O. (2011). In the shadow of misperception: Assistive technology use and social interactions. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 705–714). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/1978942.1979044` doi: 10.1145/1978942.1979044

Sodhi, R., Poupyrev, Glisson, M., & Israr, a. (2013). AIREAL: interactive tactile experiences in free air. *ACM Transactions on Graphics*, *32*(4), 134. doi: 10.1145/2461912.2462007

Tan, B. S., Wilson, E., Campain, R., Murfitt, K., & Hagiliassis, N. (2019). Understanding Negative Attitudes Toward Disability to Foster Social Inclusion: An Australian Case Study. In S. Halder & V. Argyropoulos (Eds.), *Inclusion, Equity and Access for Individuals with Disabilities: Insights from Educators across World* (pp. 41–65). Singapore: Springer Singapore. Retrieved 2019-06-03, from `https://doi.org/10.1007/978-981-13-5962-0_3` doi: 10.1007/978-981-13-5962-0_3

Taylor, J. R. (2009). The communicative construction of community: Authority and organizing. *Learning and expanding with activity theory*, 228–239.

Teal, G., & French, T. (2016). Designed engagement. In *Proceedings of the design research society 50th anniversery conference.*

Vázquez, M., & Steinfeld, A. (2012). Helping visually impaired users properly aim a camera. In *Proceedings of the 14th international acm sigaccess conference on computers and accessibility* (pp. 95–102).

Vigo, M., & Harper, S. (2013a). Challenging information foraging theory: screen reader users are not always driven by information scent. In *Proceedings of the 24th acm conference on hypertext and social media* (pp. 60–68).

Vigo, M., & Harper, S. (2013b). Coping tactics employed by visually disabled users on the web. *International Journal of Human-Computer Studies*.

Vigo, M., & Harper, S. (2014). A snapshot of the first encounters of visually disabled users with the Web. *Computers in Human Behavior*, *34*, 203–212. doi: 10.1016/j.chb.2014.01.045

Vitense, H. S., Jacko, J. A., & Emery, V. K. (2002). Multimodal feedback: establishing a performance baseline for improved access by individuals with visual impairments. In *Proceedings of the fifth international acm conference on assistive technologies* (pp. 49–56).

Voida, S., & Mynatt, E. (2009). It feels better than filing: everyday work experiences in an activity-based computing system. *Proceedings of the SIGCHI Conference on Human*. Retrieved from `http://dl.acm.org/citation.cfm?id=1518744`

Voida, S., Mynatt, E., & MacIntyre, B. (2007). Supporting activity in desktop and ubiquitous computing. *Beyond the Desktop*.

Voida, S., Mynatt, E. D., & Edwards, W. K. (2008). Re-framing the desktop interface around the activities of knowledge work. In *Proceedings of the 21st annual acm symposium on user interface software and technology - uist '08* (p. 211). New York, New York, USA: ACM Press. Retrieved from `http://portal.acm.org/citation.cfm?doid=1449715.1449751` doi: 10.1145/1449715.1449751

Völkel, T., Weber, G., & Baumann, U. (2008). Tactile graphics revised: the novel brailledis 9000 pin-matrix device with multitouch input. In *International conference on computers for handicapped persons* (pp. 835–842).

Vredenburg, K., Mao, J.-Y., Smith, P. W., & Carey, T. (2002). A survey of user-centered design practice. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 471–478).

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

W3C. (2020). *Web Content Accessibility Guidelines (WCAG).* Retrieved 2020-08-01, from `https://www.w3.org/WAI/standards-guidelines/wcag/`

Walker, B. N., Nance, A., & Lindsay, J. (2006). Spearcons: speech-based earcons improve navigation performance in auditory menus. *Proceedings of the International Conference on Auditory Display*, 95–98.

Wang, R., Yu, C., Yang, X.-D., He, W., & Shi, Y. (2019). Eartouch: facilitating smartphone use for visually impaired people in mobile and public scenarios. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1–13).

