Volume 12
May 2024
ISSN:

2024

The Journal on Technology and Persons with Disabilities

Journal Track Proceedings, Online 2024

CSUN | CENTER ON DISABILITIES

THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

Challenges of mHealth Technology for Blind Older Adults

Soyoung Choi
University of Illinois at Urbana-Champaign
soyoung@illinois.edu

Abstract

In this case study, the author explores unexpected accessibility challenges encountered during mobile health (mHealth) research involving wearable devices and smartphone apps among blind older adults. This study offers recommendations for researchers who may encounter similar challenges in the future. The author delves into the experiences of five older adults, each with different degrees and onset of severe low vision or blindness. This analysis illuminates not just system-related accessibility problems but also a range of issues stemming from user-system and user-researcher interactions. Through this study, the author aims to enhance understanding of how blind older adults engage with mHealth systems and the potential barriers they may encounter. This paper is organized as follows: The cases of four blind older adults who participated in the author's mHealth research over six months will be discussed. Based on the author's assumption that the degree of visual impairment and the attitudes and practices of individuals coexisting with disability are not generalizable, the author's research was conducted using a single-care research design. Each case is introduced, and the lessons learned from the unintentional outcomes are discussed, and recommendations for

addressing them in the future studies are provided at the end of the case analysis.

Keywords

Accessibility, blind, case studies, mobile health apps, older adults, wearables

Introduction

The burgeoning field of accessible computing has seen numerous strides towards the integration of technology that epitomizes inclusivity and diversity (Kaur and Kaur, 19). Specifically, irrespective of physical, sensory, or cognitive impairments, concerted efforts are ongoing to harness the unique potential of individuals with varying abilities, by tailoring products or systems to align with their distinctive characteristics and contexts (Wobbrock et al., 62). This trend is evident in the corpus of recent publications within the realm of accessible computing, which showcase an array of success stories and introduce groundbreaking, accessible systems for users with disabilities. However, the question persists as to whether the technologies improve the quality of life for individuals outside the standard user profile, and whether they truly derive benefits from enhanced technological accessibility. The worry is that accessibility research may inadvertently focus predominantly on satisfying academic intrigue, at the expense of providing practical, tangible benefits to its intended end-users.

This case study is not written with the intention of sharing success stories of accessible computing with other researchers, but rather to share failed cases of mobile health (mHealth) research with individuals with disabilities, assisting other researchers to avoid similar mistakes and discuss the hidden challenges in deploying technologies in their daily lives. This paper is thought to be beneficial for researchers planning intervention studies to enhance health behaviors of older adults with visual impairments through the application of mHealth systems. For the sake of transparency, the author defines the term “visual impairment” as pertaining to individuals who continue to experience severe vision loss or total blindness, even after the application of therapeutic treatments and surgical interventions (Dandoan et al., 1873)

Discussion

Case Studies

The author conducted a six-month mHealth study in Urbana, IL, in partnership with a community rehabilitation center. In this mHealth research, a group of five blind older adults was invited. As existing members of the community center's Low Vision Group, these individuals were already familiar with one another. However, the implementation intervention was individualized for each participant. Following an initial personal meeting to overview the study and confirm their participation, they were given an informed consent form, which was read to them, and verbal consent was secured. The purpose of the study was to motivate the participants to wear a Fitbit device (Fitbit) for four weeks, monitor their health metrics, perform recommended walking exercises for their health condition communicated via text message, and remodel any problematic health behaviors, including physical activity, sleep, and diet. All participants were iPhone users and had some knowledge about wearable technology like Fitbit but had never operated one. Compared to the Apple Watch, the Fitbit is more cost-effective and produces reliable health data. Considering the price and the reliability of the health data (Balbim et al., e25289), the Fitbit Charge 5 was selected for the study.

Lessons learned from Case 1

Mr. D is a 70-year-old Caucasian male who, due to a congenital optic nerve defect, has experienced total blindness since birth. Despite living his entire life without sight, he maintains an optimistic and social outlook. Upon completing high school, he pursued a career at a local rehabilitation center, instructing visually impaired individuals in the use of refreshable braille devices. He married in his mid-twenties, fathered two children, and now lives comfortably with his wife. His leisure activities include walking and swimming, as well as participating in religious events on weekends. Presently, his wife serves as a coordinator at a rehabilitation center, while Mr. D maintains a part-time job teaching braille device usage. He manages his diabetes through diet and exercise. He has been an iPhone user for over five years, primarily using it for basic tasks such as answering calls and sending text messages via Siri, seldom utilizing the iPhone's built-in accessibility features. While conducting research with Mr. D, initial difficulties arose when Mr. D was asked to handle the Fitbit device and wear it on his wrist; he struggled with the task of aligning the elastic rubber band's protruding part with the small hole at the end. His total blindness, coupled with his poor fine motor skills, hindered the execution of this simple task. After more than five minutes of effort, he managed to wear it. The subsequent step involved installing the Fitbit app on Mr. D's iPhone and guiding him to track his health data through it. However, this posed a challenge because the app installation required an Apple ID and password, which Mr. D couldn't remember. Another hurdle presented itself when it came to charging the device. Plugging the USB cable into the USB port on Mr. D's desktop proved to be a challenging and time-consuming task for him. Activities that are usually simple for sighted individuals can become significantly more tedious for blind older adults.

In the initial phase of study design, the author considered acquiring a separate smartphone, fully setting it up, and then handing it over to the participant. This device would come with a pre-set Fitbit account, ensuring a consistent flow of health data crucial for analysis. However, recognizing that most blind individuals are comfortable with their current devices and often daunted by the prospect of adapting to a new one (Herskovitz et al., 2), the author chose to have the participant install the app on their existing smartphone and share their Fitbit account for the intervention. In Mr. D's situation, the author encountered two challenges: 1) the participant possibly not knowing his personal information necessary for app installation, and 2) the design of the Fitbit not being universally accessible for all users. To conserve both time and budget, it is suggested to confirm participants' abilities to install the necessary app on their smartphones ahead of time. The age-related degradation of motor control is especially noticeable in the case of fine motor skills (Hardin, 20). Thus, when selecting consumer health products for feasibility studies, it might be advantageous to conduct preliminary testing with prospective participants to assess each product's user compatibility.

Lessons learned from Case 2

Mr. M, a 66-year-old African American male, serves as an Assistive Technology Specialist at a rehabilitation center. He assists new clients with learning about and borrowing assistive technology devices such as magnifiers and screen magnification software. He suffers from optic nerve degeneration and is gradually losing his vision to the point where he can only discern the shapes of people and objects. After earning a college degree, he worked for 20 years as a computer graphics designer. However, an abrupt vision loss led him to resign from his job, and he secured a new position through a rehabilitation center. He derives immense satisfaction from his current job and is preparing for future vision loss by learning Braille and using a screen reader. He manages his diabetes and high blood pressure and has recently invested in an indoor bicycle to aid in weight loss. In general, he has been using the zoom feature on his iPhone to view the screen and is comfortable utilizing voice commands.

Mr. M's involvement in the research, he exhibited signs of depression. This unstable emotional state further led to a diminished interest in the study, likely heightened by overall fatigue and lethargy. Upon comparing his current situation with past experiences, he attributed the contrast largely to his deteriorating vision. As a result, he became increasingly withdrawn and requested a break from the study. Two weeks after this unforeseen occurrence, Mr. M expressed his enthusiasm to resume the study. This unexpected change in attitude caught the researcher off guard, considering Mr. M's depressive symptoms and the surprising nature of his shift in demeanor. At the beginning of the study, the participant showed a strong commitment to changing his unhealthy behaviors, reflected in his proactive interaction with the author and regular monitoring of his health indicators via the Fitbit web dashboard. However, as this case illustrates, those suffering from progressive visual impairment may struggle with declining mental health and physical instability, significantly affecting their ongoing participation in the study. It is commonly assumed that individuals with visual impairments gradually adjust to their disability (van Munster et al., 2021). However, in cases of ongoing severe vision loss, strong feelings of frustration and loss can persist (Zapata and Pearlstein, 1492). These issues underscore the importance of careful consideration in terms of ethical research conduct and the adequate support that should be extended to such participants.

Lessons learned from Case 3

Mrs. J, the oldest participant in the study at 87, is a retired teacher who lives with her husband. Despite her mobility being significantly hampered, necessitating the use of a walking aid, she is deeply interested in the study and eager to learn more about her health. Mrs. J's technological knowledge is confined to the use of an iPhone and iPad for communications such as calls and texts to her children, and she has no experience using these devices for online information search. Her computer skills, which she acquired during her teaching career, are now outdated, and only relate to the use of floppy disks. Due to severe low vision caused by glaucoma and macular degeneration, she heavily relies on magnification devices to the point where any disruption can significantly impact her daily routine. In her home, a large calendar notes her and her husband's various appointments, as she prefers traditional methods for tracking engagements over digital solutions (Figure 1). While working with Mrs. J, the author faced several challenges. One significant issue was her limited dexterity and coordination in her fingers, which made interacting with touchscreen devices difficult. She also struggled to press the home button long enough for the device to recognize her fingerprint, often leaving her stuck at the lock screen. Mrs. J’s initial difficulties with installing the app highlighted the foundational knowledge required for older adults who are unfamiliar with modern digital devices. These experiences provoked important considerations about the practicality of seniors learning to use current mobile health devices and the potential benefits of gathering personal health data.

Figure 1. Pencil and paper based self-tracking

This case highlights that advanced technology may not always be the best solution for older adults, as traditional methods may at times prove more beneficial. Mobile devices have drastically revolutionized the collection of personal health data, facilitating continuous monitoring and automated data acquisition with minimal user intervention (Epstein et al., 2). However, the necessity for users to examine the gathered data, along with synchronizing devices and apps for data transfer, may pose substantial challenges for the “oldest old” demographic (Slavin et al., 1103). This age group, typically consisting of individuals aged 80 and above, is growing in population at a faster pace than any other age group worldwide. When implementing intervention studies involving technologies within the “oldest old” population, it is imperative to meticulously evaluate their foundational technological skills and the feasibility of handling the proposed device. Such assessments offer essential insights that can inform the structuring of the research design and methodologies, thereby guaranteeing their usability and efficacy specifically tailored to this unique demographic.

Lessons learned from Case 4

Mrs. N, a retired scholar, lives with her single daughter. She grapples with significant low vision, including severe loss of peripheral vision and sight obstructions caused by floaters. She often accidentally bumps into passersby and struggles with narrow entrances, frequently colliding with door frames. Despite the need to adapt to her reduced field of vision, she has not undergone any specialized rehabilitation training to aid in this process. Her mobility has significantly decreased, and she expresses fear about exploring new places alone. Regardless of these difficulties, Mrs. N maintains her ability to browse the internet and compose emails, skills attributed to her past career as a university professor. However, she finds it challenging to read her computer screen effectively and has recently started training at a rehabilitation center to learn how to use a screen reader. Mrs. N utilizes an iPhone, frequently using the magnification feature to enhance visibility on the small screen. She primarily uses voice commands for smartphone operations and has installed Alexa in her home to control various devices, such as lights and the radio, using voice commands. While she has not used wearable technology before, Mrs. N showed confidence in her capacity to learn its use. She successfully created a Fitbit account and adeptly installed the corresponding app independently, without the researcher’s help. Mrs. N understood the purpose of the study and expressed a strong interest in monitoring her daily activity levels, as fears of falling had previously limited her physical activity engagement. However, she felt uneasy about sharing her health data, revealing concerns about her sleep patterns and physical activity data being exposed to the researcher. She asked several questions about privacy issues, including whether she could change her Fitbit account's password in the future. Although she understood and agreed to the researcher's access to her health data for possible future health counseling, Mrs. N found the unrestricted sharing of her personal data with the researcher somewhat disturbing.

Individuals with visual impairments frequently express apprehensions regarding the protection of their personal information, particularly when interacting with those without visual impairments (Akter et al., 3). Given their potential limitations in perceiving their surroundings and understanding their data, they often exhibit a heightened awareness and concern for privacy issues (Ahmed et al., 3525). In the case mentioned above, an individual with a solid grounding in privacy and personal data ownership expressed discomfort about her biometric data being collected and potentially shared without her knowledge. In recent years, the trend of self-tracking for personal health information collection and sharing with family, friends, and healthcare providers has surged (Lupton, 2). More and more healthcare consumers leverage mobile health technologies for collecting both subjective health experiences and objective health metrics to assess their health status (Feng et al., e25171). Nevertheless, the necessary ethical safeguards and systems to regulate the exposure of personal health data are still insufficient (Cherif et al., 3). If researchers plan to share sensitive personal health data with participants with visual impairments as well as sighted participants, they particularly ensure a comprehensive explanation of the study's purpose, and the ethical handling of personal health data is provided. This information should be conveyed in a format that participants can easily access and comprehend. During the process of participant recruitment, it is critical to relay the comprehensive study information in an accessible manner. This can be accomplished by generating various formats of study materials that are compatible with assistive technologies such as screen readers, to enhance their accessibility to a wide range of participants with disabilities.

Conclusions

The participants in this study, although possessing basic literacy skills, lacked proficiency in digital technology usage and learning new features. Regardless of their age, they expressed an eagerness to learn, showing interest in understanding how to use the internet or operate new digital devices via their local community centers. Although a significant number of resources is directed towards creating accessible systems, there seems to be a relative neglect of the foundational digital literacy skills needed to engage with these technologies and systems. Considering that older adults require more time to assimilate new knowledge and need repeated reinforcement to retain and apply it compared to younger populations (Chiu et al., 63), there is a definite need for more training opportunities in digital technology usage. Such training should take into consideration the unique learning process and cognitive abilities of older adults.

The perception of visual impairment is inherently subjective, and it is challenging to fully encapsulate an individual’s lived experience of visual impairment through purely quantitative measures derived from ophthalmic diagnostic techniques (Ivers et al., 42). Consequently, research involving individuals with visual impairments calls for profound exploration and is resistant to broad generalizations. Moreover, incorporating the ongoing adaptation process and embodiment of visual impairment, along with a philosophical understanding of disability identity into accessible computing research is crucial (Mankoff, Hayes, and Kasnitz, 4). Unlike research involving standardized user populations, researchers dealing with individuals with disabilities must grasp the intricacies of the disability itself to facilitate effective human-computer interaction. The omission of detailed information about participants' visual impairments in studies may constrain the possibilities for synthesizing knowledge and conducting a meta-analysis of the findings.

Works Cited

Ahmed, T., et al. "Privacy concerns and behaviors of people with visual impairments." Conference on Human Factors in Computing Systems - Proceedings, 2015, pp. 3523-3532.

Akter, T., et al. "Shared privacy concerns of the visually impaired and sighted bystanders with camera-based assistive technologies." ACM Transactions on Accessible Computing, vol. 15, no. 2, 2022, doi.org/10.1145/3506857.

Balbim, Guilherme M., et al. "Using Fitbit as an mHealth intervention tool to promote physical activity: Potential challenges and solutions." *JMIR mHealth and uHealth*, vol. 9, no. 3, 2021, e25289.4.

Cherif, E., N. Bezaz, and M. Mzoughi. "Do personal health concerns and trust in healthcare providers mitigate privacy concerns? Effects on patients' intention to share personal health data on electronic health records." *Social Science and Medicine*, vol. 283, 2021, doi.org/10.1016/j.socscimed.2021.114146.

Chiu, C.-J., et al. "How to help older adults learn new technology? Results from a multiple case research interviewing the internet technology instructors at the Senior Learning Center." *Computers and Education*, vol. 129, 2019, pp. 61-70.

Dandona, Rakhi et al. "Planning low vision services in India: A population-based perspective." *Ophthalmology*, vol. 109, no. 10, 2002, pp.1871-1878.

Epstein, D. A., Caldeira, C., Figueiredo, M. C., Lu, X., Silva, L. M., Williams, L., ... & Chen, Y. (2020). Mapping and taking stock of the personal informatics literature. Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies, 4(4), 1-38. doi.org/10.1145/3432231

Feng, Shan, et al. “How self-tracking and the quantified self-promote health and well-being: systematic review.” *Journal of Medical Internet Research*, vol. 23, no. 9, 2021, e25171.

“Fitbit official site for activity trackers & more.” Fitbit, 6 July 2023, https://www.fitbit.com/global/us/home.36.

Hardin, Misty. “Assessment of hand function and fine motor coordination in the geriatric population.” *Topics in Geriatric Rehabilitation*, vol. 18, no. 2, 2002, pp. 18-27.

Herskovitz, J., et al. “Hacking, switching, combining: Understanding and supporting DIY assistive technology design by blind people.” Conference on Human Factors in Computing Systems - Proceedings, 2023, doi.org/10.1145/3544548.3581249

Ivers, Rebecca Q., et al. “Visual function tests, eye disease and symptoms of visual disability: A population-based assessment.” *Clinical & Experimental Ophthalmology*, vol. 28, no. 1, 2000, pp. 41-47.

Kaur, Navdeep, and Maninderpal Kaur. “Role of technology for Equality, Diversity and Inclusivity.” *International Journal of Information Technology & Computer Engineering*, vol. 2, no. 1, 2022, pp.19-29.

Lupton, Deborah. “Self-tracking, health and medicine.” *Health Sociology Review*, vol. 26, no. 1, 2017, pp. 1-5.22.

Mankoff, J., G.R. Hayes, and D. Kasnitz. “Disability studies as a source of critical inquiry for the field of assistive technology.” ASSETS’10 - Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility, 2010, pp. 3-10

van Munster, E.P.J., et al. “Barriers and facilitators to recognize and discuss depression and anxiety experienced by adults with vision impairment or blindness: A qualitative study.” *BMC Health Services Research*, vol. 21, no. 1, 2021, pp. 1-10.

Slavin, Melissa J., Henry Brodaty, and Perminder S. Sachdev. “Challenges of diagnosing dementia in the oldest old population.” *Journals of Gerontology: Series A: Biological Sciences and Medical Sciences*, vol. 68, no. 9, 2013, pp. 1103–1111.

Wobbrock, Jacob O., et al. “Ability-based design.” *Communications of the ACM*, vol. 61, no. 6, 2018, pp. 62-71.

Zapata, M.A., and Pearlstein, J.G. “Disability self-worth relates to lower anxiety and depression in people with visual impairment.” *Journal of Clinical Psychology*, vol. 78, no. 7, 2022, pp. 1491–1499.

THE JOURNAL ON
TECHNOLOGY AND
PERSONS WITH
DISABILITIES

A Framework for the Outdoor Wayfinding of Visually Impaired

Uddipan Das[1], Vinod Namboodiri[2]
The College of New Jersey[1], Lehigh University[2]
dasu@tcnj.edu, vin423@lehigh.edu

Abstract

In locations where accurate GPS coverage is not available, reading and following visual signs remains the most common method of navigation and receiving wayfinding information. People who are blind or visually impaired (BVI) face difficult challenges as a result of this. Wayfinding is a crucial task for self-sufficiency. However, visually impaired persons struggle a lot in their wayfinding. A low-cost, easy-to-use, and dependable auxiliary wayfinding system has a tremendous demand that complements existing satellite-based systems in both indoor and outdoor areas. This work emphasizes the need for auxiliary outdoor wayfinding tools for people with vision impairments through both a user study and a quantitative study of GPS accuracy in outdoor premises. These two studies highlighted the shortcomings of GPS-based systems in outdoor wayfinding. As such, these two studies with GPS led to the proposal of a lightweight deep learning-based image localization framework in this work for reliably giving path advancement information for outdoor wayfinding. The framework was shown to be very accurate and efficient during testing, indicating that it could be a helpful tool for future inclusion into outdoor wayfinding systems.

Keywords

Accessibility technologies, blind or visually impaired persons, outdoor wayfinding, deep learning

Introduction

Wayfinding has always been a struggle for blind and visually impaired persons. There are numerous outdoor premises such as sidewalks, and office buildings, in and around university campuses, where there is limited (or sometimes non-existent) efficiency of satellite-based systems such as global positioning systems (GPS). Even in places with decent GPS coverage, the precision is typically insufficient for fine-grained pedestrian navigation. Currently available GPS-enabled navigation systems in mobile phones have a typical localization error of around 5m; although the accuracy could be worse near buildings, trees, and other infrastructures (GPS accuracy). So, a low-cost, simple-to-use, and dependable indoor and outdoor navigation system that complements existing satellite-based systems remains critical for people who are blind or visually impaired (BVI). Understanding the current location of the user and which direction, they should go next from their current location is a basic building component for outdoor navigation. Given GPS inaccuracies, identifying the current path and where it leads is a big challenge for pedestrian navigation. This is particularly difficult at intersections when paths diverge, sometimes leading to different destinations.

Related work on building systems for wayfinding relies on solutions that depend on either new infrastructure in the environment or those that don’t require any additional infrastructure. There have been some recent efforts using BLE beacons for indoor wayfinding for the BVI, such as StaNavi (Kim et al. 2016), GuideBeacon (Cheraghi et al. 2017), ASSIST (Nair et al. 2018), PerCept (Ganz et al. 2018), NavCog (Ahmetovic et al. 2016), NavCog3 (Sato et al. 2017) and in (Yang et al. 2018). However, the biggest challenge with BLE-based localization in GPS-limited areas is the scale of infrastructure deployment needed, making it practically infeasible for outdoor environments. Localization without any infrastructure has relied on image processing, computer vision, or utilizing Wi-Fi access points. The limitations of Wi-Fi-based localization include the need for a high density of access points to achieve acceptable accuracies for navigation, especially infeasible in outdoor environments. Some recent works on image-based localization include (Ye et al. 2016), (Zhang et al. 2016), (Li et al. 2018), (Fusco et al. 2020), (Ishihara et al. 2017), and (Ishihara et al. 2018). However, in all previous studies, there has been minimal emphasis on what aspects of outdoor pedestrian navigation have proved problematic for BVI persons and how image-based localization can assist them by minimizing complexity and computational burden on the system.

This paper describes PathLookup, a deep learning-based computer vision framework that allows users to capture an image in the direction of a path they are taking (or want to take) and obtain information about it. PathLookup predicts a path en route to a destination using a convolutional neural network (CNN) and a pre-built contextual database. The PathLookup framework proposed a lightweight CNN model, as opposed to other computationally expensive heavy-weight CNN networks, for path prediction as well as used contextual information of the navigational space in providing navigational guidance to visually impaired persons.