Webaim.org. (2020). *Web Accessibility in Mind.* Retrieved 2020-08-01, from `https://webaim.org/`

Weiser, M. (1991). The computer for the 21st century. *Scientific american*, *265*(3), 94–104.

Weiss, M., Voelker, S., Borchers, J., & Wacharamanotham, C. (2011). FingerFlux: Near-surface Haptic Feedback on Tabletops. *Proceedings of the 24th Symposium on User Interface Software and Technology - UIST '11*, 615–620. doi: 10.1145/2047196.2047277

Weiss, M., Wagner, J., Jansen, Y., Jennings, R., Khoshabeh, R., Hollan, J. D., & Borchers, J. (2009). SLAP Widgets: Bridging the Gap Between Virtual and Physical Controls on Tabletops. In (pp. 481–490). New York, NY, USA: ACM. doi: 10.1145/1518701.1518779

Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge university press.

Wentz, B., Hochheiser, H., & Lazar, J. (2013, aug). A survey of blind users on the usability of email applications. *Universal Access in the Information Society*, *12*(3), 327–336. doi: 10.1007/s10209-012-0285-9

Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S., & Froehlich, J. (2011). Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing (TACCESS)*, *3*(3), 9.

Wong, E. J., Yap, K. M., Alexander, J., & Karnik, A. (2015). Habos: Towards a platform of haptic-audio based online shopping for the visually impaired. In *2015 ieee conference on open systems (icos)* (pp. 62–67).

Xu, C., Israr, A., Poupyrev, I., Bau, O., & Harrison, C. (2011). Tactile display for the visually impaired using TeslaTouch. In *Proceedings of the 2011 annual conference extended abstracts on human factors in computing systems - chi ea '11* (p. 317). New York, New York, USA: ACM Press. Retrieved from `http://portal.acm.org/citation.cfm?doid=1979742.1979705` doi: 10.1145/1979742.1979705

Yesilada, Y., Harper, S., Goble, C., & Stevens, R. (2004). Screen readers cannot see. In (pp. 445–458). Springer.

Zhou, L., Parker, A. T., Smith, D. W., & Griffin-Shirley, N. (2011). Assistive technology for students with visual impairments: Challenges and needs in teachers' preparation programs and practice. *Journal of Visual Impairment & Blindness*, *105*(4), 197.

Zuckerman, O., & Gal-Oz, A. (2013). To TUI or not to TUI: Evaluating performance and preference in tangible vs. graphical user interfaces. *International Journal of Human Computer Studies*, *71*(7-8), 803–820. doi: 10.1016/j.ijhcs.2013.04.003

Zychlinski, E., Ben-Ezra, M., & Raz, Y. H. (2016, November). Changing attitudes about disability: The impact of the 'Accessible Community' program. *Journal of Social*

*Work*, *16*(6), 742–757. Retrieved 2019-05-31, from http://journals.sagepub.com/doi/10.1177/1468017315589871 doi: 10.1177/1468017315589871

# Appendices

## A Verbal Steering Command Transcripts

### A.1 CoachS1 and PaddlerB2

[C1]: okay check, we're gonna do it by voice command b/c the coop system just cooped out.

[P1]: okay

[C1]: okay so your going to push on your left foot first.

[C1]: ok go ahead, start paddling, push hard on that left.

[P1]: I am

[C1]: perfect, now you can straighten out.

[P1]: huh?

[C1]: take your foot off the left turn press alittle bit to the right. now keep on even cuz your gonna straigh ahead. little back to your left. good. come al ittle ore left good. little bit right. now you going straigh ahead on your own

...boats collide...

[C1]: okay you want to press to your right side and paddle, hard right, paddle striaght, keep pushing to yor right, hard right hard right keep going to your right. stay ont he right stay ont hr right, stay on the right, stay on the right, stay on the right, okay even them out a little bit straight ahead little bit to your left, k, litte right, good, k, little bit left, little right, gonna be that way the whoel time, little left, i think b/c the ama, k little right, just a hair to the left, press right, press right again, ok. now little bit left. little bit right. little left. k. little right.

@9:57 ....25 seconds...

[C1]: little push right. 4 second. left. right. left. right. ok try not to tap to hard so just littles bits at a time. k left, left. k, right. left. little right. small tap left...small tap right...small tap right...small tap left....small tap right....small tap left.small tap right...small tap left...small tap right, right, hold it right hold it right, good, keep going. small tap left, small tap left, small tap left...small tap right...small tap left...this currents a pain in the ass too....small tap right, small left, small left............small left...small left..........small tap right...k small left..i think that's the other thing, we need right up the dialoge and go over it with them so they no small tap, hold, so I guess this is good, ok small tap right.......small tap left.......small tap right......small tap left..there you go..left a bit.small tap right..small tap left...left, left...ok you are about 5 feet from shore, slow down...stop stop stop stop stop stop stop STOP STOP STOP.

## A.2 PaddlerB1 and CoachS3

[P1] tap right

[P1] keep working it

[P1] tap left, that's it

[P1] tap right, there you go, hold that line

[P1] [louder] tap right, and hold that line

[P1] doing great bud, keep working it. tap left, and hold it.

[P1] tap right, that's it

[P1] tap left, that's it

[P1] tap right, find center,

[P1] there you go, keep working it.

[P1] tap right, that's it...find center...tap left Andrew, find center. Alright.

[P1] tap left

[P1] tap right...and tap left, you're kind of slaloming back and forth, tap right. right there.

[P1] tap left, there you go, peftect, keep that line

[P1] alirght, tap right, alright right there, hold that line.

[P1] tap right

[P1] tap left, there you go

[P1] [louder] tap right, there you go, right there, hold that line

[P1] [louder] tap left...there you go...keep working it, getting near the bridge

[P1] tap right, and hold it there, tap left, remember to find the center after you tap.

[P2] ok.

[P1] tap left, yep

[P1] work on that line, tap left, perfect.

[P1] tap right, that's it.

[P1] tap left, there you go.

[P1] tap right, TAP RIGHT.

[P1] tap left, that was a little hot. there you go, find that center and hold that

[P1] tap left, there you go.

[P1] doing great buddy, coming up to the bridge, and then it will just be the home stretch. after the bridge we'll let you take a break and then we'll head home okay?

[P2] yep.

[P1] right on.

[P1] tap right. okay, find center, there you go. uh, tap left, find center...good.

[P1] tap right, there's your center. tap left. [actually taps right again] woah, where you going, there you go.

[P1] tap left. Find your center. There you go, right there.

[P1] tap left, that's it.

[P1] tap right

[P1] alright tap left, right there, hold that line, keep it center.

[P1] tap right, there you go, hold that line.