Discussion

Motivation

To demonstrate the motivation and to formulate the problem, two studies are reported in this section that were conducted on two different test sites.

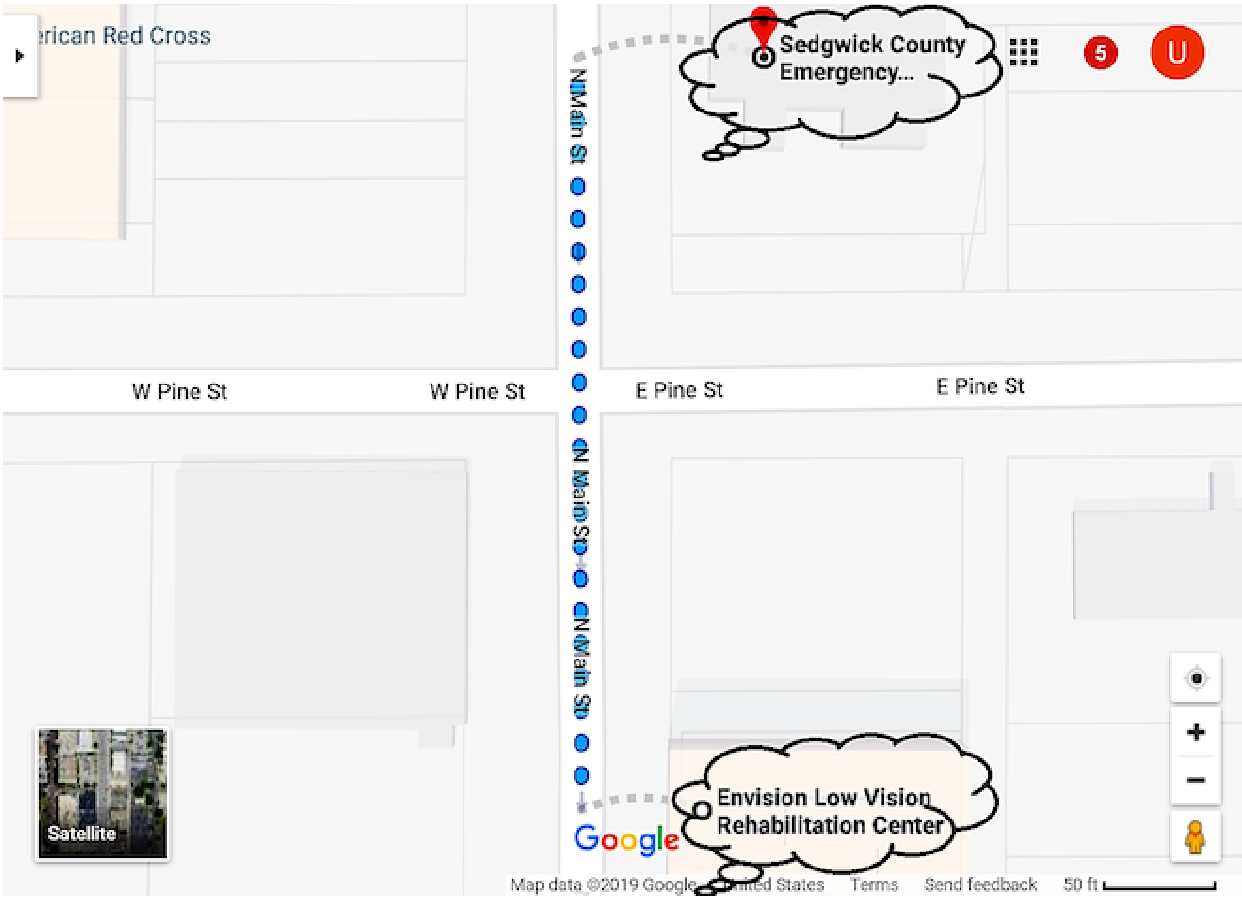


Fig. 1. A motivating Google Maps scenario.

A User Experience study of relying only on GPS

A short user study was carried out to investigate the constraints of real-time outdoor wayfinding and to solicit feedback from potential BVI individuals. Five visually impaired individuals were recruited and asked to navigate from a starting point outside a building to a destination (338 ft away on the shortest route) (Fig.1). The area provided adequate GPS coverage. As can be seen in Table 1, participants took substantially longer to complete a route that should have taken only a minute. Furthermore, the number of steps taken by individuals to arrive at their destination shows a substantial variance. It was a quite straightforward testing path where there were not a lot of turns involved. Moreover, the number of times the traveler needs to ask for assistance was not considered as a metric because it was a study where travelers would completely rely on GPS. However, Google Maps was unable to direct users in the correct way when crossing a street and in locations close to the parking lot. The findings in this study clearly show that employing a GPS-based mapping application alone is insufficient for the task of outdoor wayfinding, even when GPS coverage is acceptable and pedestrian pathways are near roadways with automotive traffic.

Table 1: A user study of relying only on GPS

Participants	**Vision description**	**Time taken (minutes)**	**Steps needed (number)**	**Remarks / Feedback**
P-1	Cane user; Light perception (LP) on both eyes	4.2833	170	Used Google Maps before, feels that Google Maps isn’t fine-grained enough for pedestrian navigation.
P-2	Cane user; LP on right eye, 20/500 on left eye	5.3833	282	Used Google Maps before for pedestrian navigation; thinks that it doesn’t help much.
P-3	Guide dog user; No vision in one eye, LP on the other eye	8.25	404	Used Google Maps before but didn’t help; Tried another app named “BlindSquare” as well for pedestrian navigation but didn’t help either.
P-4	Visually impaired	N/A	N/A	Didn’t feel comfortable using Google Maps; So didn’t participate in testing by using Google Maps.
P-5	Visually impaired	N/A	N/A	Didn’t feel comfortable using Google Maps; So didn’t participate in testing by using Google Maps.

The Live View feature in Google Maps could provide an improved outdoor localization performance (Live View). However, this feature depends on the device’s compatibility with ARKit/ARCore and good Street View coverage of the walking area. There are only a small number of big cities where this feature is available. So, the issues of device compatibility and good street view coverage further restrict the usage of this feature even if it is available. Apple also provides a similar feature using Geotracking which is only available in specific geographic locations (ARGeoTracking). The Geotracking captures views from public streets and routes accessible by car, however, it doesn't include images of gated or pedestrian-only images. As such, this feature at the current stage might not be that useful for visually impaired persons.

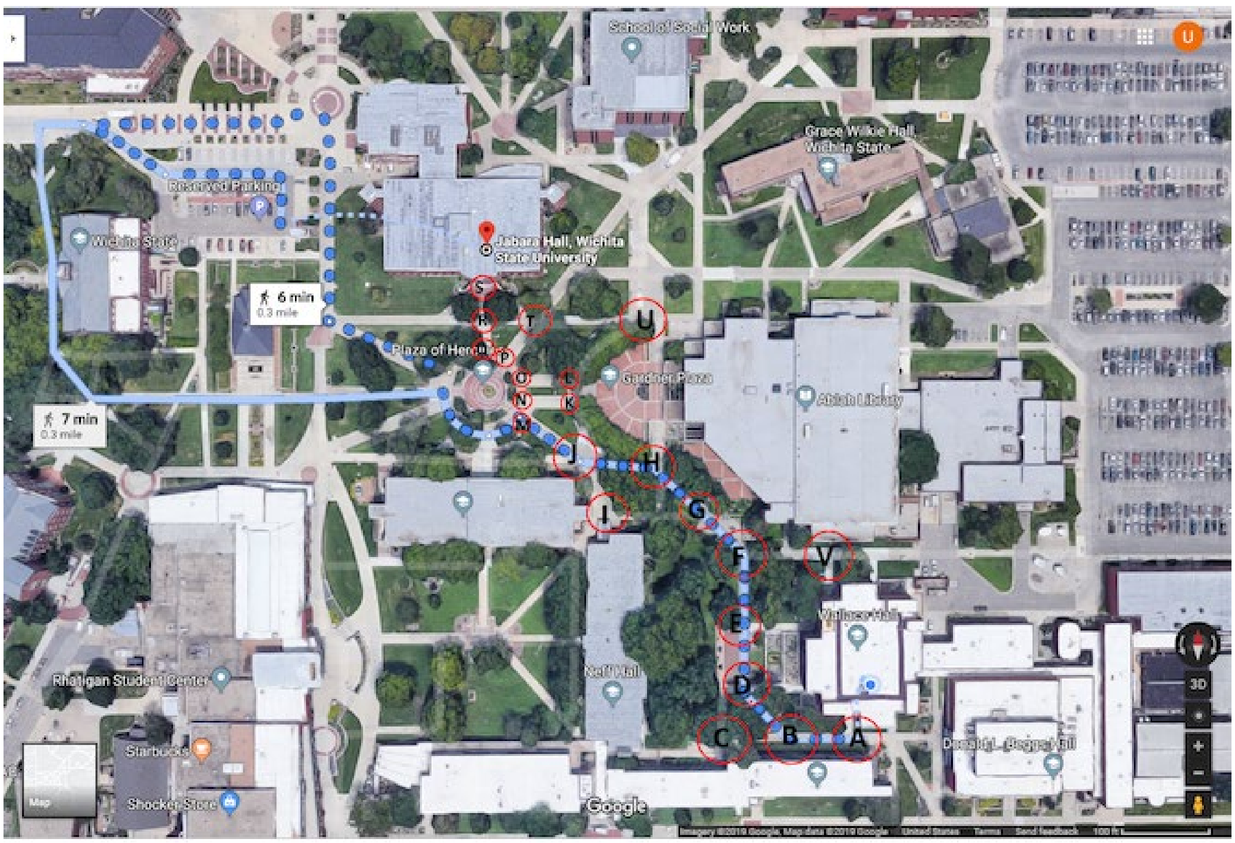


Fig. 2. Pedestrian navigation of a test scenario (PoI marked as red circle).

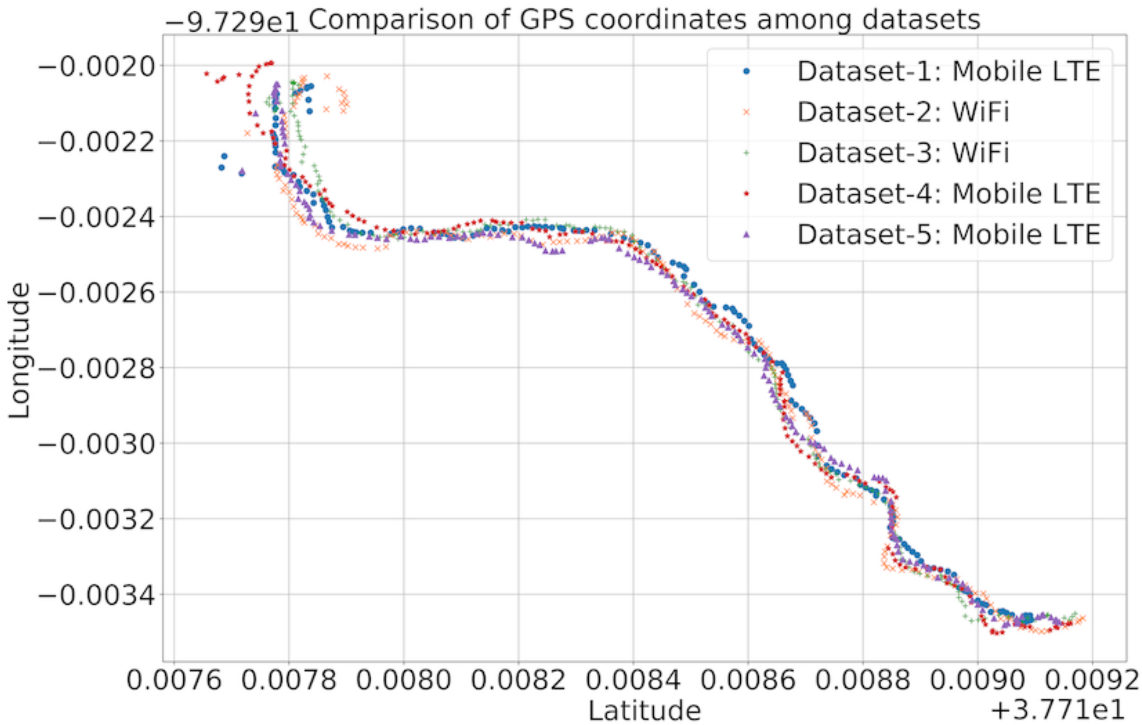


Fig. 3. Variation of GPS coordinates in a test site.

A Quantitative Study of GPS Accuracy

A quantitative study of GPS accuracies for pedestrian navigation was done on another test location after it was determined that Google Maps (which uses underlying GPS signals) is insufficient for outdoor navigation. The pedestrian navigation path in the test scenario depicted in Fig. 2, acquired from Google Maps, begins from Wallace Hall (source) to Jabara Hall (destination) at Wichita State University, Wichita, KS. Many neighboring buildings, trees, intersections, and circles on the path, among other things, make up this test location. As can be shown in Fig. 2, Google Maps is unable to provide a path that is close to the shortest path from source to destination; thus, even a sighted person may be puzzled by Google Maps if this location is unfamiliar to them.

Five distinct sets of data were collected, utilizing a smartphone, on a clear sky day to assess GPS coordinate accuracies received on this route (Wallace Hall to Jabara Hall). In this study, GPS coordinates in terms of longitude and latitude, as well as GPS accuracy, were measured. The variance of GPS coordinates along this path is depicted in Fig. 3. A user may be confused about reaching the destination properly and within the expected time due to the variety of the GPS coordinates provided by the mapping tool in the same path as shown by the plot. It was observed that the acquired GPS coordinates (Fig. 3) had varying degrees of accuracy for each navigation iteration (Fig. 4). The GPS accuracy varied from 3.2m to 18m on the same path. So, any navigation application built using the Google Maps API and relying solely on GPS coordinates would be ineffective for a BVI person in outdoor wayfinding.

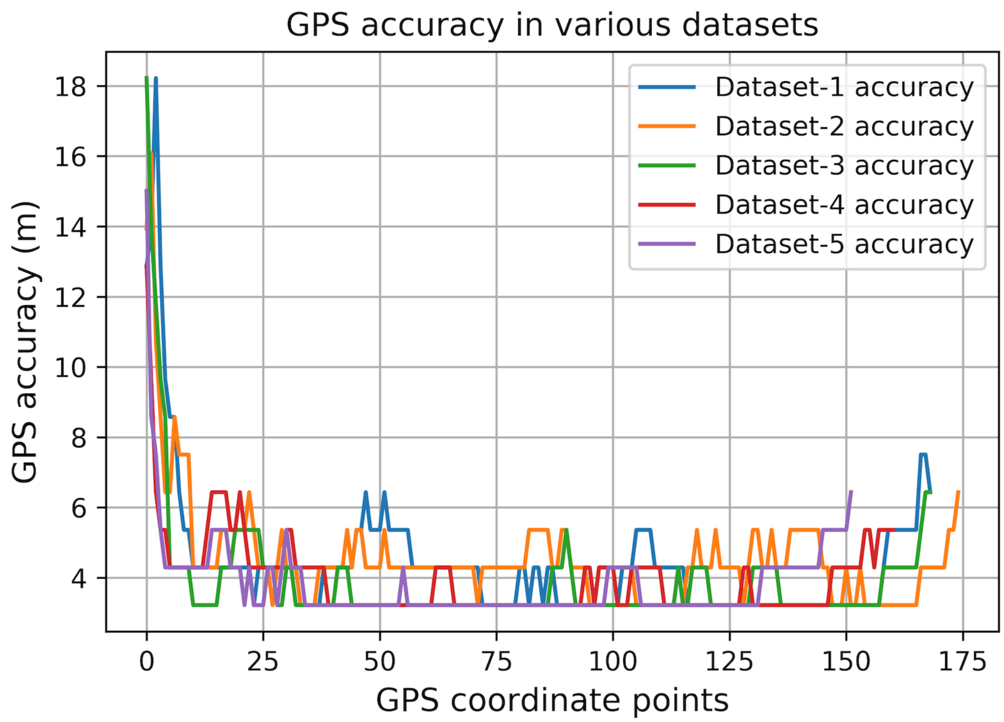


Fig. 4. Variation of GPS accuracies in a test site.

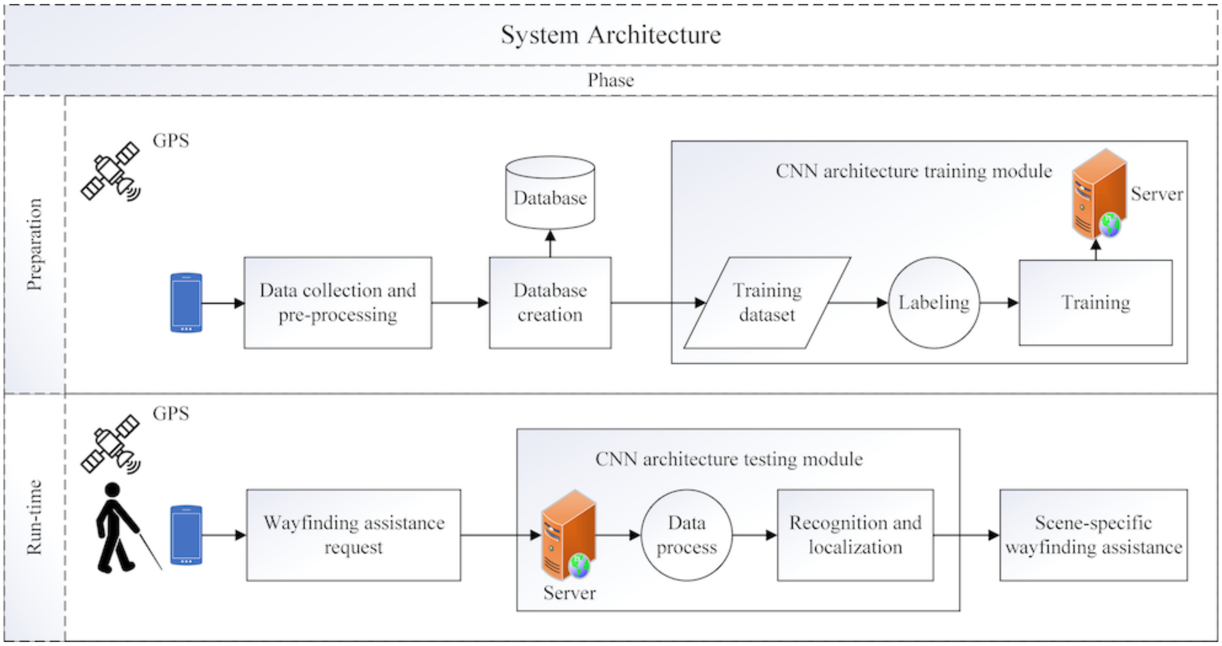


Fig. 5. System architecture of the PathLookup.

The two studies discussed above highlighted the necessity for an auxiliary wayfinding system to assist people with vision impairments in outdoor settings, in addition to GPS-based mapping apps. As such, new methods are required to overcome the limitations of GPS-based outdoor wayfinding for the visually impaired.

System Design of the Framework

Since computer vision techniques can accomplish localization without the usage of tags in the environment, a CNN-based framework named PathLookup (Fig. 5) and its associated algorithm (Fig. 6) are proposed in this section. The proposed PathLookup solution involves training a deep-CNN-based framework in the server, with end-users capturing an image from their smartphone and offloading it to the server for lookup while navigating, and then receiving the results on the smartphone over the network (Fig. 5).

Algorithm 1: PathLookup algorithm (PLA-Sys)

Input : Graph $G(V, E)$ where V is a candidate set of buildings/structures and E is a set of connecting edges of V; image dataset X of paths collected within the graph G; ground-truth GPS coordinates

Output: Path advancement information θ_i along the path of source and destination within the graph G

$\alpha \leftarrow \{b_1, b_2, b_3, ..., b_n\}$, where α is a candidate set of buildings/structures, b_i, consists of potential source and destination

$\beta \leftarrow \{s, d\}$, where β denotes a set of source and destination pair, s is source, and d is destination

$\gamma \leftarrow \{A, B, C, ..., X, Y, Z\}$, where γ denotes a set of possible points of interest (PoI) along the path of β within the graph G

for PoI *in* γ **do**
    $f_i : \sigma_i \rightarrow \gamma_i$, where σ_i is the ground-truth GPS coordinate labeled to γ_i

$\sigma \leftarrow \{\sigma_1, \sigma_2, \sigma_3, ..., \sigma_n\}$, where σ is a set of all labeled ground-truth GPS coordinates on γ

while *there exists available directional path on* $PoIs$ *in* γ **do**
    $\psi \leftarrow \{\{A_1, A_2, A_3, ..., A_n\}, \{B_1, B_2, B_3, ..., B_n\}, \{C_1, C_2, C_3, ..., C_n\}, ..., \{Z_1, Z_2, Z_3, ..., Z_n\}\}$, where ψ is a set of all available directional paths on $PoIs$

for *directional path in* ψ **do**
    $f_i : \xi_i \rightarrow \psi_i$, where ξ_i is the clock orientation labeled to ψ_i and specific to path of β

$\xi \leftarrow \{\xi_1, \xi_2, \xi_3, ..., \xi_n\}$, where ξ denotes a set of all clock orientations labeled on ψ

for *directional path in* ψ **do**
    $f_i : \lambda_i \rightarrow \psi_i$, where λ_i denotes path description corresponds to ψ_i

$\lambda \leftarrow \{\lambda_1, \lambda_2, \lambda_3, ..., \lambda_n\}$, where λ is a set of all path descriptions

$\nabla_i \leftarrow \{\psi_i, \sigma_i, \xi_i, \lambda_i\}$, where ∇_i denotes tuples in a created contextual database, ∇, containing essential attributes

$X \leftarrow \{x_1, x_2, x_3, ..., x_n\}$, where $x_1, x_2, x_3, ..., x_n$ are n number of images labeled with corresponding ψ and X is denoted as an image dataset

$\Omega_i \leftarrow \arg\max_{i \in \psi} P(c_i \mid x_j)$, where Ω_i is the maximum probability prediction of a path, s.t. $\Omega \ni \psi$, as per the implemented CNN and $P(c_i \mid x_j)$ is the probability of directional path class c_i given image x_j

$\Theta_i \leftarrow \{\Omega_i, \nabla_i\}$, where Θ_i denotes the path advancement information based on path prediction by CNN and contextual database

return Θ_i

Fig. 6. PathLookup algorithm.