# B Qualtitative Analysis Report: Themes and Codes

## B.1 safety and awareness

### Uncategorized

[1.76.5] learning to trust novel [7.19.1] risks of shared activities [10.82.7] avoiding risks [11.39.2] risks of not properly communicating context and situation [20.28.2] learning brings risks [25.42.2] situation awareness of controller [26.77.5] water conditions can create unsafe situations [39.73.4] limiting responsibility to increase focus on shared control [41.77.5] situation awareness [43.77.5] take control when necessary [49.76.5] building trust with steerperson [50.23.1] conflicting goals adds stress [53.43.2] situation awareness between non-controlled paddlers and steerperson [55.78.5] safety [56.9.1] safety through shared knowledge and experience [71.77.5] concern over personal safety [73.37.2] anxiety and stress vary across individuals [79.43.2] communicating situational changes [83.78.5] naturally occuring obstacles affect safety [84.78.5] natural obstacles [85.45.2] steerperson and boat coordination [86.75.5] removing responsibility reduces stress [90.43.2] situation awareness between all paddlers and other watercraft [91.46.2] loss of situational control increases stress [94.42.2] need for coordination between all parties [101.79.5] mitigating challenges to improve safety increases dependency on others [107.73.4] uncertainty about steering more than one person [108.79.5] artificial obstacles increase complexity [110.75.5] paddling can be stressful [119.82.7] staying safe. [126.77.5] situation awareness [129.45.2] maintaining situation awareness across multiple roles [139.76.5] takes time to build trust in steerperson [140.37.2] stressors affect individuals by familiarity and experience [144.78.5] compensating for lack of confidence in unsafe conditions [147.77.5] awareness increases safety [148.40.2] communication effects on stress [154.77.5] hidden obstacles affect safety [155.43.2] fluid situtation leads to stressfulness [162.46.2] situation awareness [171.33.2] orientation and confidence varies by paddler [173.78.5] situation awareness [174.38.2] reduce situational complextity through pre-planning [178.13.1] safety concerns in shared control [179.13.1] challenges of sharing control [184.79.5] sacrificing desired behavior for safety and reduce complexity [187.75.4] pushing paddlers keeps them safe [188.73.4] reasoning coaches require focus [195.77.5] trust is more important than safety [207.81.6] recognizing risks [220.28.2] safety [224.46.2] safety measures often require additional support [240.79.5] shared control creates challenges [245.77.5] water conditions change risks and safety [250.23.1] complex device interaction increases error rate [251.46.2] changing conditions affect safety

## B.2 building common language from shared analogies

### Uncategorized

[2.69.4] poor spatial awareness affects paddling performance [3.70.4] cognitive ability affects technique and coaching strategy [4.28.2] language is built through time and repetition [8.69.4] finding suitable metaphors [29.34.2] metaphors drawn form common knowledge [57.35.2] use of visual language is unclear [69.37.2] shared experiences [80.36.2] applying metaphors from shared experiences [117.34.2] using familiar metaphors drawn from shared experiences [118.70.4] procedures that build muscle memory [120.36.2] ambiguous coaching language [131.35.2] ambiguous coaching language [153.84.8] adhoc metaphor use [206.70.4] practice and repetition [209.69.4] need to identify useful metaphors [211.69.4] metaphor success varies across paddlers [215.83.7] communication breakdown from undefined metaphors [222.70.4] muscle memory through repetition in place of visual mirroring [225.34.2] finding the right metaphors [229.70.4] metaphor success draws on personal experience

## B.3 challenges of community coordination

### Uncategorized

[5.15.1] communication challenges [6.10.1] patience [21.28.2] time constraint and opportunity limits engagement [27.72.4] weather constrains coaching and practice strategy [197.10.1] coordination

## B.4 accommodating different sensory abilities

### causes of audibility issues

[9.42.2] complex verbal statements [13.38.2] wind makes verbal commands inaudible [96.42.2] boat positioning affects audibility [116.42.2] variations in audibility [156.37.2] natural conditions affect audibility [175.42.2] ambient sounds affect audibility [177.25.1] proximity for verbal commands

### sensory feedback for safety

[38.35.2] visual versus tangible outcome variation [45.11.1] appropriating new sounds [112.77.5] artificial [221.78.5] natural obstacles [230.42.2] paddler limitations to accommodate inaudibility

### resolving audibility issues

[30.38.2] simplyfying verbal commands for audibility [46.43.2] simple verbal communication ease audibility [89.42.2] acknowledging impediments to verbal commands [103.44.2] simplifying commands for audibility [105.34.2] accommodating variation in audibility [109.42.2] two way radio improves audibility [142.42.2] raising voice to workaround [158.42.2] proximity improves audibility [182.35.2] simplifying verbal instruction to improve audibility [227.38.2] simplifying verbal commands for audibility [249.42.2] acknowleding audibility issues by simplifying verbal statements