Preparation Phase of the System

A set of potential points of interest (PoI) was identified as γ (Fig. 2) in a particular outdoor premise and was labeled using ground-truth GPS coordinates, σ. Each PoI in γ is produced with all possible directional paths ψ. The clock orientation, ξ, is identified for all possible directional paths, ψ, and thus used to label those elements. ∇ was constructed as a contextual database to present crucial information about all defined directional paths. ∇ includes, among other things, available directional paths ψ for each PoI, ground-truth GPS coordinates σ for each PoI, clock orientation ξ, and path description λ. In the run-time phase of PathLookup, a directional path query to this database coupled with path prediction, produced by the computer vision element of the proposed algorithm (Fig. 6), would be employed as a wayfinding aid. A large image dataset, X, corresponding to ψ, was produced for the CNN network. The maximum probability prediction of a directional path would be the result of the CNN model.

Run-time Phase of the System

During the run-time phase, a BVI user would capture a picture from their smartphone camera on the route, following proper orientation instructions, through the app while navigating in an outdoor premise and an uncertain circumstance. The app would then transfer the taken image to the server for localization in a specific scene. After receiving a captured image/frame at the server, it would be tested on the CNN model and the results would be sent across the network to the app on the smartphone. The outcome would be in the Θ_i format and could be integrated into any navigation app. This Θ_i could provide BVI individuals in a specific place with accurate scene-specific wayfinding assistance in route decisions.

Table 2: Image dataset for CNN network

	Set 1	Set 2	Set 3
Weather	Rainy/cloudy	Sunny	Sunny
Timing	12pm - 2pm	10am - 12pm	2pm - 4pm
Total no. of images	8589	8485	8821

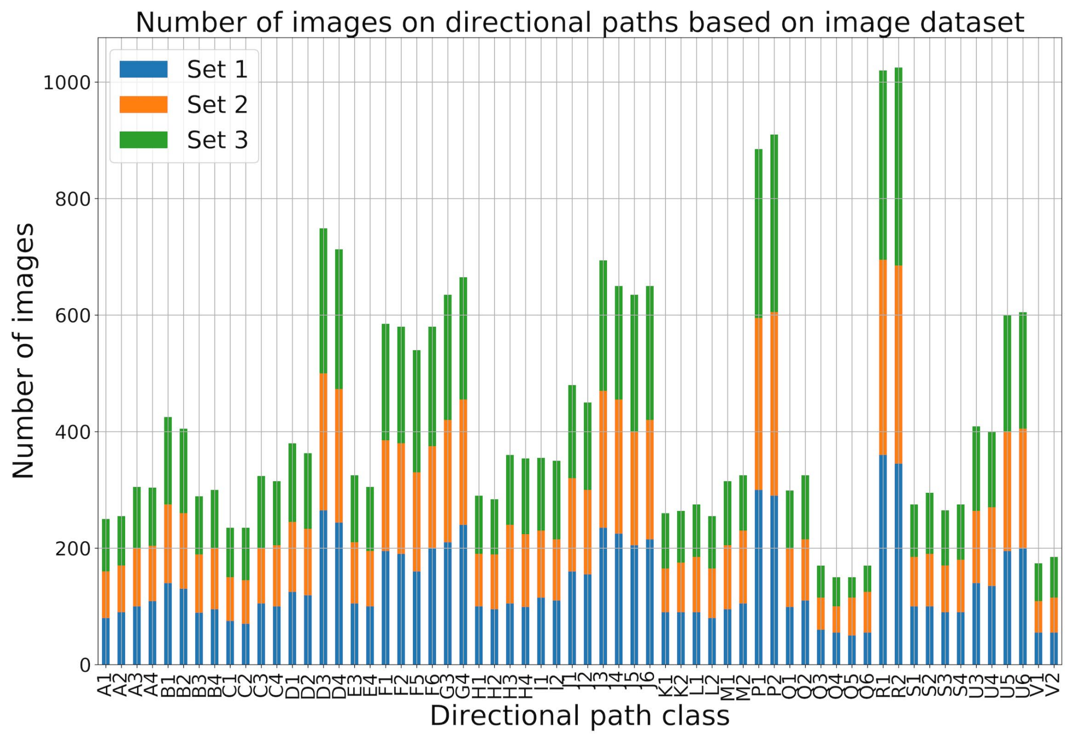


Fig. 7. Distribution of image dataset.

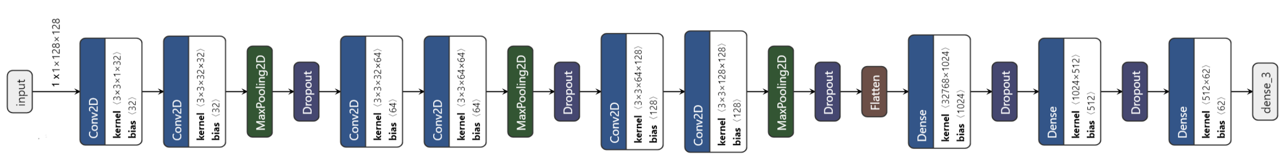


Fig. 8. Architecture of the proposed CNN network of PathLookup.

CNN Architecture Model

In this work, a CNN architecture model (Fig. 8) is proposed and implemented for PathLookup to train the image dataset associated with ψ. To prepare the image dataset, along the path of β = {Wallace Hall, Jabara Hall}, several high-definition videos of ψ were captured by a Samsung Galaxy Note 8 smartphone on the Wichita State University campus and were then converted to images/frames. Three different sets of videos were captured in different weather conditions and timings to create a diverse image dataset for the localization network (Table 2). The smartphone's camera was held at chest level, as a user might when taking an image for path lookup. For the image localization network, a total of 20716 images are used in the training dataset and 5179 images are used in the validation dataset. The distribution of the image dataset across three different sets of data for ψ is depicted in Fig. 7.

To reduce computational complexity, all input images were resized to a resolution of 128 x 128 and only one channel (grayscale) was used in the training. The CNN network is composed of a stack of convolutional modules that perform feature extraction. A final fully connected layer containing the number of ψ units is created for the image localization network. The CNN model is trained to output a probability over ψ directional path classes for each given image. It is noted that the proposed CNN network in PathLookup is a lightweight model that decreases the number of parameters in the network as compared to the classical network models. The balance between model accuracy and calculation speed is maintained by essentially decreasing the complexity of the network and fine-tuning the hyperparameters of the CNN network.

Experimental Results of the CNN Network

The loss and accuracy curves for the training and validation datasets are shown in Fig. 9. The plot of training loss and validation loss, as shown in Fig. 9, declines to a point of stability after epoch 15, and validation loss has a tiny gap with training loss. These results imply that our model's learning curves are well-fitting. Furthermore, the CNN model obtained 97% training accuracy and 99.52% validation accuracy at the end of epoch 40.

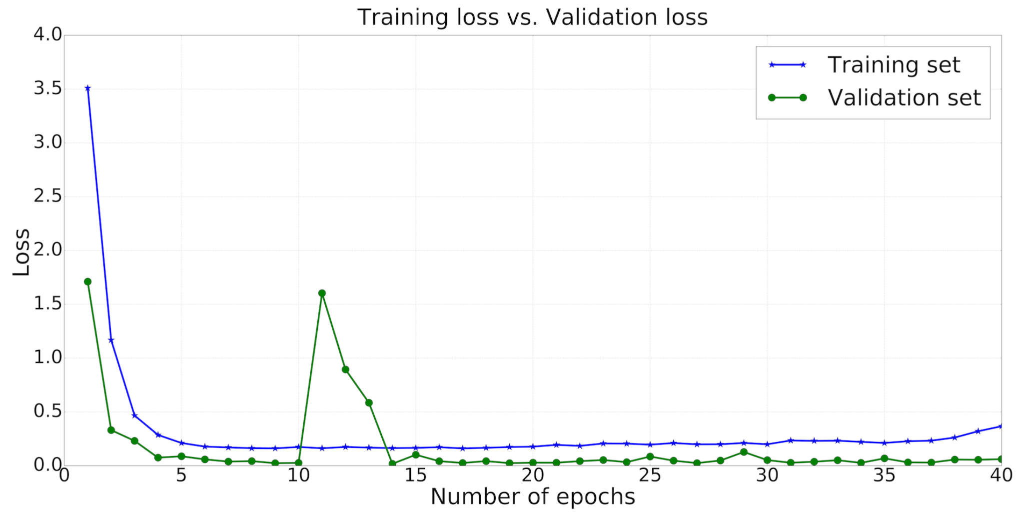


Fig. 9a. Loss curve.

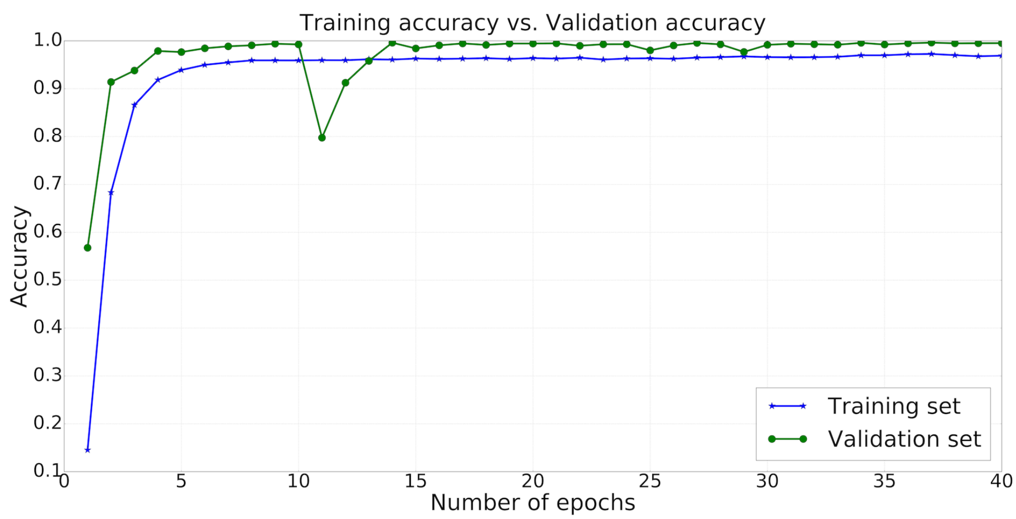


Fig. 9b. Accuracy curve comparison between training and validation dataset.

The classification performance of our model in predicting a directional path in the test site is depicted in Fig. 10. The quantity of how much our model properly predicts out of all the directional path classes in ψ is defined as “Precision” in this plot. As can be observed from the plot (Fig. 10), the value of “Precision” is 1 in 50 of the 62 directional path classes, while it varies between 0.94 and 0.99 in the remaining classes. Furthermore, “Recall” determines how well our model predicts out of all the positive classes. The value of “Recall” is 1 in 48 out of 62 classes, as shown in Fig. 10. The value of “Recall” in the remaining classes ranges from 0.90 to 0.99. Overall, in ψ, a greater value of “Precision” and “Recall” was obtained for all directional path classes. “F1-score” is sometimes a better measure to keep a balance between “Precision” and “Recall” by taking the harmonic mean of those (Fig. 10).

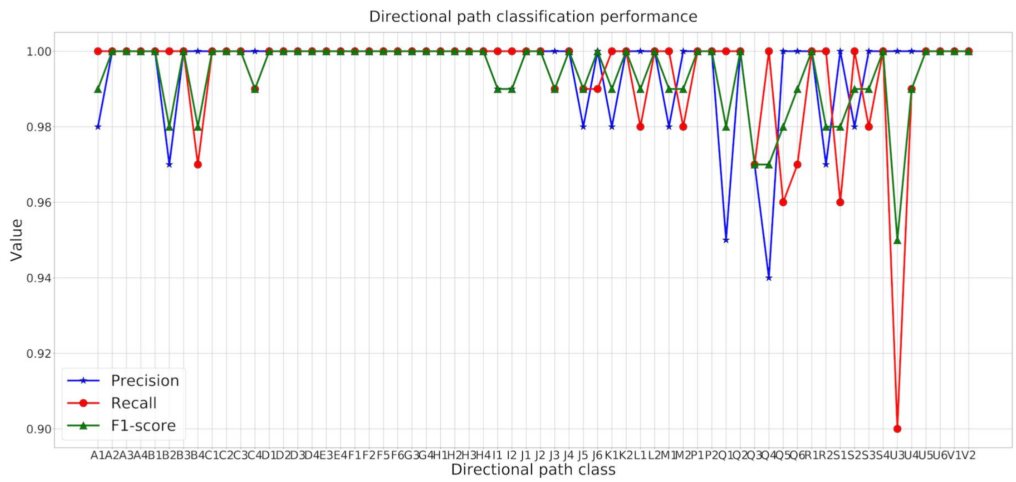


Fig. 10. Classification performance of all directional path classes.

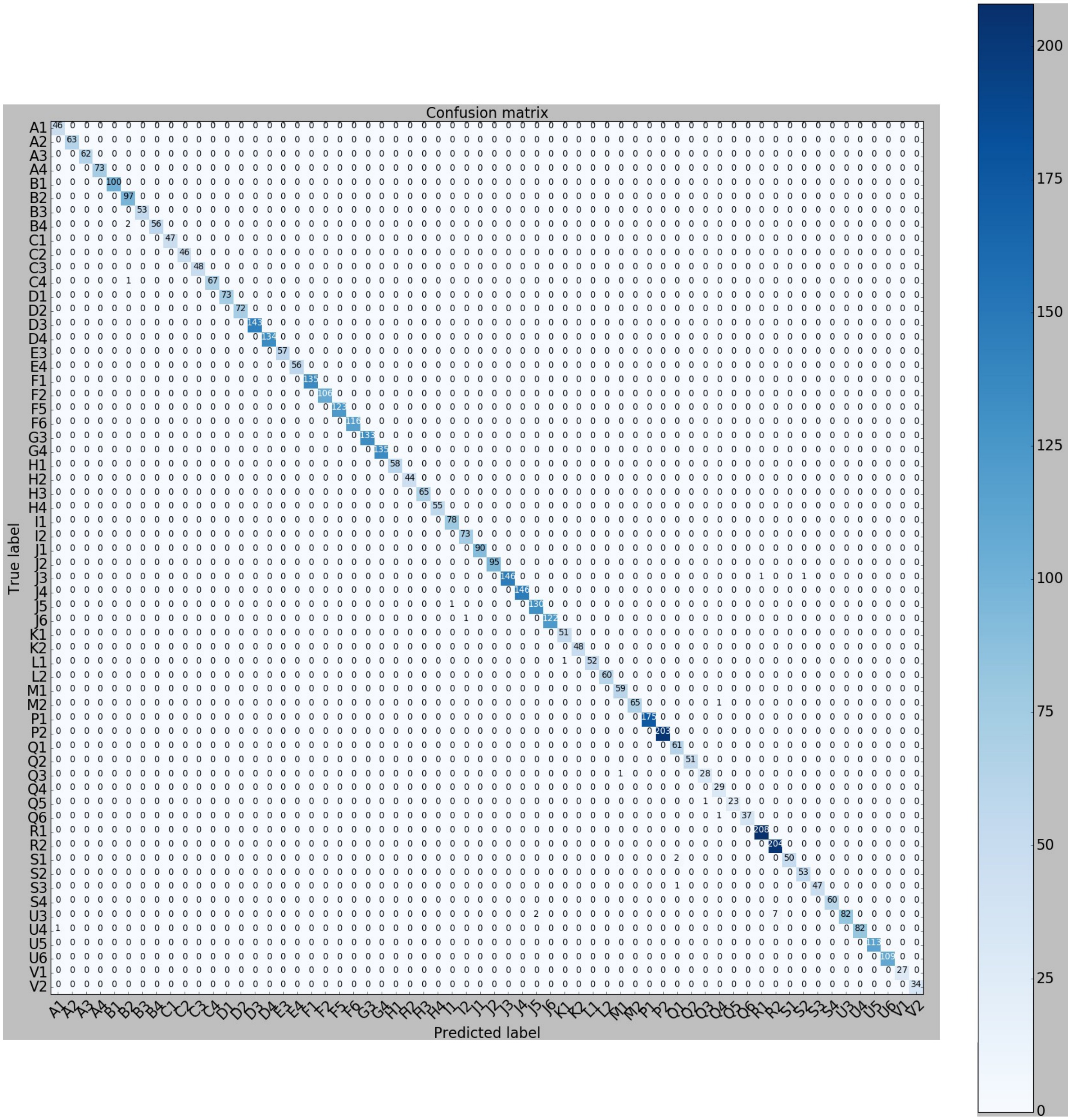


Fig. 11. Confusion matrix of all directional path classes.

Figure 11 illustrates the confusion matrix of our defined model where the diagonal elements represent the number of occurrences for which the predicted label of the directional path is equal to the true label of the directional path in the test site. Off-diagonal elements in this matrix are mislabeled by the model. Since the diagonal values are high and equal to the test values in most cases, it is obvious that the model proposed in this paper works well enough to predict a path.

Fig. 12a. Outcome testing pictures.

Fig. 12b. Outcome testing pictures.

Analysis of Outcome and Approach of the Framework

Outcomes of the Framework

As an example of how the framework performs to predict a path and provide information to the user in their wayfinding, two pictures (not part of the training and validation dataset) shown in this section (Fig. 12) were tested by the framework. The CNN model of this framework predicted these two pictures correctly and within a few milliseconds. The path prediction result illustrates that picture (a) belongs to the directional path class “L1” and the related attribute for “L1” in the contextual database is ∇_{52} = {‘L1’, ‘(37.718894, -97.293127)’, ‘9 o'clock’, ‘This path leads to the Plaza of Heroines circle, the path toward Jabara Hall will be on the right side of it’}. Moreover, the path prediction for picture (b) shows that the picture belongs to the directional path class “Q3” and the related attributes for “Q3” in the contextual database is ∇_{74} = {‘Q3’, ‘(37.718992, -97.293466)’, ‘12 o'clock’, ‘You will have 2 paths ahead; the right one leads to the Jabara Hall; the sharp left one is the adjacent path to the Plaza of Heroines circle that could potentially lead to the Rhatigan Student Center’}. Essentially, this type of information provided by the framework could be utilized in the underlying navigation app for the visually impaired.

Aspects of Offloading Paradigm

Since we proposed leveraging the processing and computational power of a server, not that of a smartphone, the framework that we proposed in this work would be scalable. Beyond just being accurate, PathLookup can perform an image prediction on the order of a few milliseconds using only a commodity laptop computer as the back-end server. The storage, battery, and computational limitations of smartphones are circumvented by exploiting a server’s resources through an offloading paradigm (Shi et al. 2014). The bottleneck of PathLookup will thus be the network connection to the server and the time it takes to get the image file to the server. Given current Internet link speeds, it is possible to get responses from a server in just a few seconds. For instance, with most networks capable of providing speeds of the order of 1 Mbps and more, for an image of size 1 MB, the round-trip delays should be no more than 10 seconds; for faster networks, lookups could be performed well under 5 seconds.

Conclusion

This work established the need for a method to assist BVI individuals in outdoor wayfinding in addition to existing GPS-based technologies by conducting a user study and a quantitative study of GPS accuracies. Following that, PathLookup, a deep learning-based computer vision framework for image localization was demonstrated. The CNN network's path prediction output when combined with other location information from a contextual database as constructed PathLookup can be quite useful for BVI individuals in outdoor wayfinding.

Acknowledgment

This work was supported in part by the US NSF award CNS # 1951864. The content of this work was presented partially at the MPAT workshop of the International Conference on Pervasive Computing and Communications (PerCom), Kassel, Germany, in March 2021.

Works Cited

Ahmetovic, Dragan, et al. "NavCog: a navigational cognitive assistant for the blind." *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 2016.

ARGeoTracking. "ARGeoTrackingConfiguration". Available: https://developer.apple.com/documentation/arkit/argeotrackingconfiguration#3611323

Cheraghi, Seyed Ali, Vinod Namboodiri, and Laura Walker. "GuideBeacon: Beacon-based indoor wayfinding for the blind, visually impaired, and disoriented." *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2017.

Fusco, Giovanni, and James M. Coughlan. "Indoor localization for visually impaired travelers using computer vision on a smartphone." *Proceedings of the 17th international web for all conference*. 2020.

Ganz, Aura, et al. "PERCEPT navigation for visually impaired in large transportation hubs." *J. Technol. Persons Disabilities* 6.1 (2018): 336-353.

GPS accuracy. "Official U.S. government information about GPS and related topics". Available: https://www.gps.gov/systems/gps/performance/accuracy/

Ishihara, Tatsuya, et al. "Beacon-guided structure from motion for smartphone-based navigation." *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2017.

Ishihara, Tatsuya, et al. "Deep radio-visual localization." *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018.

Kim, Jee-Eun, et al. "Navigating visually impaired travelers in a large train station using smartphone and Bluetooth low energy." *Proceedings of the 31st Annual ACM Symposium on Applied Computing*. 2016.

Li, Bing, et al. "Vision-based mobile indoor assistive navigation aid for blind people." *IEEE transactions on mobile computing* 18.3 (2018): 702-714.

Live View. "A new sense of direction with Live View", October 01, 2020. Available: https://blog.google/products/maps/new-sense-direction-live-view/

Nair, Vishnu, et al. "ASSIST: Personalized indoor navigation via multimodal sensors and high-level semantic information." *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. 2018.

Sato, Daisuke, et al. "Navcog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment." Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility. 2017.

Shi, Cong, et al. "Cosmos: computation offloading as a service for mobile devices." *Proceedings of the 15th ACM international symposium on Mobile ad hoc networking and computing*. 2014.

Yang, Zhuorui, and Aura Ganz. "A sensing framework for indoor spatial awareness for blind and visually impaired users." *IEEE Access* 7 (2018): 10343-10352.

Ye, Cang, et al. "Co-robotic cane: A new robotic navigation aid for the visually impaired." *IEEE Systems, Man, and Cybernetics Magazine* 2.2 (2016): 33-42.

Zhang, He, and Cang Ye. "An indoor navigation aid for the visually impaired." *2016 IEEE international conference on robotics and biomimetics (ROBIO)*. IEEE, 2016.

THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

Intelligent Low-Cost 3D LiDAR Tripping Hazard Scanner

Brandon Cai[1], Aditya Patankar[2], Nilanjan Chakraborty[2]
Garcia Research Scholar Program, Stony Brook University[1], Stony Brook University[2]
brandoncai888@gmail.com, aditya.patankar@stonybrook.edu, nilanjan.chakraborty@stonybrook.edu

Abstract

A high-resolution 3D LiDAR scanner was developed to effectively detect tripping hazards with small elevation changes, such as potholes, doorsteps or curbs. Although these obstacles are rarely a problem for people with normal vision, they are the most common tripping hazards for blind and visually impaired people. These small elevation differences can blend in with background road surface, making them difficult to detect by sonar and 2D LiDAR. This paper introduces a new way to turn a low-cost 2D LiDAR into a high-resolution 3D LiDAR and generate a live 3D terrain model that can effectively identify elevation changes as small as 2cm. It can detect a pothole down to 2cm(W)x7cm(L)x2cm(D) in size within 2 meters distance and is operational up to 8 meters.

Inspired by autonomous driving and robotic vision technologies, the ground in the 3D terrain model was removed using Random Sample Consensus (RANSAC) to create segmented point clouds of the tripping hazards, followed by finding the Oriented Bounding Box (OBB) using Convex Hull and Rotating Calipers techniques. With innovative software, each obstacle is not only isolated from the background with a clear location marked, but also has its orientation and dimensions quantified in 3D using an OBB.

Keywords

3D LiDAR scanner, tripping hazard detection, obstacle sensing, computer vision, surface detection, Oriented Bounding Box (OBB).

Introduction

Traditionally a swinging white cane is used by blind and visually impaired people to inspect the ground surface ahead for obstacles. If the cane hits something, then one knows an obstacle is there. The white cane swinging back and forth can only survey the ground at a very coarse resolution that could miss smaller obstacles but still serves as the most effective way to detect tripping hazards with small elevation changes, such as potholes, doorsteps, or curbs. Sometimes, the white cane can be a hazard to other people such as when it taps people's ankles or slips between their feet. Although it is generally respectful to give the white cane user space, a cane by its very nature must impinge on other people's space because the users need to scope out the ground around them on which other people may be standing (“How Dangerous Are White Canes?”, BBC News).

Other alternative assistive technologies using sonar and 2D LiDAR are only effective in detecting large above ground objects and cannot detect small elevation differences that blend in with the background road surface (“The UltraCane”) (Shaoul) (Spiess). On the other hand, many computer-vision and Machine Learning (ML) based approaches can identify known objects (Kim, 703) (EdjeElectronics), such as people and street furniture in front of the visually impaired person but lack accurate distance/depth perception to detect and estimate the size of potholes and other small elevation changes with unpredictable shape and color. With centimeter accuracy in distance sensing, LiDAR technology holds promise in distinguishing small evaluation differences required for pothole detection. However, high end commercial 3D LiDARs designed for autonomous driving, such as the $75,000 Velodyne HDL-64E with 64-line scanning capability, are too expensive to be an assistive technology solution.

Experiments and Discussion

Design and Build the High-Resolution 3D LiDAR

This paper introduces a new way to turn a low cost 2D LiDAR into a high-resolution 3D LiDAR and generate a live 3D terrain model that can effectively identify elevation changes as small as 2cm. As shown in Fig.1, this low-cost 3D scanner is built by combining a $300 Slamtec RPLiDAR 2D LiDAR scanner with a 3rd dimension added by controlling the tilt angle of a $17 LewanSoul LX-224 digitally programmable servo motor. The LiDAR is designed for vacuum robot and home robot applications with an FDA Class 1 laser certification that is considered safe under all circumstances. The LiDAR can perform a 2D 360° scan 10 times a second with a 12-meter range. For applications that do not require fast image refresh, the servo motor can have a scanning resolution down to 0.5°, comparable to an expensive 3D LiDAR's resolution.

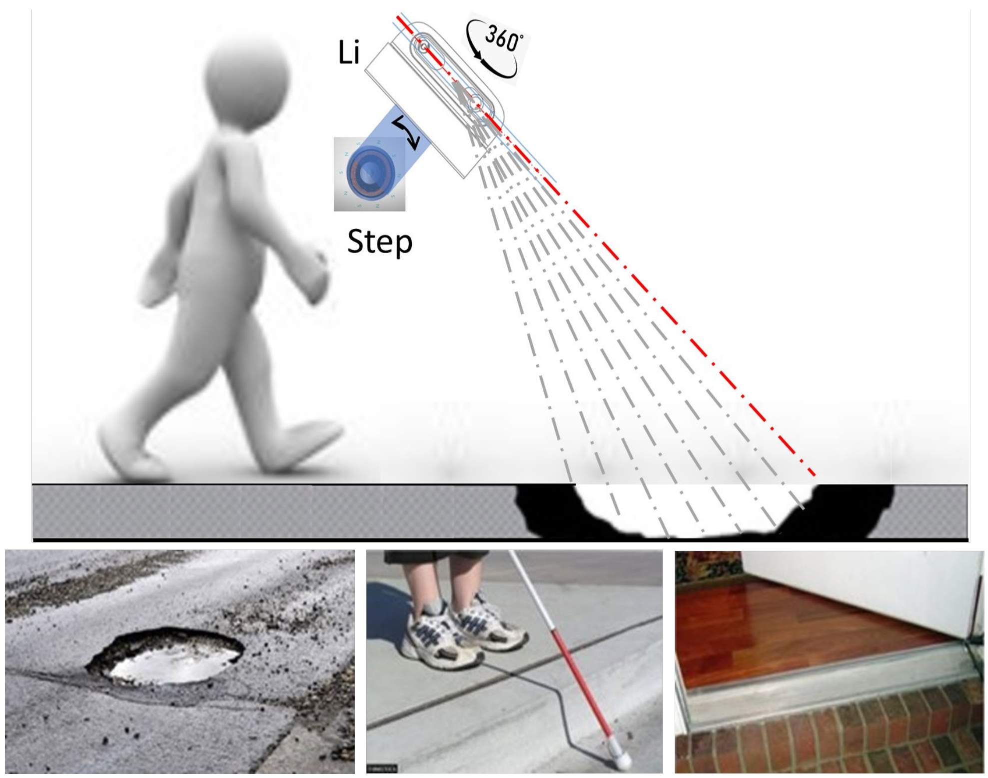


Fig. 1. Conceptual drawing: homemade 3D LiDAR scanner in this project for uneven road surface detection for the visually impaired.

The system is controlled by a $59 mini Raspberry Pi 4 computer connected to the LiDAR and motor through high speed USB3 ports. Software development is done in Python under a Linux operating system. To securely mount the LiDAR on top of the swinging servo motor, a custom-made mounting base was 3D printed with PLA plastic. The full integrated 3D scanner hardware system is shown in Fig.2.

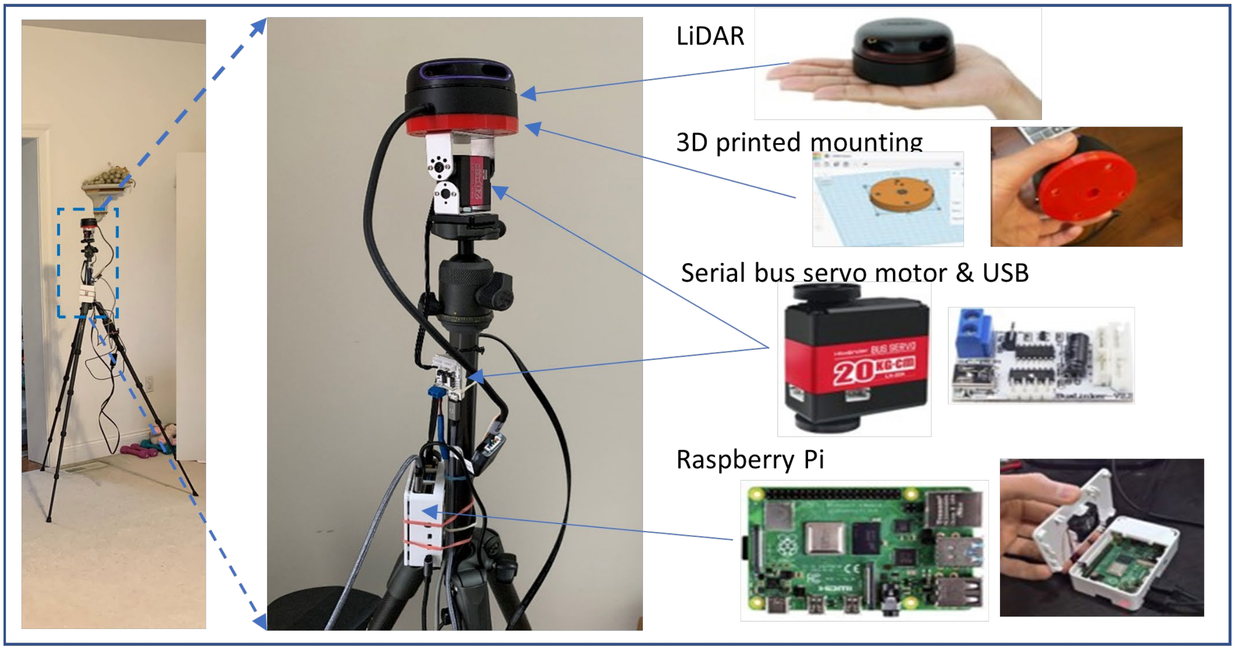


Fig. 2. Pictures of the 3D LiDAR system and with key components illustrated.

Develop 3D Scan Software and Algorithm to Generate the 3D Terrain Model

The first block of the software was developed to enable a 2D LiDAR scanner to acquire the distance and associated panning angle and then save the data in an array for data processing (Hyun-je). The second block of code developed was for controlling the servo motor tilt through the Raspberry Pi computer. Then, the LiDAR and servo motor codes were integrated into a coherent 3D scanner based on the flow chart shown in Fig.3. In essence, the servo motor tilts the LiDAR base to a certain angle and waits for the 2D LiDAR to scan 360° on that tilted plane and then increments the tilt to the next angle. The raw data captured contains the 3D terrain profile in terms of tilt angle of the servo, panning angle of the lidar, and distance (θ, φ, r).

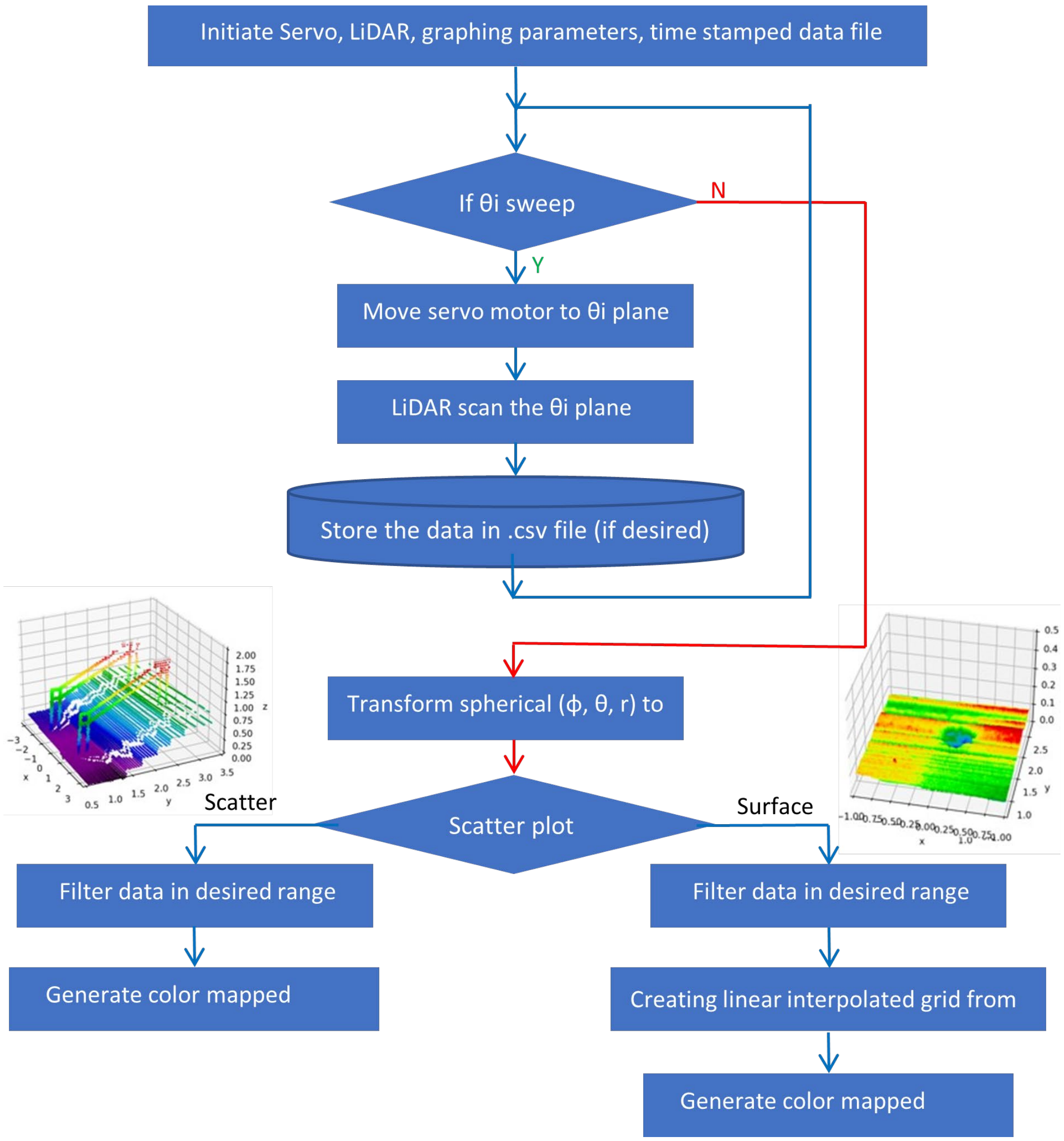


Fig. 3. Flowchart of the Python program developed for the 3D scanner integrating LiDAR control, servo motor control, data processing and 3D graphing capabilities.

The actual coordinate transformation used is explained in Fig.4. The LiDAR acquired data is referenced around its local polar (φ, r) frame. This coordinate transformation is needed to account for the LiDAR base rotating around the servo motor axle by some angle θ and to project the coordinates to the user's Cartesian frame. A systematic error is introduced by the LiDAR scan plane sitting 6.7cm away from the servo motor axle. The correction algorithm is also illustrated in Fig.4.

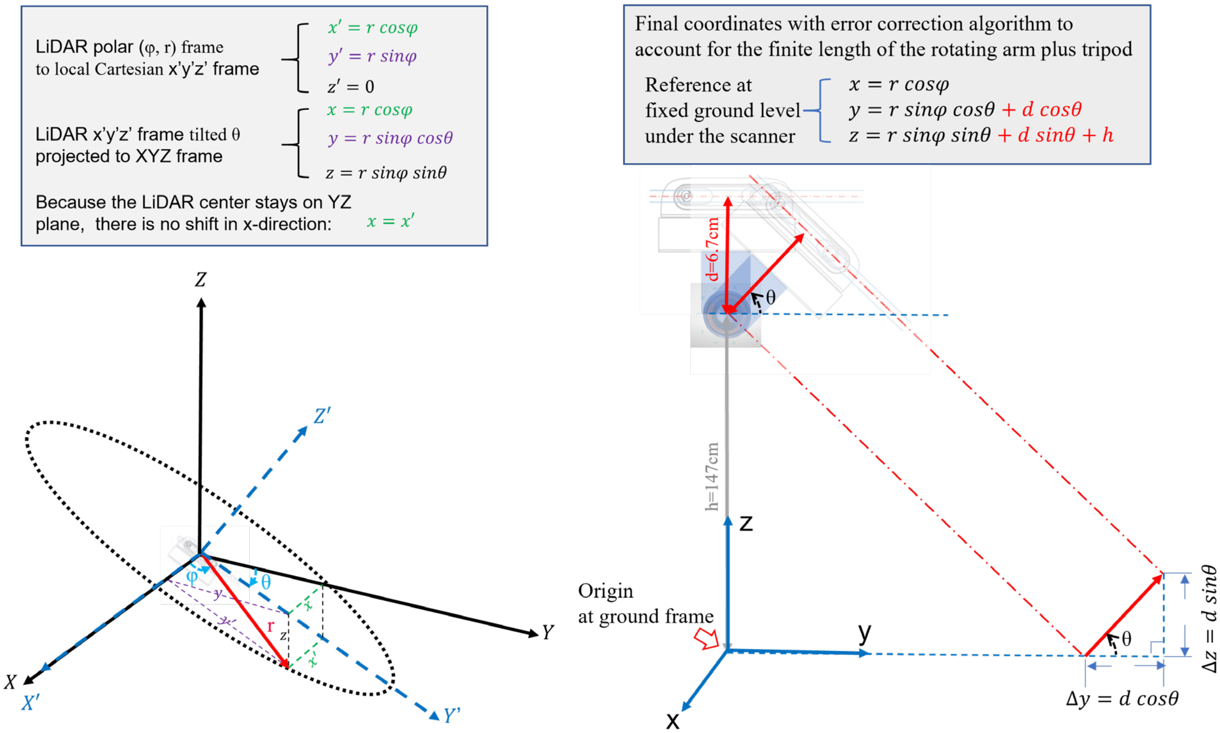
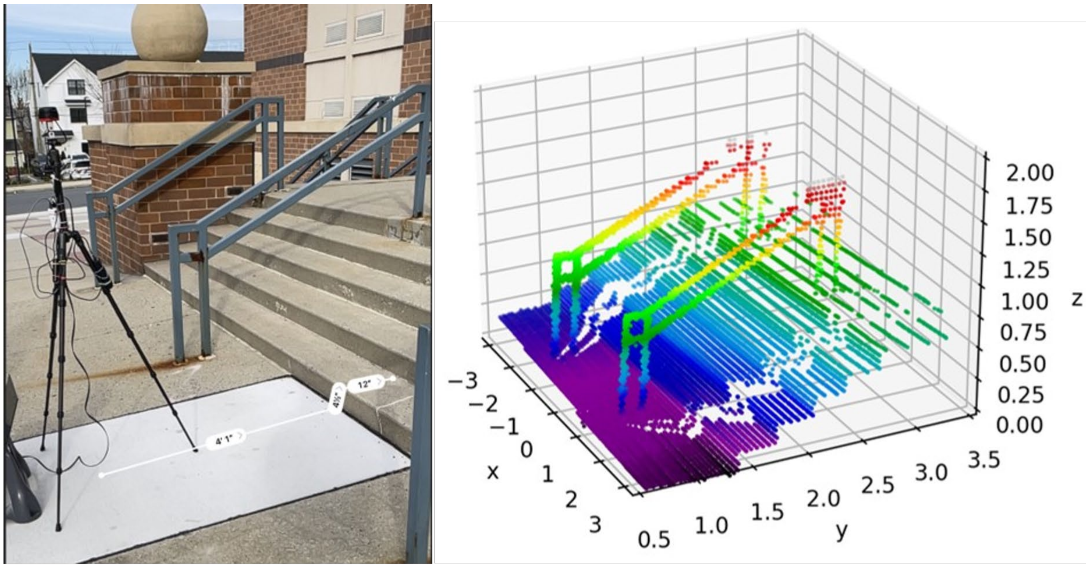


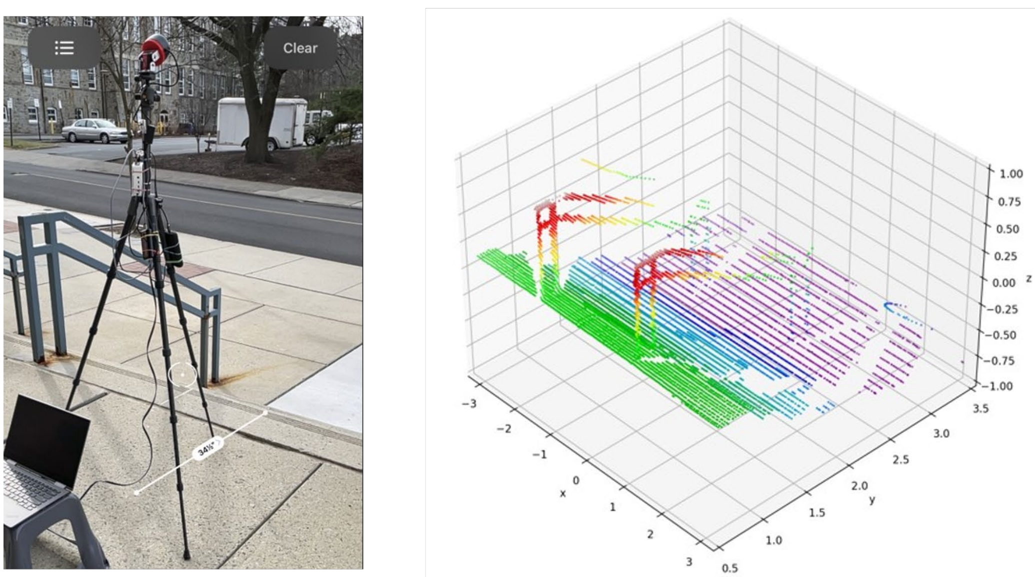
Fig.4. The coordinate transformation developed included error correction terms to remove the systematic error introduced by the LiDAR located off from the motor axle.

Outdoor Experiment Results for Detecting Real Life Tripping Hazards

To demonstrate the LiDAR system's spatial resolution in detecting the challengingly small elevation changes, the elevation difference is highlighted in the 3D plot as different colors using the Python matplot.lib color map as shown in Fig.5.



(A) 1.25 meter from the bottom of the stairs



(B) 0.88 meter from the top of the stairs

Fig. 5. Two 3D LiDAR scans of the stairs were performed with the scanner placed at: (A) 1.25m away from the bottom of the stairs showing the increase in elevation, and (B) 0.88m away from the top of the stairs showing the descent in elevation.

Real world small elevation change tripping hazard detection capability with the 3D LiDAR can be found in the 3D surface plots shown in Fig.6. With the complete xyz-terrain model extracted from LiDAR, the user has the liberty to generate any cross-sectional views.

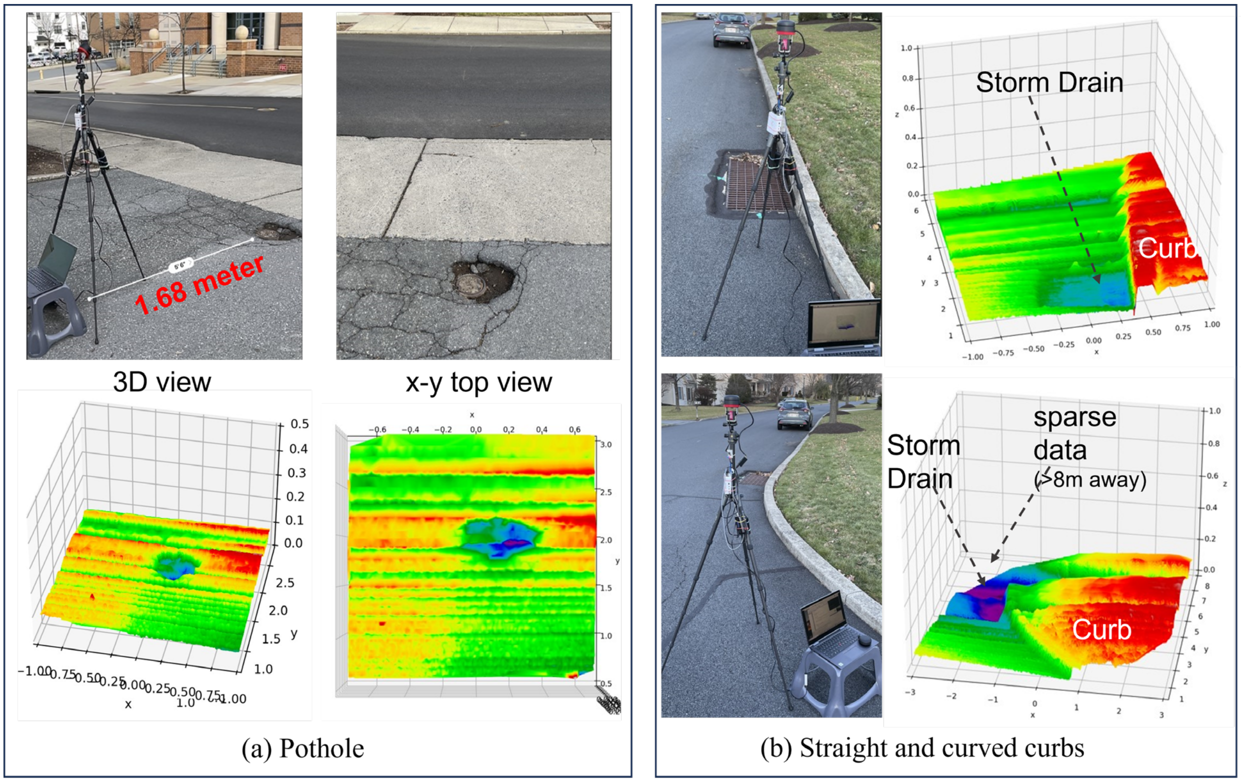


Fig. 6. The 3D scanner captured a pothole, straight and curved curbs, and storm drains.

Quantitative Accuracy Evaluation of the 3D LiDAR

Inspired by techniques for evaluating photography image sharpness, 3D imaging of the sharp edges of long wood boards was used to quantify the measurement resolution in multiple directions with an automatic edge detection code developed to maximize the number of edge points included along the 4 edges of two boards. As shown in Fig.7, the maximum deviation from a curve fit is the worst-case resolution.

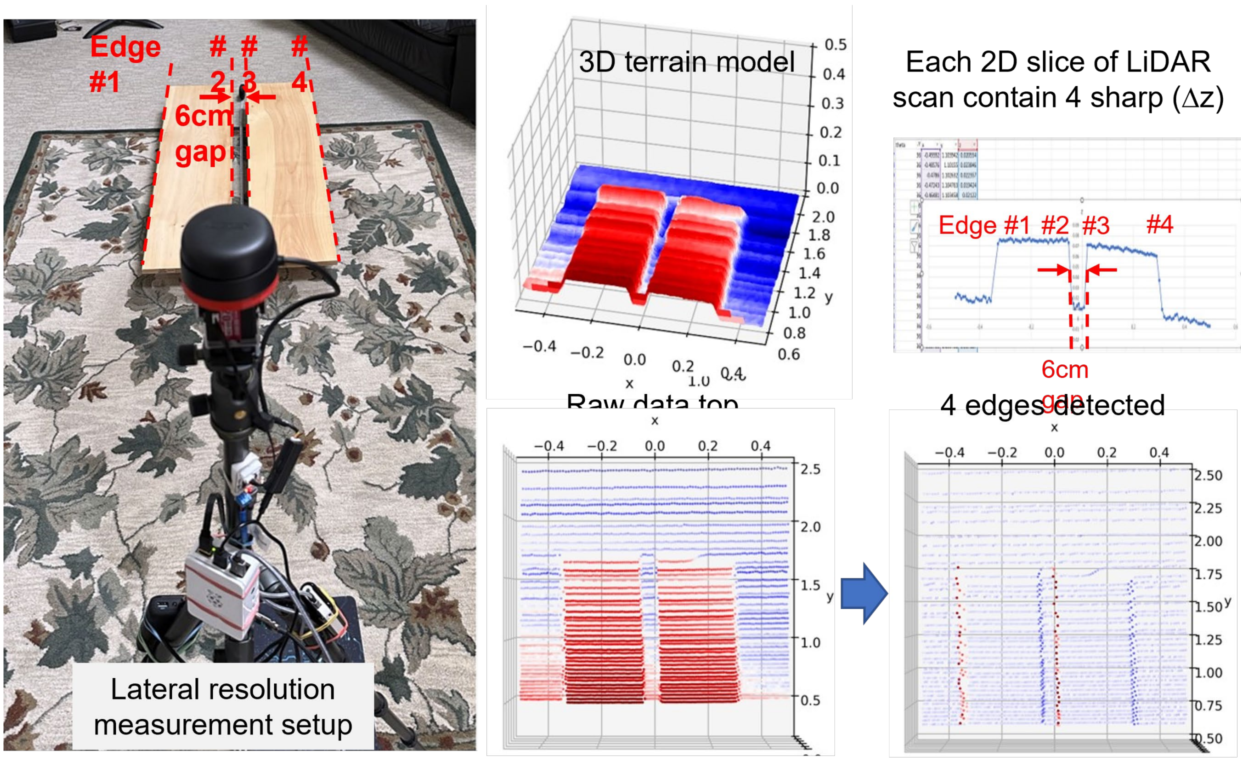


Fig. 7. Lateral resolution measurement setup with two wood boards were placed parallel with a 6cm gap in between.

The feasibility of this 3D sensor has been demonstrated to reliably identify potholes with minimum size of 2cm(W)x7cm(L)x2cm(D) at a distance up to 2 meters ahead of the traveler and is operational at up to 8 meters ahead (>4x the range of traditional white canes).

Computer Vision based 3D Object Detection Algorithm

In order for visually impaired and blind people to decipher the road condition obtained from the 3D LiDAR scanner data, further data analysis is needed to intelligently abstract the potential road hazards in descriptive terms for a voice alert. Inspired by computer vision and robotic vision technologies, the raw LiDAR 3D point cloud data is fed into a digital data processing algorithm that has been developed to isolate each obstacle and mark it with a bounding box, which provides approximate location and dimensions of the obstacle. Multiple steps of data processing are required to achieve this goal.

Separating Obstacles from the Ground Surface

We convert the irregular point clouds obtained from the 3D LiDAR into a voxel grid—a regular 3D grid. In the second step, we remove the ground using Random Sampling Consensus (RANSAC). RANSAC finds the plane of the ground by sampling 3 random points to calculate a plane through those points. Then, it calculates the distance from each of the other points in the point cloud to that plane, and if that distance is less than the distance threshold, it adds that point as an inlier. This process is repeated a fixed number of times and the plane with the maximum number of inliers is used as the ground, which is then excluded while the outliers are considered for object detection as shown in Fig. 8.

Fig. 8. Ground detected using RANSAC with the pothole identified as an outlier point cluster.

With outliers off the ground surface isolated as segmented clusters, very small clusters of points are removed as small imperfections on the road surface as they are not big enough to be considered obstacles. This makes the object detection robust against noise in the LiDAR data. To group local point cloud clusters together, an Open3D density-based clustering algorithm DBSCAN was used. Multiple objects can be isolated into separate clusters, which enables this LiDAR data processing algorithm to report multiple tripping hazards in the same point cloud, as shown in Fig.9.

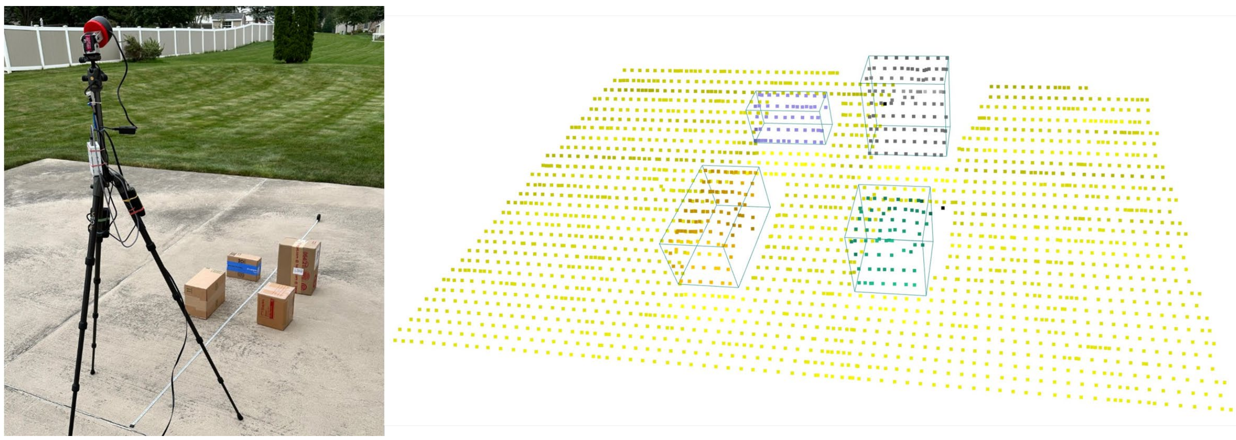


Fig. 9. Multiple outlier clusters detected and isolated into individual objects for OBB calculation.

Abstracting Location and Dimension Info from Each Obstacle

The next step in the computer vision algorithm is to mark each detected object with a bounding box, a rectangular prism that surrounds an object to provide the dimensions. For a bounding box to show the object location and size accurately, we experimented with the most basic axis aligned bounding box, which simply calculates the minimum and maximum x-, y-, and z-coordinates of all the points to create a bounding box with edges parallel to the axes even if the object is not aligned with the coordinate axes, which always overestimates object size for objects with a misaligned base. As shown in Fig.10, there is an 84% estimation error in volume.

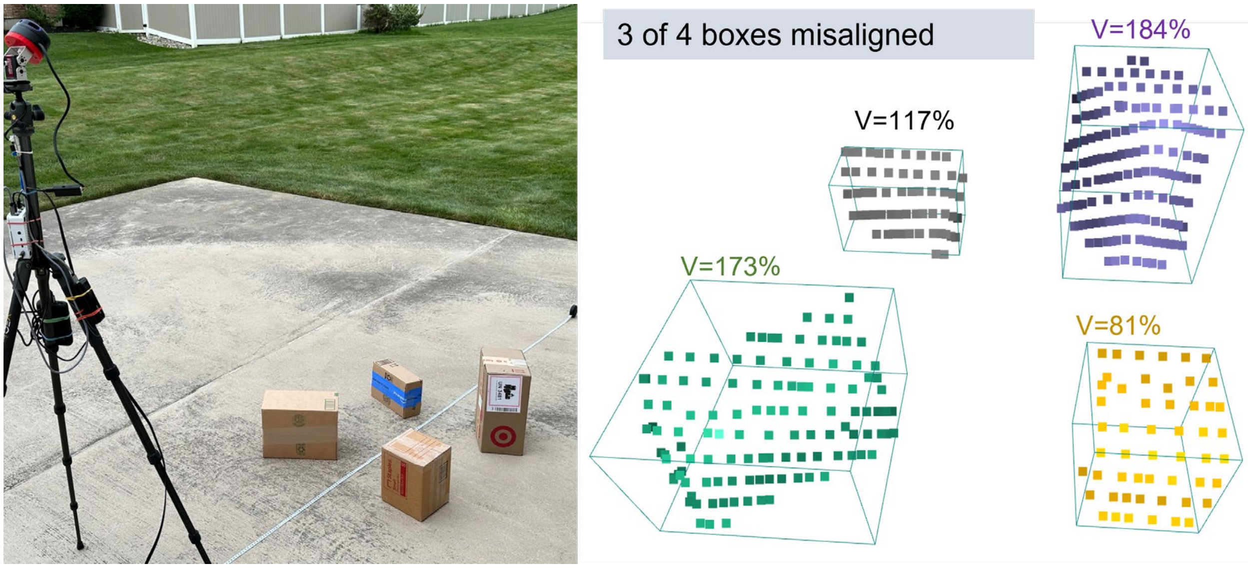


Fig. 10. The limitation of using axis aligned bounding box for size estimation is that misaligned objects will have exaggerated sizes.

To properly estimate the size of the object, the orientation of each object needs to be detected. The Open3D library does have the ability to generate Oriented Bounding Box (OBB) using Principal Component Analysis (PCA) by calculating covariance matrix of the point cloud to detect the orientation. But because the objects are on the ground, we can assume the bottom side lays flat on the ground and simply project the point cloud down to the xy plane and calculate the 2D PCA bounding box before adding the height for the final OBB. This includes calculations of both the variance and covariance in the x and y directions as a matrix. However, as shown in Fig.11, this method can be fooled by an uneven distribution of the points as every point in the cloud carries the same weight in orienting, resulting in misalignment.

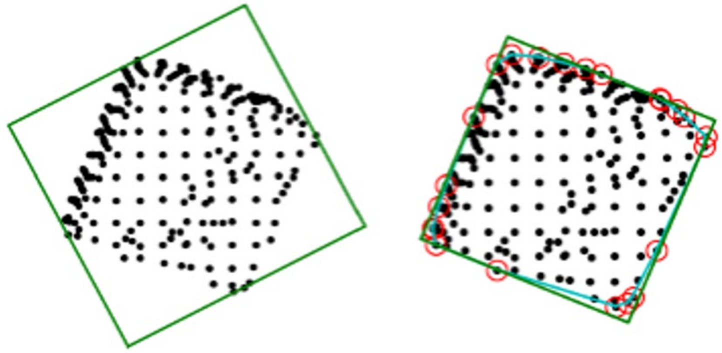


Fig. 11. The limitation of PCA based orientation bounding box for size estimation. Left – Non-optimal PCA based OBB, Right – Optimal OBB (desired).

The main achievement for this project in object estimation was the creation of a customized bounding box algorithm using the Convex Hull (Burger) and advanced Rotating Calipers algorithm (Naujoks, 1) (Chang, 112) (Arnon) to guarantee the best fitting bounding box generation. As shown in Fig.12, the convex hull of the point cloud is calculated. The convex hull is a polygon containing some points of the cloud as vertices such that all the other points are inside the polygon. We can thus ignore all the points inside the convex hull when calculating the optimal bounding box as they do not affect the optimal bounding box of the entire point cloud. Now, for every edge of the convex hull, we rotate the hull such that that edge is parallel to one of the axes and calculate the axis aligned bounding box. After doing this for all the edges, we pick the one with the minimum area and extend the 2D bounding box vertically to get a 3D Oriented Bounding Box parallel to the ground.

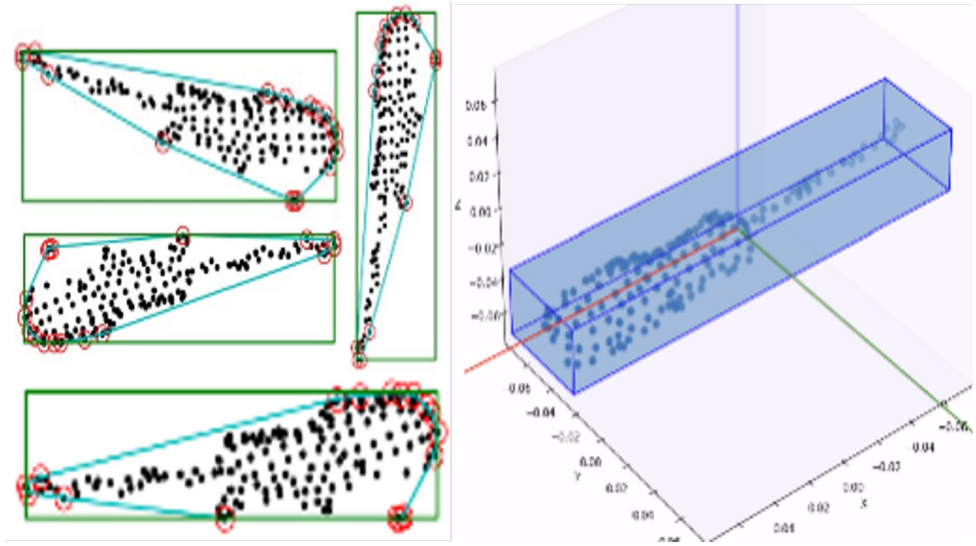
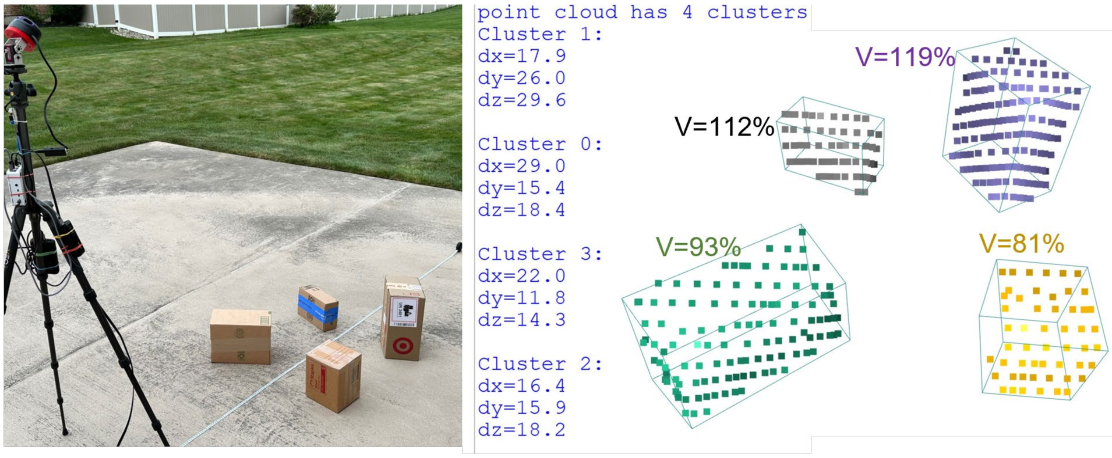
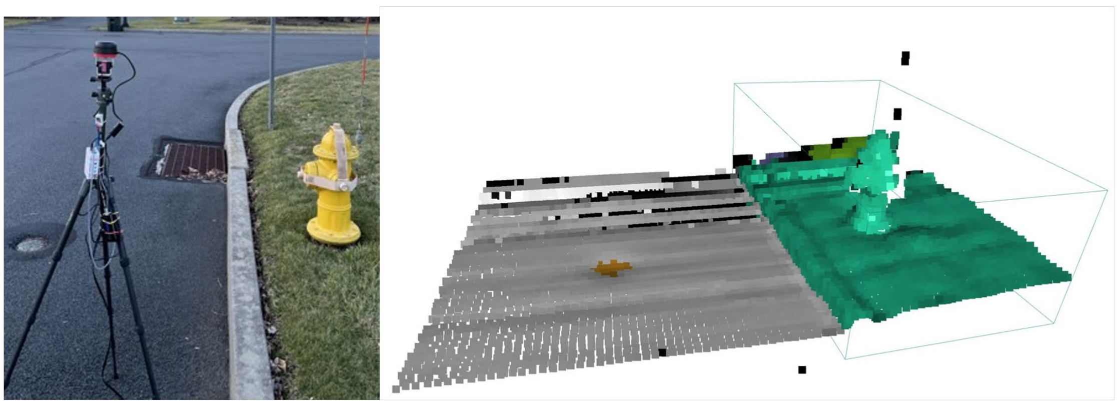


Fig. 12. Creating a customized OBB algorithm using Convex Hull and Rotating Calipers. Left – Screwdriver rotated to find optimal bounding, Right – Oriented Bounding.

The Oriented Bounding Box (OBB) enhanced the accuracy of object size estimation. As shown in Fig.13 (a), +/-20% accuracy in volume estimation of various small boxes with 14cm to 30cm sides was observed. Fig. 13 (b) illustrates a detection of multiple objects after removing the road surface (shown in gray) including the green cluster for the curb and fire hydrant to the right of the user and the brown cluster for a small indent in the middle of the road.



(a) Optimal OBB captured boxes with different orientations:
OBB volumes (V) are with ±19% of the actual



(b) Multiple object clusters isolated: curb & fire hydrant, a small indent, road surface

Fig. 13. Rotating Calipers based optimal Oriented Bounding Box calculation effectively detected rotated orientation.

In a test on a complex real-world environment, as shown in Fig.14, the LiDAR data was automatically processed and abstracted into 5 obstacles with bounding boxes around them. This includes the doorstep right in front of the user, the plant to the right, the plant and the wall on the left, the wall around the door, and the closed door a few inches indented into the wall.