### problems with verbal and auditory dependence

[16.23.1] unclear communication [87.66.3] issuing verbal commands gets tiring [97.46.2] acknowledging audio overload [143.39.2] increased risk due to inaudibility of verbal commands [172.46.2] auditory overload for paddler [185.42.2] inconsistent audibility. [204.75.5] absence of verbal commands improves focus [208.75.5] annoyed by verbal commands [214.42.2] difficulty hearing [236.25.1] technical workarounds to audibility issues

### verbal communication is necessary

[48.37.2] providing context to environmental sounds [115.44.2] communicating situational changes [124.77.5] desire for more information and situation awareness [196.66.3] verbal cadence replaces visual synchronization cues [226.36.2] verbally communicating auditory context [235.77.5] desire for context

### sonification provides contextual clues

[12.11.1] reappropriating the absence of visual cues with sounds [146.67.3] tradeoffs between audible artifacts

### supporting range of sensory perception

[42.32.2] paddlers use tactile cues to orient in the canoe [152.69.4] reliance on proprioception [198.46.2] visual cues remain important for low vision paddlers [217.77.5] tangible communication

### Uncategorized

[95.43.2] context

## B.5 navigating expectations of mixed-ability communities

### Uncategorized

[14.8.1] thankfulness [34.12.1] excitement over partnership [59.5.1] underestimating community enthusiasm [61.8.1] gratitude for support [66.10.1] understanding [70.29.2] accommodating disability through exception [76.4.1] exhibiting feelings of gratitude [77.6.1] relaxed state of coordination and communication [81.10.1] timeliness [93.45.2] unmet expectations [104.45.2] frustration [127.10.1] gratitude [133.79.5] accommodating shortcomings of shared technology [145.73.4] financial limitations [159.12.1] miscommunications [161.40.2] complexity of technology affects norms [166.32.2] tension between different ability groups [181.15.1] gratefullness [183.13.1] gratitude [192.10.1] gratitude [194.30.2] inspiration porn [202.27.1] building awareness [212.12.1] gratitude [219.73.4] constrain paddling opportunity [243.17.1] financial challenges

## B.6 personal growth and learning

### Uncategorized

[15.73.4] solo paddling forces effort [17.73.4] unable to cheat [31.81.6] growth through experience [33.36.2] reinforcing improvements from coaching [36.80.6] generating new opportunities for activity [58.33.2] gaining confidence quickly [62.81.6] learning [63.74.4] more types of activities increase social opportunties [64.74.4] solo paddling builds endurance [106.79.5] shared experience takes time to learn [113.76.5] benefits across multiple skill levels [122.23.1] shared control experience [141.70.4] verbal instruction [151.76.5] new possibilities bring new experiences [200.28.2] creating opportunities for inclusiveness [201.80.5] changed experiences from adding technology [244.74.4] rigorous activity pushes paddlers to grow [246.75.5] artificial barriers stifle self-improvement

## B.7 shared control changes structure of activity

### Uncategorized

[18.81.6] participating in a 'real' activity [23.45.2] complexity of shared control [24.46.2] coach-paddler bonding [35.72.4] time constraints influence coaching strategy [37.66.3] tangible expands ways to coach effectively [47.70.4] physical ability affects paddling technique and coaching strategy [98.75.5] assistive technology to break through barriers [99.28.2] creating new opportunities for activity [210.78.5] normalizing shared experiences [228.79.5] requires familiarity