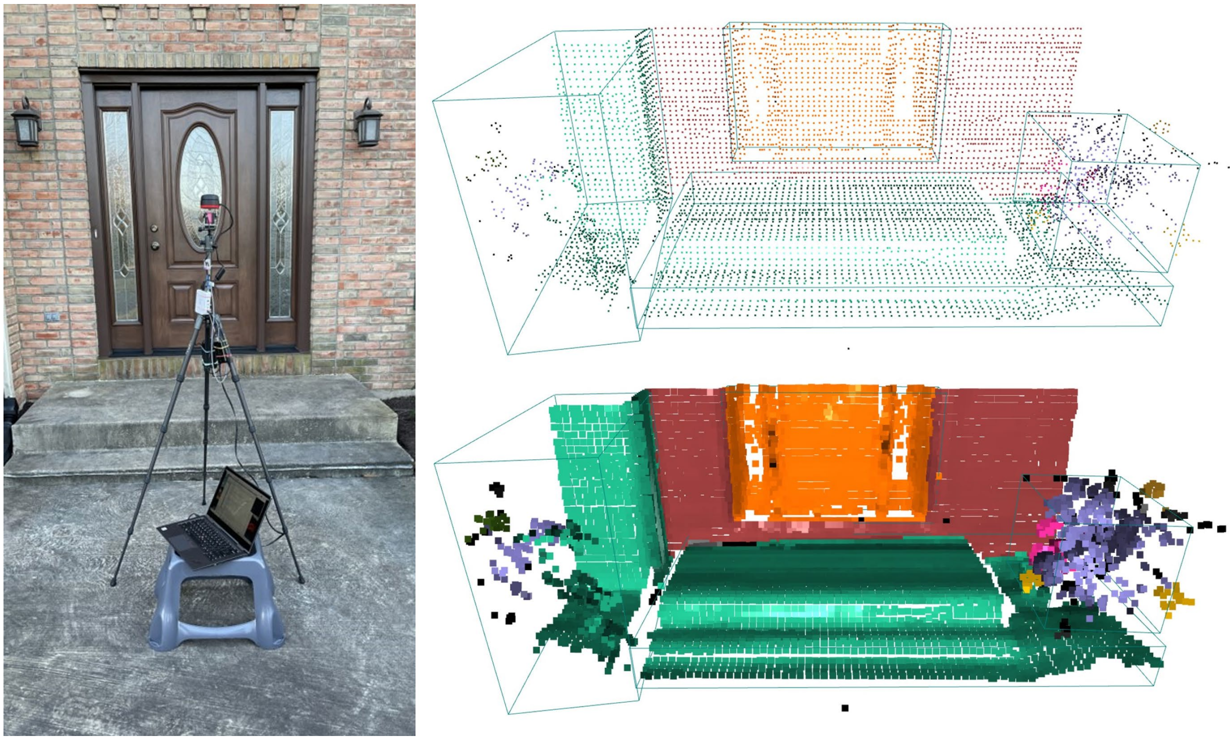


Fig. 14. A more complex case with many obstacles detected by the LiDAR equipped with computer vision capability abstracted the surrounding into 5 objects.

Conclusion

This project successfully demonstrated that a high-resolution 3D LiDAR scanner can be made by adding the third dimension to a 2D LiDAR using a digital servo motor. A fully functional software algorithm for generating and graphing the 3D terrain profile has been created in Python with coordinate transformations, 3D interpolations, a unique LiDAR off-axle error correction algorithm, and a high contrast color-coded surface plot of the terrain. The 3D LiDAR scanner can reliably identify potholes of a minimum size of 1x5x2cm(WxLxD) up to 1 meter in distance, 2x7x2cm(WxLxD) up to 2 meters in distance, and 2x13x4cm(WxLxD) up to 4 meter in distance.

Inspired by autonomous driving and robotic vision technologies, an automatic object detection and dimension estimation algorithm has been developed in Python. As demonstrated, the ground in the 3D terrain model was removed using Random Sample Consensus (RANSAC) to create segmented point clouds of each tripping hazard, followed by finding an Oriented Bounding Box (OBB) using Convex Hull and Rotating Calipers techniques. In this way, each obstacle is not only isolated from the background with its clear location marked, but also the dimensions and orientation are quantified in 3 dimensions using an OBB.

This technique opens the door to 3D LiDAR being used for high resolution ground level tripping hazard detection, which is not feasible using sonar or 2D LiDAR. With new MEMS-based 3D LiDAR technology in development (Yoo, 408) or very high-resolution mechanical scanning LiDAR with a much faster scanning speed, this pothole detection technique could even be expanded to fast moving vehicles on the road.

Future Research

There are two focused areas for future improvements. First is to improve the scanning speed. Despite of the fast 10Hz scan rate and a 4.8Gbps USB3 data link, the Slamtec RPLiDAR has an unexpected requirement of a disconnect command before reading the acquired data, this repetitive connect and disconnect resulted in a scan time as long as 30 to 50 seconds depending on the desired range. Besides trying new LiDAR models without the dead time, the next prototype in development will also include a stereo camera with a 30fps frame rate.

Secondly, with the target being a wearable assistive device, the next prototype in development includes an integrated gyroscope for camera angle detection and text-to-speech to deliver voice alerts as the user interface for those who are blind or have severely compromised vision.

Acknowledgement

In the summer of 2023, the coauthor Brandon Cai was fortunate enough to find a research opportunity at Dr. Nilanjan Chakraborty's Interacting Robotic System Laboratory at Stony Brook University through the Gacia Research Scholar Program. After Brandon created the Rotating Calipers based OBB algorithm for robot manipulation planning, it inspired Brandon to introduce the computer vision technologies to his independent project in building assistive device for the visually impaired. In 2022 before the work at Stony Brook, Brandon already constructed the 3D LiDAR with the software developed for spatial transformation and graphing. For the early stage of LiDAR development, Brandon would like to acknowledge the support from Dr. Luis Rodriguez from the Naval Surface Warfare Center Dahlgren Division for suggestions on general practice in data collection and possible sun glare on the LiDAR sensor. This suggestion also inspired the idea of using this LiDAR sensor as a night vision device.

Works cited

"About the UltraCane." *UltraCane*, https://ultracane.com/about_the_ultracane.

Arnon, Dennis S., and John P. Gieselmann. "A linear time algorithm for the minimum area rectangle enclosing a convex polygon." Computer Science Department, Purdue U, 7 Dec. 1983. *Purdue e-Pubs*, https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1381&context=cstech.

Burger, Patrick, and Hans Jaochim Wuensche. "Fast multi-pass 3D point segmentation based on a structured mesh graph for ground vehicles." *IEEE Intelligent Vehicle Symposium*, 2018, https://ieeexplore.ieee.org/abstract/document/8500552.

Chang, Chia Tche, et al. "Fast oriented bounding box optimization on the rotation group SO(3,ℝ)." *ACM Transactions on Graphics*, vol. 30, no. 5, 22 Oct. 2011, https://dl.acm.org/doi/10.1145/2019627.2019641.

EdjeElectronics, Evan. "How to run TensorFlow Lite Object Detection Models on the Raspberry Pi (with Optional Coral USB Accelerator)." *GitHub*, https://github.com/EdjeElectronics/TensorFlow-Lite-Object-Detection-on-Android-and-Raspberry-Pi/blob/master/deploy_guides/Raspberry_Pi_Guide.md.

"How Dangerous Are White Canes?" *BBC*, 18 Nov. 2015, https://www.bbc.com/news/disability-34855311.

Hyun-je. *pyrplidar*. *GitHub*, 5 Oct. 2019, https://github.com/Hyun-je/pyrplidar.

Kim, Taehyeong, and Seung Ki Ryu. "System and method for detecting potholes based on video data." *Journal of Emerging Trends in Computing and Information Sciences*, vol. 5, no. 9, Sept. 2014, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7543c374269e45cc08ae16fc03ef048d6b2ca57a.

Naujoks, Benjamin, and Hans Joachim Wuensche. "An orientation corrected bounding box fit based on the convex hull under real time constraints." *IEEE Intelligent Vehicle Symposium*, 2018, https://ieeexplore.ieee.org/document/8500692.

Shaul, Yorai. "A full autonomous stack, a tutorial | ROS + Raspberry Pi + Arduino + SLAM." *Yoraish*, 8 Sept. 2021, https://yoraish.wordpress.com/2021/09/08/a-full-autonomous-stack-a-tutorial-ros-raspberry-pi-arduino-slam/.

Spiess, Andreas. "The ugly truth about LIDARs for obstacle avoidance (RPLIDAR A1, Arduino, ESP8266)." *Youtube*, 27 Sept. 2018, https://www.youtube.com/watch?v=AERfie9wlWM

Yoo, H.W., et al. "MEMS-based lidar for autonomous driving." *E&I Elektrotechnik Und Informationstechnik*, original series, vol. 135, 31 July 2018, pp. 408-415.

The Journal on Technology and Persons with Disabilities

Use of Braille in the Workplace by People Who Are Blind

Michele C. McDonnall[1], Rachael Sessler-Trinkowsky[2], Anne Steverson[1]
National Research & Training Center on Blindness & Low Vision[1],
University of Massachusetts Boston[2]
m.mcdonnall@msstate.edu, r.sesslertrinkowsky@umb.edu,
asteverson@colled.msstate.edu

Abstract

Interest in the benefits of braille for people who are blind is high among professionals in the blindness field, but we know little about how braille is used in the workplace. The broad purpose of this study was to learn how employed people who are blind use braille on the job. Specific topics investigated included: work tasks refreshable braille technology (RBT) is used for, personal and job characteristics of RBT users compared to non-users, and factors associated with RBT use among workers with at least moderate braille skills. This study utilized data from 304 participants in a longitudinal research project investigating assistive technology use in the workplace by people who are blind. Two-thirds of our participants used braille on the job, and more than half utilized RBT. Workers who used RBT did not necessarily use it for all computer-related tasks they performed. RBT use was generally not significantly related to job characteristics, except for working for a blindness organization. RBT use was not significantly related to general personal characteristics but it was significantly different based on disability-related characteristics. Only older age and higher braille skills were significantly associated with RBT use on the job in a multivariate logistic regression model.

Keywords

Blind, legally blind, braille use, assistive technology, employment

Introduction

Interest in the benefits of braille use by people who are blind is high among professionals in the blindness field. There are many potential benefits to using braille as opposed to only using audio output for people who are blind. Braille is an important "literacy medium" (Wormsley) in a world that has continued to expand the availability of information through technology. Benefits to using braille include tactile access to spelling and text structure, as well as a method to quickly scan text (Rempel). Braille technology has made braille access even more readily available. A qualitative study by Martiniello et al. identified misconceptions about braille, including braille is only for those who are totally blind, braille is too difficult to learn, and braille is only useful for lengthy readings. Many people have asserted that braille use has decreased substantially over the past several decades, but recent research determined that this assertion is not based on fact, as the data to determine this is simply not available (Sheffield et al.).

Research related to braille has primarily focused on teaching and learning braille (Hall and Newman; Hoskin et al.; Martiniello et al.; Emerson et al.; Bickford and Falco). Correlates of braille use have also been a focus area of a small body of research. A few studies investigated the relationship between regular braille use and employment and found a positive association (Bell and Mino; Bell and Silverman; Ryles). Another study investigated the relationship between braille reading status and several outcomes. This study found that braille readers since childhood had higher life satisfaction, self-esteem, and employment rates compared to all others, and that people who learned to read braille at a later point in life had higher outcomes in the areas compared to people who never learned to read braille (Silverman and Bell).

Very little research has investigated how people actually use braille in their daily lives. Only two research studies that addressed braille use were found, and no studies investigated how braille is used at work. A 1998 study focused on braille communication modes taught by teachers of students with visual impairments (TVI) and used in practice by 71 adults (Allman). The other study, published in 2012, was a qualitative study of 12 braille users aged 16 to 20 who were high school or college students (D'Andrea). This study provided information about devices and tools the students used for reading and writing, as well as preferences for using braille. Both studies are dated and used relatively small samples.

Other sources have provided personal accounts of how people use braille in their daily lives. The National Federation of the Blind's publication *The Braille Monitor* has published many such stories. A journal article provided a description of four people's use of braille from a panel session at a conference in 1989 (Huebner). This article provided many examples of the ways in which braille was being used by the four panelists in everyday life. While existing information provides support for the usefulness of braille to people on a personal level, it does not inform us about braille use at work.

Given the importance that professionals in the field place on braille, we thought it would be valuable to explore how braille is being used by employed people who are blind. The broad purposes of this study were to learn how many employed people use braille and how braille is being used on the job. Five research questions (RQ) were addressed.

1. What percentage of employed people who are legally blind utilize braille on the job and what percentage utilize refreshable braille technology (RBT) on the job?
2. How did braille technology users obtain the RBT they utilize on the job?
3. What percentage of braille technology users utilize RBT for specific work tasks?
4. How do the personal and job characteristics of RBT users compare to those of people who do not use RBT at work?

5. What personal factors are associated with RBT use on the job among people who report at least moderate braille skills?

Method

Data Source & Sample

Data were obtained from a longitudinal survey study about the use of assistive technology (AT) in the workplace by people who are blind or have low vision. The purpose of the longitudinal study is to increase knowledge about what AT is used at work, how AT use changes over time, satisfaction with AT, and challenges people experience using AT at work. The overall study consists of collecting data from the same group of participants in four annual surveys conducted between 2021 and 2024. Criteria for participation in the study included: (1) being blind or having low vision, (2) age 21 or older, (3) currently working, either for an organization or self-employed, (4) using AT on the job, (5) planning to work for the next four years, and (6) residing in the United States or Canada. For the present study, we restricted the sample to only include people who reported being totally blind or legally blind, used a computer at work, and used a screen reader to access the computer, resulting in a sample of 304 people. Data from Survey 1 and Survey 2 were used for this study.

Data Collection

Data were collected via Qualtrics in an online survey. Participants were provided the option of completing the survey by phone; 24 people were asked the questions verbally by phone and a research assistant entered their responses in Qualtrics. Data collection occurred in May through September 2021 (Survey 1) and May through September 2022 (Survey 2). Some participants joined the study in 2022; those 48 people completed Survey 1 and Survey 2 questions at the same time, although a few of the Survey 1 questions were changed slightly, as discussed in the Variables section. Most data used for this study came from items in Survey 1.

Variables

Participants were asked to select all AT that they used at work from a list of 28 AT. We identified *braille users* as anyone who reported using a braille device at work, including refreshable braille displays, braille notetaking devices, braillewriters, or braille labeling systems. We identified people as *refreshable braille technology* (RBT) *users* if they reported using a refreshable braille display, a braille notetaker, or both.

Participants were asked how they obtained each AT they use at work in Survey 2. They were provided with a list of five options (see Table 1 for the list) and an "other" option in which they could write in a response. We reported their responses to this item for the two RBTs.

Participants were asked whether they perform 14 specific tasks at work ("Do you perform the following tasks on your job?", e.g., "use a computer to access the Internet, use email or create text documents"). We then asked them to identify which AT they use for the work tasks they perform. In the first implementation of Survey 1, people were able to provide the primary AT and secondary AT they use for the tasks. To streamline the implementation of the combined Survey 1 and Survey 2, we did not ask participants to identify secondary AT used for work tasks. Because people could have selected screen reader or RBT as their primary AT (and the other as a secondary AT), we only used data from the original Survey 1 respondents for these variables. We included 7 of the 14 work tasks for which RBT was utilized for this study, as identified in Table 2. Sample sizes for the work task items vary because people reported performing only some of the tasks for their jobs.

We included nine personal characteristics as independent variables to address RQ 4. *Sex* was a dichotomous variable (Male, Female). *Age* was continuous but was classified into five categories (21-30, 31-40, 41-50, 51-60, 61 or older). *Race* was dichotomized to white race or minority. *Hispanic ethnicity* (yes to "Are you of Hispanic, Latino, or Spanish origin?) and *non-visual disability* (yes if participant reported having one or more other disabilities or chronic health conditions) were also dichotomous variables. *Education level* was grouped into three categories: less than bachelor's degree, bachelor's degree, and graduate or professional degree. *Level of vision loss* included three self-identified categories: totally blind, legally blind with minimal functional vision, legally blind with some functional vision. To determine *age of blindness onset*, participants were asked how old they were when they started having serious difficulty seeing, even with glasses or magnification. Their responses were grouped into three categories: pre-school (age 0-4), K-12 (age 5-18), post-school (age 19 or older). *Braille skill* was based on participants' self-report of their skill level from four possible responses: (1) no braille skills, (2) minimal braille skills, such as using uncontracted Grade One braille, (3) moderate braille skills, such as some use of contracted Grade Two braille, and (4) proficient braille skills, fluent in contracted Grade Two braille.

We included five job characteristics as independent variables to address RQ 4. The first was the *Bureau of Labor Statistics* (BLS) *occupational group* for the person's job. Participants reported their job titles and employers, which was used to identify corresponding BLS job titles and codes. Participants' jobs were in 14 BLS groups, listed in Table 4. Only the 10 groups with 5 or more observations were used in the data analyses. We also categorized the participants' jobs based on whether they were *disability-related* (e.g., including jobs such as TVI, AT instructor, and rehabilitation counselor; National Industries for the Blind jobs; jobs related to accessibility), *employed by a blindness organization* (regardless of type of job), or *government job* (federal, state, or local jobs, including public school teachers). All three variables were dichotomous.

Annual earnings (from all jobs) was a 6-category variable in the survey and was further grouped into three categories, as listed in Table 6.

Data Analyses

We utilized descriptive statistics (frequencies) to address RQs 1, 2, and 3. Chi-square was used to answer RQ 4, and logistic regression was used to address RQ 5. We utilized a subset of the personal characteristic variables, including basic demographic factors and blindness-related factors, for the logistic regression model and included only participants who reported moderate or proficient braille skills (n=244). We sequentially removed non-significant variables from the initial model and used change in Akaike information criterion (AIC) values to determine variables to retain in the final model. Odds ratios provided a measure of effect size.

Results

Demographic information about participants is provided in Table 3. Two-thirds of study participants (66.1%, n = 201) used braille on the job, while just over half (52.6%, n = 160) used RBT on the job. Refreshable braille display use was more common, with 73 people who used only that RBT, 51 people who used both RBTs, and 36 people who only used a braille notetaker. Participants who used braille notetakers on the job mainly purchased the device themselves, whereas participants who used braille displays mainly reported their employers paid or purchased it for them. See Table 1 for full results about how participants obtained their RBT. The percentage of participants who utilized RBT for specific work tasks is presented in Table 2.

Table 1. How refreshable braille technology used on the job was obtained

Method	Braille Notetaker (N = 73)	Braille Display (N = 121)
I purchased it	39.7	23.1
Someone gave it to me	5.5	4.1
Agency or governmental program paid for it	31.5	25.6

Method	**Braille Notetaker (N = 73)**	**Braille Display (N = 121)**
Employer paid/purchased	20.6	44.6
It was free	0.0	0.8
Other	2.7	1.7

Note. All numbers are percentages.

Table 2. Percentage of refreshable braille technology (RBT) users who utilize RBT technology for specific work tasks

Work Task	**N^***	**Use RBT**
Take notes in a meeting	105	82.9
Make formal presentations	73	64.4
Use a computer to access the Internet, use email, or create text documents	112	55.4
Use a computer to access the organization's database or software system	77	42.9
Use a computer to create spreadsheets	81	42.0
Use a computer to create presentations	48	41.7
Use a computer to participate in a meeting	112	35.7

$^*N$ represents the number of RBT users who reported performing the work task.

Tables 3 through 6 present the percentages of RBT and non-RBT users by various personal and job characteristics. RBT use differed significantly for 4 of the 9 personal characteristics, all of which were related to vision loss or other disability (see Table 3). Participants who used RBT were more likely to experience vision loss early in life, have no or limited functional vision, and have higher braille skills. They were less likely to report a non-visual disability. RBT use differed significantly on only 1 of the 5 job characteristics, with people who work for blindness organizations less likely to use RBT. There were some large differences for job occupational group, but these differences did not reach statistical significance.

Table 3. Refreshable braille technology (RBT) use by personal characteristics

Personal Characteristic	***N***	**Sample (%)**	**RBT User (%)**	**Non-RBT User (%)**	χ^2	***df***
Sex					0.03	1
Female	181	59.5	53.0	47.0		
Male	123	40.5	52.0	48.0		
Education level					2.04	2
Less than bachelor’s degree	62	20.4	50.0	50.0		
Bachelor’s degree	114	37.5	57.9	42.1		
Graduate or professional degree	128	42.1	49.2	50.8		
Age category					5.52	4
Age 21 to 30	37	12.2	55.6	44.4		
Age 31 to 40	84	27.6	51.2	48.8		
Age 41 to 50	78	25.7	48.7	51.3		
Age 51 to 60	69	22.7	47.8	52.2		
Age 61 or older	36	11.8	70.3	29.7		
Race					1.39	1
White	251	82.6	54.2	45.8		
Other	53	17.4	45.3	54.7		
Hispanic ethnicity					0.65	1
Yes (Hispanic)	32	10.5	59.4	40.6		
No (Hispanic)	272	89.5	51.8	48.2		
Age of blindness onset					38.32**	2
Pre-school (Age of onset)	202	66.5	64.4	35.6		
K-12 (Age of onset)	52	17.1	40.4	59.6		
Post school (Age of onset)	50	16.5	18.0	82.0		
Level of vision loss					28.90**	2
Totally blind	206	67.8	62.1	37.9		
Legally blind with min. functional vision	73	24.0	39.7	60.3		
Legally blind with some functional vision	25	8.2	12.0	88.0		
Non-visual disability					4.52*	1
Yes (nonvisual disability)	110	36.2	44.6	55.5		
No (nonvisual disability)	194	63.8	57.2	42.8		
Braille skill					108.94**	3

Personal Characteristic	***N***	**Sample (%)**	**RBT User (%)**	**Non-RBT User (%)**	χ^2	***df***
No braille skills	21	6.9	0.0	100.0		
Minimal braille skills	39	12.8	5.1	94.9		
Moderate braille skills	47	15.5	25.5	74.5		
Proficient braille skills	197	64.8	74.1	25.9		

* $p < .05$, ** $p < .0001$

Table 4. Percentage of refreshable braille technology (RBT) users by Bureau of Labor Statistics (BLS) occupation groups.

BLS Occupation Group	***N****	**RBT User**	**Non-RBT User**
Management	58	50.0	50.0
Educational Instruction and Library	47	63.8	36.2
Office and Administrative Support	46	43.5	56.5
Community and Social Service	45	48.9	51.1
Computer and Mathematical	40	57.5	42.5
Business and Financial Operations	19	63.2	36.8
Healthcare Practitioners and Technical	16	37.5	62.5
Arts, Design, Entertainment, Sports, and Media	12	75.0	25.0
Sales and Related	6	16.7	83.3
Legal	5	60.0	40.0
Life, Physical, and Social Service	4	0.0	100.0
Architecture and Engineering	3	66.7	33.3
Protective Service	1	100.0	0.0
Healthcare Support	1	100.0	0.0

Note. Fisher's exact test with Monte Carlo estimation, including only groups with 5 or more observations: $p = .19$.

* *N* represents the number of people who work in each occupation group.

Table 5. Refreshable braille technology (RBT) use by job type

Job Type	N^*	RBT User	Non-RBT User	$\chi^2(1)$
Disability-related	169	51.5	48.5	0.06
Employed by a blindness organization	112	43.8	56.3	5.06*
Government	98	59.2	40.8	3.30

* $p < .05$

* N represents the number of people who work in each job type.

Note. Sample $N = 290$ for Disability-related job and Blindness organization job. Sample $N = 292$ for government job.

Table 6. Refreshable braille technology (RBT) use by annual earnings.

Annual earnings	RBT User (n=149)	Non-RBT User (n=134)	χ^2 (2)
Less than $40,000	35.6	42.5	1.86
$40,000 - $79,999	39.6	38.1	
$80,000 or more	24.8	19.4	

The initial logistic regression model to identify factors associated with RBT use at work included seven independent variables (age [continuous variable], sex, education level, age of onset, vision level, nonvisual disability, and braille skill). The model was statistically significant ($X^2(10, n = 244) = 59.84$, $p <.0001$, however, several variables in the model were not related to RBT use and were removed sequentially to establish the final model. [Insert Table 7 here.]

The final model included five independent variables (see Table 7); the model was statistically significant based on the likelihood ratio test $X^2(6, n = 244) = 57.89$, $p <.0001$, and explained 29% of the total variance in RBT use at work. The Hosmer and Lemeshow goodness-of-fit test of the final model was not significant, $X^2(6, n = 244) = 3.65$, $p = 0.89$, indicating good model fit. Only *braille skill* and *age* were significant predictors of RBT use. Participants who had proficient braille skills were 11.63 times more likely than participants with moderate braille skills to use RBT at work. Odds of using RBT at work increased by 4.2% for each additional year of age. For example, compared to a 30-year-old, a 50-year-old had 2.28 times higher odds of using RBT at work while a 60-year-old had 3.44 times higher odds of using RBT at work.

Discussion

This study provides new information about how many workers who are legally blind use braille and RBT, as well as how they are using RBT on the job. Most people who use braille on the job (in any form) also use RBT – only 13.5% reported only using a braillewriter or braille labeling system. Not surprisingly, a higher percentage of people with more significant visual impairments used braille and RBT on the job. However, as Martiniello et al. discussed, braille could be beneficial to adults with various levels of low vision.

Participants reported a variety of methods for obtaining the RBT that they used on the job. Even though these devices were used on the job, far less than half of them were paid for or purchased by employers. Almost 40% of braille notetaker users reported that they purchased these devices themselves, as did almost a quarter of refreshable braille display users. Since many popular braille notetakers are built on Android operating systems, employers may consider them personal devices, similar to mobile phones or tablets. Braille notetakers are also generally more expensive compared to refreshable braille displays with a similar number of braille cells. Considering the relatively high percentage of participants who purchased RBT on their own, more employers and governmental programs should consider providing RBT for braille users who could utilize this expensive technology on the job.

Although more than half of the sample used RBT on the job, they did not necessarily use it for all work tasks. Only 55.4% to 35.7% of RBT users utilized their devices for specific computer-related tasks. These relatively low rates may be associated with a need for more training in how to effectively utilize RBT. About one-third of RBT users reported that they would benefit from more training on using these devices (McDonnall et al.). Alternatively, it may be that some tasks lend themselves to the need for RBT more than others.

It is surprising that people who work for blindness organizations were less likely to use RBT. This suggests that potentially fewer people who teach or can serve as role models for consumer receiving services use RBT, as compared to other employment settings. Although differences did not reach statistical significance, our results suggest that the use of RBT may be more common, and perhaps more important, in certain occupation groups. Three-quarters of people in the Arts, Design, Entertainment, Sports, and Media group were RBT users. There are a wide variety of jobs in this group, including for example an editor, interpreter, and transcriber who used RBT in our sample. Other BLS occupation groups with higher than average RBT users were Educational Instruction and Library, which included AT instructors in our sample, Business and Financial Operations, and Architecture and Engineering. RBT would be of great value to those in highly technical fields, where it would be beneficial to have access to tactile information when making precise edits to text, such as those exploring programming code for accessibility or software development. Other groups had a much lower percentage of RBT users, such as Sales and Related occupations.

Only disability-related personal characteristics were significantly associated with RBT use in univariate analyses, including age of blindness onset, vision level, having a nonvisual disability, and self-rated braille skill. As expected, braille proficiency was a key factor associated with whether people use RBT on the job. Those with blindness onset prior to school age were much more likely to use RBT on the job, while fewer who experienced vision loss during their schooling and far fewer who experienced vision loss after high school used RBT on the job. This may speak to a need for more opportunity and encouragement for braille instruction for people who lose vision later in life, as research has supported positive outcomes for those who learn braille later (Silverman and Bell).

When utilizing multivariate analyses with people who had proficient or moderate braille skill, only having proficient braille skill and age were significant factors. As age increased, the odds of RBT use increased, but proficient braille skill was clearly the key factor associated with using RBT on the job. If we presume the value of RBT use, then it is important to understand how someone becomes proficient with braille and who is proficient with braille. When age was categorized, it was evident that members of the oldest age group (61 or older) were much more likely to use RBT. Those 61 or older would have completed high school prior to the emergence and availability of personal computers in the 1980s, as well as the availability of the Internet in the 1990s. It is likely that their educational experiences would not have been impacted by the misconception that screen readers and other AT replace the need for braille, including the need for braille instruction and braille support in educational settings.

We should acknowledge limitations to this study, which include that all data is self-reported and our sample may not be representative of all employed people who are legally blind. In addition, we categorized participants into BLS occupation groups based on their job title and employer, but accuracy cannot be certain due to limited details about the jobs.

Conclusion

There has been a complete lack of information about how braille is used at work, and this study helps to fill that gap in research and literature. Our findings indicate that slightly more than half of workers use RBT in some form, but not for all computer tasks. People who lose vision during school age or later in life were less likely to use RBT, likely associated with limited braille skill. Teenagers and adults who experience vision loss should have the opportunity for braille instruction, as well as training with RBT, and this should particularly be emphasized for those who are pursuing employment. People of all ages with less severe visual impairment should also have the option of learning braille. Future research should explore braille proficiency to evaluate who achieves it and how they achieve it, how RBT use benefits workers, and further investigate a potential relationship between RBT use and occupations.

Works Cited

Allman, C. B. “Braille Communication Skills: What teachers teach and visually impaired adults use.” *Journal of Visual Impairment & Blindness*, vol. 92, no. 5, 1998, pp. 331–37.

Bell, Edward C., and Natalia M. Mino. “Blind and visually impaired adult rehabilitation and employment survey: Final results.” *Journal of Blindness Innovation & Research*, vol. 3, no. 1, 2013, pp. 1–35, doi:10.5241/2f1-35.

Bell, Edward C., and Arielle M. Silverman. “Rehabilitation and employment outcomes for adults who are blind or visually impaired: An updated report.” *Journal of Blindness Innovation & Research*, vol. 8, no. 1, 2018, doi:10.5241/8-148.

Bickford, James O., and Ruth A. Falco. “Technology for early braille literacy: Comparison of traditional braille instruction and instruction with an electronic notetaker.” *Journal of Visual Impairment & Blindness*, vol. 106, no. 10, 2012, pp. 679–93, doi:10.1177/0145482x1210601012.

D’Andrea, Frances Mary. “Preferences and practices among students who read braille and use assistive technology.” *Journal of Visual Impairment & Blindness*, vol. 106, no. 10, 2012, pp. 585–96, doi:10.1177/0145482X1210601003.

Emerson, Robert Wall, et al. “Acquisition of literacy skills by young children who are blind: Results from the ABC Braille Study.” *Journal of Visual Impairment & Blindness*, vol. 103, no. 10, 2009, pp. 610–24, doi:10.1177/0145482X0910301005.

Hall, Anthony D., and Slater E. Newman. “Braille learning: Relative importance of seven variables.” *Applied Cognitive Psychology*, vol. 1, no. 2, 1987, pp. 133–41, doi:10.1002/ACP.2350010206.

Hoskin, Elizabeth R., et al. "Effectiveness of technology for braille literacy education for children: A systematic review." *Disability and Rehabilitation: Assistive Technology*, 2022, doi:10.1080/17483107.2022.2070676.

Huebner, K. M. "Daily uses of braille as told by four adult braille users." *Journal of Visual Impairment & Blindness*, vol. 83, no. 6, June 1989, pp. 308–09, doi:10.1177/0145482X8908300612.

Martiniello, Natalina, et al. "Enablers and barriers encountered by working-age and older adults with vision impairment who pursue braille training." *Disability and Rehabilitation*, vol. 44, no. 11, 2022, pp. 2347–62, doi:10.1080/09638288.2020.1833253.

McDonnall, Michele C., et al. "Assistive technology use in the workplace by people with blindness and low vision: Perceived skill level, satisfaction, and challenges." *Assistive Technology*, May 2023, pp. 1–8, doi:10.1080/10400435.2023.2213762.

Rempel, Johan. "Voices from the field: The importance of braille during a pandemic and beyond." *Assistive Technology Outcomes and Benefits*, vol. 16, 2022, www.atia.org/atob.

Ryles, Ruby. "The impact of braille reading skills on employment, income, education, and reading habits." *Journal of Visual Impairment & Blindness*, vol. 90, no. 3, 1996, pp. 219–26, doi:10.1177/0145482X9609000311.

Sheffield, Rebecca M., et al. "How many braille readers? Policy, politics, and perception." *Journal of Visual Impairment & Blindness*, vol. 116, no. 1, 2022, pp. 14–25, doi:10.1177/0145482X211071125.

Silverman, Arielle Michal, and Edward C. Bell. "The association between braille reading history and well-being for blind adults." *Journal of Blindness Innovation & Research*, vol. 8, no. 1, 2018, doi:10.5241/8-141.

Wormsley, D. P. “Braille as the primary literacy medium: General guidelines and strategies.” *Instructional Strategies for Braille Literacy*, edited by D. P. Wormsley and F. M. D’Andrea, AFB Press, 1997, pp. 1–16.

VR for the Psychosocial Support of Orofacial Cleft

Tiffany Thang[1], Achi Mishra[1], Sri Kurniawan[1], Grace Peters[2], Christian Henry[2]
University of California Santa Cruz[1], Smile Train[2]
tthang@ucsc.edu, amishra1@ucsc.edu, skurnia@ucsc.edu, gpeters@smiletrain.org, chenry@smiletrain.org

Abstract

Individuals with orofacial cleft lip and/or palate (CLP) will commonly experience psychosocial issues, including low self-esteem and social avoidance. In this study, we examine the feasibility of a virtual reality (VR) game designed with psychologists and preadolescents with CLP in providing psychosocial support to this population in the interest of understanding whether this technology may help improve access to such care. In collaboration with global cleft organization, Smile Train, the game was developed with psychologists and preadolescents with CLP from their global network of partnering hospitals and clinics using co-design and user-centered design methods. The resulting game was evaluated through a case study where the game was used as a supplement to psychosocial support from a psychologist. Data collected from the study suggests that while there were no significant impacts to psychosocial development, feedback from preadolescent participants and the psychologist suggest that the game may be feasible in helping to promote improvements in psychosocial skills, such as empathy, communication, and self-reflection.

Keywords

Orofacial cleft lip and/or palate, virtual reality, psychosocial support, serious games, social skills.

Introduction

Orofacial cleft lip and/or palate (CLP) is a congenital condition that is commonly associated with a visible facial cleft, speech disorders and hearing deficits (Zeraatkar et al. 14; Wehby and Cassell 3; Crerand et al 2). While physiological symptoms of this condition are often emphasized, individuals with CLP will commonly experience psychosocial issues, including low self-esteem and social avoidance (Stiernman et al. 326; Zeraatkar et al. 14; Stock et al. 187). Individuals with CLP and their families may seek psychosocial support for these issues from psychologists or clinicians, however, this type of support is often inaccessible. Research has found that 57% of families with a child with CLP in the United States live over an hour away from their CLP healthcare clinics, which may not always provide psychosocial support (Kaye et al. 1084). Additionally, the cost of psychosocial support has been reported to cause increased financial strain on individuals with CLP and their families, even with health insurance coverage (Kaye et al. 1084; Rochlin et al.). Research has suggested that the use of VR technologies may be able to promote access to psychosocial support to individuals with disabilities, while maintaining the effectiveness of existing psychological methods of treatment (Gardini et al. 10; Rowland et al. 15). In this study, we examine the feasibility of a virtual reality (VR) game for the psychosocial support of preadolescents with CLP.

Related Work

Research has demonstrated that the use of VR technologies to provide psychosocial support to individuals with psychosocial disorders is promising in improving symptoms of social anxiety and social avoidance (Beidel et al. 1360; Gardini et al. 10). Reviews on VR interventions for the treatment of social anxiety disorder (SAD) have found that they are successful in improving symptoms of these conditions and that treatment outcomes of VR interventions are comparable to current psychological methods of treatment (Beidel et al. 1360, Rowland et al. 15). Existing research also suggests that the use of VR interventions in providing psychosocial support is effective for children with psychosocial disorders, where a study evaluating the use of a VR intervention for children with SAD found improvements in social anxiety across participant self-report, parent-reports, and clinical measures (Beidel et al. 1360). VR for the psychosocial support of individuals with CLP has yet to be thoroughly explored, however, research has evaluated the use of similar technologies, such as augmented reality (AR) and non-immersive simulations to provide support to individuals with CLP, suggesting that the technology is well-received (Lo et al. 6; Schendel et al. 226; Vyas et al. 628).

Based on the literature discussed above, we believe that the development of a VR game for the psychosocial support of individuals with CLP will be feasible in promoting psychosocial development. In our study, we have chosen to focus on designing and developing this game for preadolescents, based on research suggesting that this stage is the most critical for the development of psychosocial skills (Guillén et al. 121).

Methods

User and Expert Interviews

Prior to designing the game 6 preadolescents and adolescents with CLP between the ages of 11 and 17 were interviewed on their experiences with psychosocial development and self-esteem, along with 2 parents of these participants who helped provide further clarity on their child's responses. Interviews were conducted remotely over Zoom and were transcribed by a member of the research team. Thematic analysis was conducted with the interview data, and the following themes were identified in relation to psychosocial development: discomfort in new social environments, aversion to interactions with unfamiliar peers and difficulty accepting one's differences. These themes were evaluated by 4 psychologists with experience working with preadolescents with CLP from Smile Train's network of partnering hospitals and clinics, who were interviewed about their approach to helping them with developing psychosocial skills. On average, these psychologists had 7.75 years of experience in clinical psychology, with a range of 1 to 20 years, specializing in providing psychosocial support to children and adolescents with CLP. Psychologists validated the themes identified in the thematic analysis and provided insight into the methods they used to promote psychosocial development with preadolescent patients with CLP, focusing on modifying negative perceptions of their condition and themselves.

Collaborative Design Procedures and Game Development

Following the interviews, collaborative storyboarding procedures were conducted with 3 psychologists. This was done to further design and develop the game's content and format, using the themes identified from the interview data. Collaborative storyboarding procedures took place over Zoom, using a digital drawing tool and Miro (Miro). Psychologists were asked to review the themes and then reflect on their experiences working with these individuals on developing psychosocial skills. They were then asked to discuss these themes in more detail, elaborating on how they might appear in their psychosocial support sessions.

Psychologists stated that preadolescents with CLP they work with often hyper-focus on their condition, worrying about how others will judge their appearance or speech, causing them to avoid new environments or interacting with new people. Psychologists advise them to focus on traits beyond their CLP so that they can be encouraged to develop and build confidence around these traits. From this, a superhero theme was designed to help promote the idea that patients have "superpowers", or traits, that they can grow and build confidence around in the game. Psychologists also incorporated “Super Pet Pals” into the storyline, such as SuperPup shown in Figure 1, to accompany patients in the game and provide them with moral support.

Fig. 1. SuperPup is shown in the cyberbullying module, where players address mean comments in the form of computer bugs.

Upon further development of the game’s storyline, 3 modules were created encompassing the themes identified from the thematic analysis, with the inclusion of an additional theme addressing cyberbullying. Psychologists noticed an increase in discussions about cyberbullying with preadolescents they worked with during the pandemic, suggesting that this addition was necessary to fully address clients’ psychosocial experiences. The resulting 3 modules focused on addressing discomfort in new social environments, aversion to interactions with unfamiliar peers, and difficulty accepting one’s differences and cyberbullying, which are further described in Table 1. Collaborative storyboarding procedures also resulted in the design of the game’s format, which included homework that preadolescents would complete in real-life to help with generalizing lessons learned in the game. Psychologists also wanted to speak with preadolescents during the game to discuss the social scenario and interactions, leading to the development of a unique UI for psychologists where the game could be paused for this purpose. The psychologist UI is shown in Figure 2, which can only be viewed and interacted with by the psychologist.

Table 1: Game Module Descriptions.

Module Name	**Theme(s) Addressed**	**Storyline Description**
Module A: Social Soccer	Aversion to interactions with unfamiliar peers, discomfort in new social environments	The player is asked to retrieve a precious gemstone from a soccer trophy and must join a tournament to retrieve it. To join the tournament, they must join a team by working up the courage to ask their peers if they can play with them.
Module B: Cafeteria Calculations	Aversion to interactions with unfamiliar peers, discomfort in new social environments	The player meets a student who is new at their school and was asked by their teacher to help collect lunch orders. The new student is too shy to speak with anyone and now has to figure out everyone's lunch order based on their payment amount. The player must help the student and give him advice on what to do when his classmate becomes curious.
Module C: Cyberbully Bugs	Accepting one's differences, cyberbullying	The patient is called to assist with someone's computer overheating, but learns that it is the work of The Mad Scientist, a social media scientist being cyberbullied by her peers. The player must help discourage the bullies from harming The Mad Scientist with mean comments and give advice to The Mad Scientist.

Fig. 2. The psychologist UI is shown with the option to progress the game according to whether the patient advises the character to lie or tell the truth in this social scenario.

A task analysis was also conducted to further understand how the game could be incorporated into existing psychosocial support sessions led by psychologists. The task analysis focused on understanding how sessions were structured and how the game may be incorporated into these sessions. It was found that they prioritized having clients reflect on their thoughts and behaviors with respect to a social situation, which led to the design of a branching narrative game format, allowing clients to reflect on how their thoughts and behaviors influence their actions and how their actions may result in a specific social consequence.

With the data from these procedures, the game was then developed using the Unity game engine for the Meta Quest 2 VR headset. Multiplayer capabilities were also incorporated into the game to allow for psychologists and preadolescents with CLP to play together, which was implemented using Photon PUN 2, a multiplayer networking package.

Case Study

A case study was conducted to evaluate the impact of the game on the development of psychosocial skills for preadolescents with CLP. A pre-test/post-test user study was designed to determine whether playing the game as a supplement to psychosocial support sessions with a psychologist would be beneficial. Participants of the study were asked to work with the psychologist once a week over the course of 4 weeks on developing psychosocial skills using the game. Psychosocial development was measured using the Need to Belong (NTB) scale (Leary 624), Rosenberg Self-Esteem (RSE) scale (Rosenberg), and the Behavior Assessment System for Children, 3rd Edition (BASC-3) (Merenda). These measures were chosen by researchers and the psychologist participating in the case study to measure changes in social behaviors or self-esteem in relation to psychosocial development. Qualitative feedback surveys were also given to participants to understand their experiences with the game and contained open-ended questions regarding their experiences relating to and understanding the scenarios presented in each module.

Results

There were 6 preadolescents with CLP who participated in the study, with an age range of 11 to 13, with an average age of 11.8. There were 4 females and 2 males, all with a current diagnosis of CLP and currently participating in regular psychosocial support sessions with a psychologist. Results from the pre-test and post-test BASC-3 are presented in Table 2, where several factors that contribute to the social behaviors and relationships of preadolescent participants are presented. For each of the factors measured by the BASC-3, a greater score indicates that the individual is more likely to exhibit that behavior or factor. For each factor, a paired t-test with a statistical significance threshold of $p<0.05$ was used. While there was no significant effect for any of the factors measured, it is noted that there was a decrease in average scores for atypicality, anxiety, feeling inadequate and social stress after playing the game.

Table 2. BASC-3 Pre- and Post-test Results.

Behavioral and emotional status	**Average pre-test value**	**Average post-test value**	**t-value**	**p-value**
Atypicality	63.3	49.5	t(5)= 0.80	p= 0.46
Anxiety	40.2	36.7	t(5)= 0.75	p= 0.49
Depression	34.2	46.7	t(5)= -1.06	p= 0.34
Feeling Inadequate	60.8	47.5	t(5)= 0.90	p= 0.41
Interpersonal Relationships	78.3	68.5	t(5)= 0.84	p= 0.44
Locus of Control	36.0	41.8	t(5)= -0.74	p= 0.50
Negative Attitudes Towards School	77.3	71.7	t(5)= 0.89	p= 0.41
Negative Attitudes Towards Teachers	44.0	47.8	t(5)= -0.29	p= 0.78
Parent Relationships	72.8	61.7	t(5)= 0.68	p= 0.53
Self-Confidence	49.7	45.5	t(5)= 0.25	p= 0.81
Self-Esteem	73.2	64.8	t(5)= 0.44	p= 0.68
Social Stress	46.7	43.3	t(5)= 0.31	p= 0.77

For the NTB and RSE scales, there was also no significant difference in scores from the pre and post-test when performing a paired t-test with a statistical significance threshold of $p<0.05$. The results for the NTB scale are as follows, ($M=-0.3$); $t(5)=-0.12$, $p=0.91$. It is worth noting that the negative mean difference denotes that the feeling of needing to belong decreases after playing the game. The results for the RSE scale are as follows, ($M=1.2$); $t(5)=0.82$, $p=0.45$.