## B.8 supporting independence and agency

### Uncategorized

[19.81.6] taking pride in new found agency [22.35.2] desire to be independent [28.75.5] personal responsibility for success [40.74.4] freedeom to operate outside of official team activities [44.78.5] desire for self reliance [51.77.5] some dependence is necessary [60.82.7] Feeling more secure in multi-person canoe [67.24.1] unsatisfied in team experience [68.73.4] individual vs team coaching [72.11.1] self-disclosure of disability [75.78.5] maintaining independence and responsibility [82.81.6] desire for independence [88.81.6] desire for independence [92.74.4] solo paddling adds freedom [100.84.8] dependency on multiple people [102.82.7] tradeoffs between independence versus teamwork [111.24.1] desire for personalized instruction [114.82.6] agency [128.72.4] time constraints reduce opportunity for individual training. [136.81.6] freedom [137.81.6] independent paddling brings new experiences [150.28.2] support family needs [160.82.6] and controlling ones own personal goals [168.82.7] independence versus teamwork [170.81.6] taking control [180.45.2] balancing disability and respecting individual agency [191.74.4] solo paddling caters to individual strengths [199.79.5] combining roles to reduce dependency on others [203.71.4] coop perceived to hinder team blending [205.81.6] nervous [213.81.6] exhibiting independence [216.81.6] control of ones schedule [223.40.2] expectations.

[232.11.1] pride in achievements [234.28.2] shared control creates opportuntities and independence [237.82.6] independence [238.81.6] taking pride in new found agency [241.75.5] satisfaction of paddling solo [242.80.6] enthusiasm [248.81.6] avoid burdening others [252.45.2] multiple roles and responsibilities

## B.9 cooperative design with community

### Uncategorized

[32.40.2] supporting technical issues through flexible design [52.6.1] diy approach to problem solving [65.7.1] finding off-the-shelf solutions to problems [74.25.1] technical mistakes on the water are costly [78.13.1] support [123.5.1] shared knowledge and experiences from strangers [125.5.1] managing expectations of the community [130.5.1] willingness of community to help solve a problem together [132.7.1] comfortable with community [134.5.1] unsolicited advice [138.13.1] community engagement [149.13.1] design challenges [157.7.1] letting go of perfection [163.13.1] cooperative designing [165.7.1] tradeoff between speed and strength during iterative prototyping [176.83.8] recovering from error [189.23.1] poor user interaction design [190.13.1] design challenges [193.76.5] unproven assistive technology

## B.10 mixed ability communities

### Uncategorized

[121.30.2] public engagement raises awareness [167.73.4] access to equipment [169.78.5] supporting independence and responsibility encourages engagement [186.27.1] public activity transforms perspectives [231.32.2] free use equates to less important [247.6.1] community has diy mindset

### B.11 coordination and communication on the water

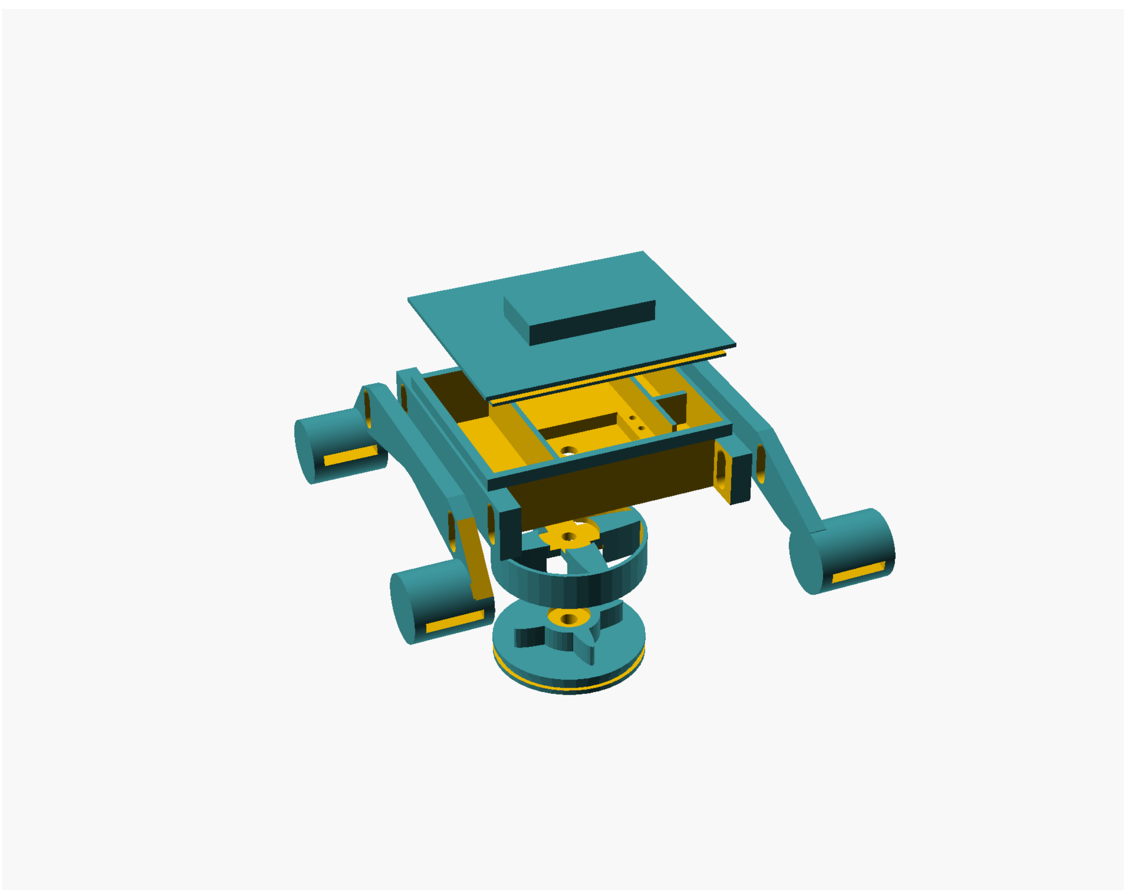
#### New Category

#### Uncategorized

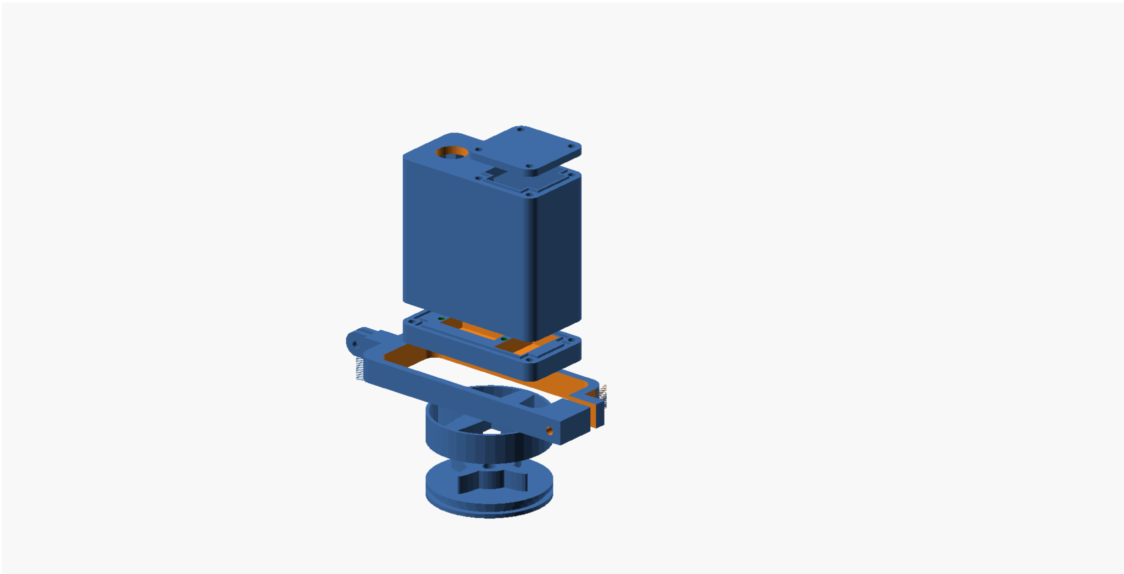
[135.23.1] unclear communication [164.79.5] become familiar [239.23.1] mixed goals

# C CAD Renderings for CoOP

## C.1 CoOP Version 1



## C.2 CoOP Version 2



## C.3 CoOP Version 3