Qualitative data collected from the qualitative feedback surveys suggest that preadolescent participants and the psychologist found the game useful in promoting empathy, as well as self-reflection and communication. They all mentioned communication as a key skill learned through the game, where they discussed learning and practicing how to communicate with others in a friendly manner. One participant described their experience with the scenarios presented in the game saying, “some of the scenarios I have lived and others I can live”, indicating that they were able to not only relate them, but could imagine experiencing the scenarios in the game in the future. Preadolescent participants generally enjoyed playing the game and felt that the modules were relatable and would help them prepare for similar scenarios in real life.

A qualitative feedback survey was also given to the psychologist to understand their experience using the game with preadolescents with CLP as a part of their psychosocial support sessions. The psychologist’s overall thoughts on the game were that it was a “really good tool to support therapy processes”, where the role-playing aspect of the game allowed patients to receive practice while being able to work with the psychologist during the process. They also expressed that their clients “could really relate [to the module] and [that] their answers show … a reflection of what they’ve been through”. In terms of impact of the game, the psychologist felt that the game encouraged “a more open conversation” with clients, where they would more readily talk about whether they had experienced the scenario in the game and how they felt about it. While the quantitative data presented no significant changes in psychosocial development or behavior, the psychologist did mention that their clients seem to be more reflective during sessions with the VR game. They stated that their clients “showed more reflection of what they went through … right after [playing] the module”.

The psychologist also suggested minor modifications to the game to help promote psychosocial development for preadolescents with CLP, including further personalization of the scenarios in the game modules. When asked to reflect on their overall experience using the game as a supplemental tool to their psychosocial support sessions with preadolescents with CLP, they stated that the game "is a great compliment to therapy … especially because kids are more into technology … sometimes it's easier for them to feel comfortable if that's the medium [they] use to do therapy … it's a great way to be closer to patients".

Discussion

Based on qualitative data from the preadolescent participants and the psychologist, it is evident that the developed VR game may be feasible in promoting psychosocial development in preadolescents with CLP. Participants found the game to be helpful in learning and practicing communication skills, as well as in understanding how they might approach challenging social situations they are experiencing or may encounter in the future. The psychologist participating in the study found that clients became more reflective of their similar real-world experiences after playing the game and felt that it was a "good tool to support therapy processes", indicating that VR games may be beneficial as a method of providing supplemental psychosocial support to preadolescents with CLP. This effect may be attributed to user-centered and co-design methods used to design the game, where the incorporation of data from user and expert interviews, as well as the collaborative storyboarding procedures and task analysis with psychologists, allowed for the creation of more realistic social scenarios in the game that preadolescents with CLP could easily relate to and identify with.

However, quantitative data collected from this case study suggests that the game has yet to have a significant impact on psychosocial skills for this population, which may be due to the study's small sample size (n=6). This effect may also reflect upon the short amount of time the game was used with patients, where observations conducted over a longer period of time may yield more significant results. Additionally, the measures used to evaluate the impact of the game on psychosocial development could be improved, where future research may benefit from the creation of specialized quantitative measures that directly evaluate the impact of the VR game based on participants' specific psychosocial needs.

Overall, the information presented by the case study is sufficient to encourage further studies regarding the use of serious VR games to assist with the development of psychosocial skills for preadolescents with CLP. The use of user-centered and co-design methods with preadolescents with CLP and psychologists should continue to be explored, where future studies should aim to understand whether this technology may be effective in promoting psychosocial development as not only a supplement to existing psychosocial support methods, but as a standalone resource.

Conclusion

This study presents the design and evaluation of a serious VR game developed with preadolescents with CLP and psychologists to understand its impact on promoting the development of psychosocial skills for preadolescents with CLP. Our findings suggest that the use of serious VR games may be feasible as a supplement to existing psychosocial support for this population. Qualitative data collected from participants show that the game was relatable and useful for preadolescents with CLP, encouraging greater expression of empathy and reflection. They also easily identified with the scenarios presented and could relate them to their real-life experiences, suggesting that they may be able to easily translate skills learned from the game to similar real-life situations. Future research on this topic should aim to further evaluate the impact of serious VR games on the psychosocial development of preadolescents with CLP by continuing the process of designing these technologies with stakeholders, such as psychologists and individuals with CLP. Consideration should also be given to the factors of VR games that may contribute to the development of psychosocial skills, so that the field may seek to improve upon these features to further benefit its use for those in need of support.

Works Cited

Beidel, Deborah C., et al. “Treating childhood social anxiety disorder with virtual environments and serious games: A randomized trial.” *Behavior Therapy*, vol. 52, no. 6, Nov. 2021, pp. 1351–63. *ScienceDirect*, doi.org/10.1016/j.beth.2021.03.003.

Crerand, Canice E., et al. “The Americleft psychosocial outcomes project: A multicenter approach to advancing psychosocial outcomes for youth with cleft lip and palate.” Clinical Practice in Pediatric Psychology, vol. 5, 2017, pp. 65–76. APA PsycNet, doi.org/10.1037/cpp0000172.

Gardini, Valentina, et al. Journal of Clinical Medicine, vol. 11, no. 21, Oct. 2022, p. 6463. PubMed Central, doi.org/10.3390/jcm11216463.

Guillén, Ana Ruiz, et al. “Psychological variables in children and adolescents with cleft lip and/or palate.” *Journal of Clinical Pediatric Dentistry*, vol. 44, no. 2, Jan. 2020, pp. 116–22. *Silverchair*, doi.org/10.17796/1053-4625-44.2.9.

Kaye, Alison, et al. “Assessment and management of psychosocial needs: Social work utilization in comprehensive cleft team care.” *The Cleft Palate-Craniofacial Journal*, vol. 55, no. 8, Sept. 2018, pp. 1081–91. *SAGE Journals*, doi.org/10.1177/1055665618760620.

Leary, Mark R., et al. “Construct validity of the need to belong scale: Mapping the nomological network.” *Journal of Personality Assessment*, vol. 95, no. 6, Nov. 2013, pp. 610–24. *Taylor and Francis+NEJM*, doi.org/10.1080/00223891.2013.819511.

Lo, Steven John, et al. “The cleft lip education with augmented reality (CLEAR) VR phase 2 trial: A pilot randomized crossover trial of a novel patient information leaflet.” *The Cleft Palate Craniofacial Journal*, Jan. 2022, p. 10556656211059708. *SAGE Journals*, doi.org/10.1177/10556656211059709.

Merenda, Peter F. “BASC: Behavior assessment system for children.” Measurement and Evaluation in Counseling and Development, vol. 28, no. 4, 1996, pp.229-232.

“Miro.” *Miro,* 2 Nov. 2023, https://miro.com/.

Rochlin, Danielle H., et al. “Out-of-pocket costs and provider payments in cleft lip and palate repair.” *Annals of Plastic Surgery*, Feb. 2022. *PubMed*, doi.org/10.1097/SAP.0000000000003081.

Rosenberg, Morris. *Society and the Adolescent Self-image*. Princeton University Press, 1965.

Rowland, Dale P., et al. “A decade in review: A systematic review of virtual reality interventions for emotional disorders.” Psychosocial Intervention, vol. 31, no. 1, Dec. 2021, pp. 1–20. DOI.org (Crossref), doi.org/10.5093/pi2021a8.

Schendel, Stephen, et al. “A surgical simulator for planning and performing repair of cleft lips.” *Journal of Cranio-Maxillofacial Surgery*, vol. 33, no. 4, Aug. 2005, pp. 223–28. *ScienceDirect*, doi.org/10.1016/j.jcms.2005.05.002.

Stiernman, Mia, et al. “Parental and health care professional views on psychosocial and educational outcomes in patients with cleft lip and/or cleft palate.” European Journal of Plastic Surgery, vol. 42, no. 4, Aug. 2019, pp. 325–36. Springer Link, doi.org/10.1007/s00238-019-01530-0.

Stock, Nicola Marie, et al. “Promoting psychosocial adjustment in individuals born with cleft lip and/or palate and their families: Current clinical practice in the United Kingdom.” The Cleft Palate-Craniofacial Journal, vol. 57, no. 2, Feb. 2020, pp. 186–97. SAGE Journals, doi.org/10.1177/1055665619868331.

Vyas, Raj M., et al. “Using virtual augmented reality to remotely proctor overseas surgical outreach: Building long-term international capacity and sustainability.” *Plastic & Reconstructive Surgery*, vol. 146, no. 5, Nov. 2020, pp. 622e–9e. *DOI.org (Crossref)*, doi.org/10.1097/PRS.0000000000007293.

Wehby, Gl, and Ch Cassell. “The impact of orofacial clefts on quality of life and healthcare use and costs.” Oral Diseases, vol. 16, no. 1, 2010, pp. 3–10. Wiley Online Library, doi.org/10.1111/j.1601-0825.2009.01588.x.

Zeraatkar, Maryam, et al. “A qualitative study of children’s quality of life in the context of living with cleft lip and palate.” Pediatric Health, Medicine and Therapeutics, vol. 10, Jan. 2019, pp. 13–20. www.dovepress.com, doi.org/10.2147/PHMT.S173070.

Low-Cost Wayfinder Using Bluetooth Angle-of-Arrival

Brandon Cai
Independent Researcher
brandoncai888@gmail.com

Abstract

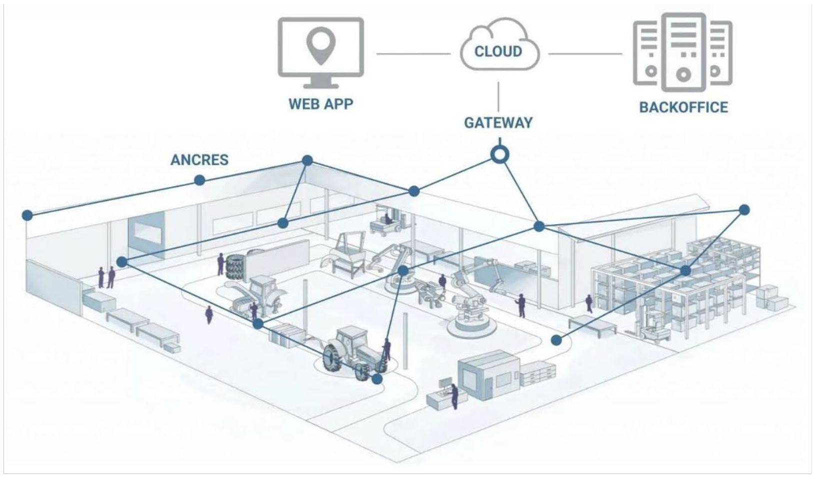
A fully functional location-tag based point-to-point wayfinder prototype system was created using three Bluetooth 5.1 boards and a program developed in Python. This new wayfinder is designed to help visually impaired people navigate public facilities such as airports and shopping centers. To minimize deployment cost and promote wide adoption, this point-to-point Angle-of-Arrival (AoA) architecture concept moved the antenna array onto the user device, allowing each beacon to mark a point of interest and be reduced to a simple single-antenna transmitting beacon without the need for computation and measurement capabilities. As proof of concept, this reversal of the whole premise mesh Indoor Positioning System (IPS) configuration opens the door to a low-cost wayfinder without a mesh network. This wayfinder prototype demonstrated directional angle accuracy of 10° within a 15 meters distance with merely 3mW transmitting power, and a radar-like graphical interface to display direction derived from AoA and the distance to the target derived from the Received Signal Strength Indicator (RSSI). For more seriously visually impaired users, a voice notification system can be built as an extension. With future smart phones and tablets with multiple Bluetooth antennas built in, this proposed wayfinder can be deployed as an app.

Keywords

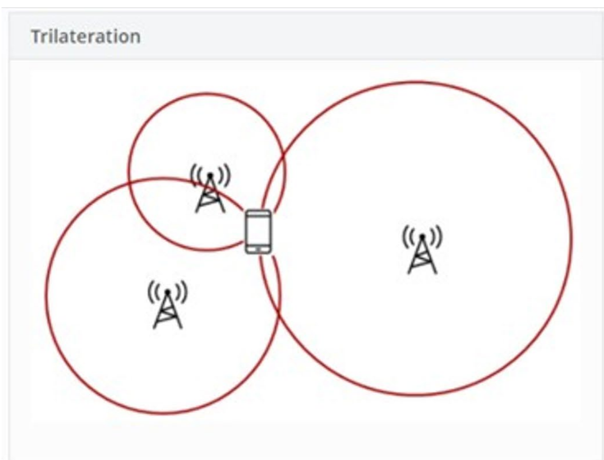
Angle of Arrival (AoA); wayfinder; Bluetooth beacon; indoor positioning; target angle

Introduction

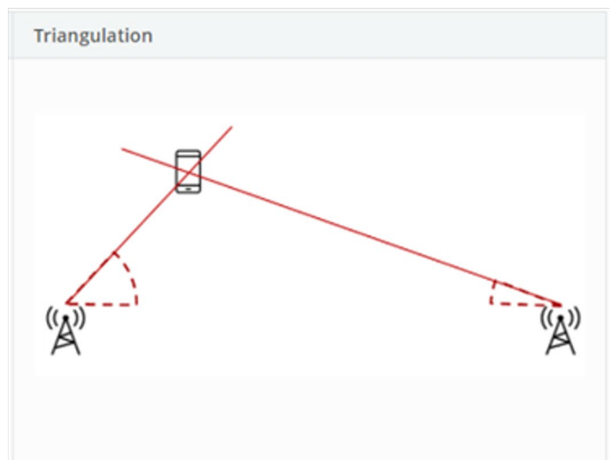
First generation IPS calculates location through distance trilateration by finding the intersection of the three circles with measured radii to centers at three nearby anchor nodes with known positions (Lee) as shown in Fig.1 (A) and (B). This requires a mesh network of Bluetooth beacons at about 10-meter intervals as location reference points throughout an entire premise (Solovev). In the new 2019 Bluetooth version 5.1, the Angle of Arrival (AoA) feature was introduced for direction measurement with the intention of improving IPS accuracy (Cominelli). An AoA beacon consists of an antenna array with two or more antennas. The transmitter's direction can be derived by detecting the received signals' phase difference. As illustrated in Fig.1 (C), to determine the actual location, two AoA antenna arrays at known anchor locations are needed for triangulation. Therefore, this new AoA based IPS system would require a similar beacon mesh network with additional antennas and computational ability for each beacon.



(A) IPS mesh network



(B) Trilateration



(C) Triangulation

Fig. 1. (A) IPS mesh network of anchors provide reference points. (B) Original IPS system calculates location by distance trilateration with 3~5m accuracy. (C) The improved AoA triangulation method based on two antenna arrays can improve the accuracy to 1-2m accuracy. (Pierre; “RTLS Toolbox”).

When AoA is used, the IPS accuracy will be dependent on angle measurement accuracy instead of distance measurement accuracy, but the extent of improvement achievable in real world IPS accuracy is still not well documented. A 2019 report showed a 0.85-meter accuracy in 95% of cases using a test bed built with Software Defined Radio to emulate the AoA function before any Bluetooth 5.1 AoA hardware was commercially available (Cominelli). The test was performed in a 25x21-meter open space. Another published accuracy analysis achieved 0.6-meters but with the target placed only 2 meters away (Andersson) compared to a real world distance up to 10-15 meters. In two recent 2022-2023 publications from the same research team, an AoA angle accuracy study was performed indoors using 2 orthogonal antenna arrays (Girolami; Mavilia). The angle measurement accuracy was reported in Maximum Absolute Error (MAE) as around 10°. Since this research setup used only a single anchor location, the result is a good indicator for AoA angle measurement accuracy for a real world IPS system, but not an apples-to-apples comparison to IPS localization accuracy with multiple AoA anchors.

In summary, AoA based IPS can offer higher accuracy down to about 1 to 2 meters. However, with the added complexity and cost of AoA antenna arrays at each anchor point, this AoA IPS system will be even more expensive to deploy, and that is not feasible for the cost-sensitive assistive technology sector for public service of the visually impaired community. In this paper, a new low-cost alternative is presented.

Experiments and Discussion

Proposed Solution: A Point-to-Point AoA based Wayfinder

To avoid an expensive mesh system, this paper presents a point-to-point wayfinder which leads users to a target beacon of their choice. As illustrated in Fig.2, the wayfinder derives the direction bearing using AoA, while the distance is estimated using traditional Received Signal Strength Indicator (RSSI). In essence, this point-to-point AoA wayfinder is a reversal of the AoA based IPS system. Whereas the typical IPS configuration requires every mesh point marker to hold an antenna array and actively calculate and communicate angle data to the user, the proposed wayfinder moves the required antenna array onto the user device for angle measurement. In this wayfinder setup, each beacon marking a point of interest can be just a simple single-antenna beacon transmitter without any computation and measurement requirements.

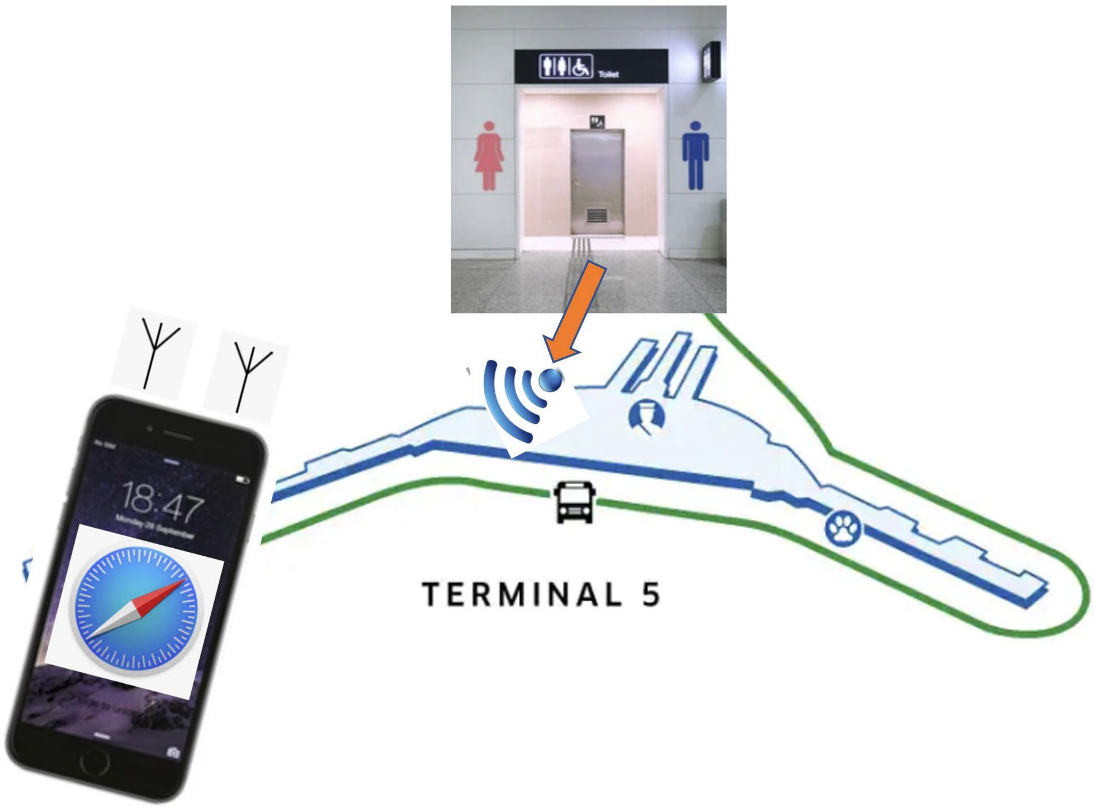


Fig. 2. Proposed concept of a low-cost wayfinder design using point-to-point Bluetooth AoA to avoid the high infrastructure cost of a full-premise mesh network required in an IPS system.

The simple beacon for this architecture is just a <$30 low-cost broadcaster. On top of the reduction in beacon cost, the complete mesh network is reduced to only selected points of interest. In this way, the number of beacons required is greatly reduced as well, leading to a more affordable initial deployment cost to encourage wider adoption. The deployment of the wayfinder system is also incrementally scalable, allowing a quick initial deployment with a few most needed public facilities first, such as airport gates or restrooms, followed by the addition of other destinations such as shops and restaurants in future expansion.

Designing and Building the Wayfinder Prototype

As shown in Fig.3 for target angle measurement, a wave front will reach Antenna 1 first and travel an extra distance of $d \cdot cos(\theta)$ to reach the Antenna 2, causing a phase delay of:

$$\Delta\varphi = \frac{d \cdot \cos(\theta)}{\lambda} 2\pi$$

d = distance between antennas (known)

θ = AoA angle (want to measure)

λ = wavelength (known)

Δφ = phase difference between antennas (measured)

In this application, the phase difference $\Delta\varphi$ is measured to calculate the AoA direction angle θ:

$$\theta = arccos\left(\frac{\lambda \cdot \Delta\varphi}{2\pi d}\right)$$

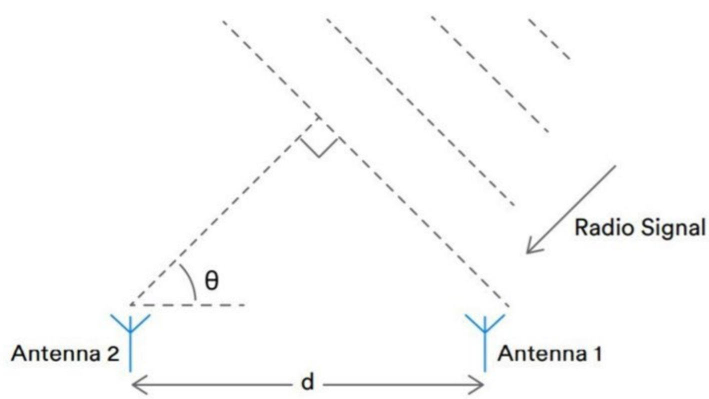


Fig. 3. Conceptual illustration of AoA angle calculation using the wave front phase difference between Antenna 1 and Antenna 2 (Armas).

Target distance measurement is based on the RSSI of the received signal strength and energy conservation, as shown in Fig.4. The power density p is given by:

$$p = \frac{\boldsymbol{P_T}}{4\pi R^2}$$

$\boldsymbol{P_T}$ = *total power transmitted from the beacon antenna*

R = *distance from the transmit antenna*

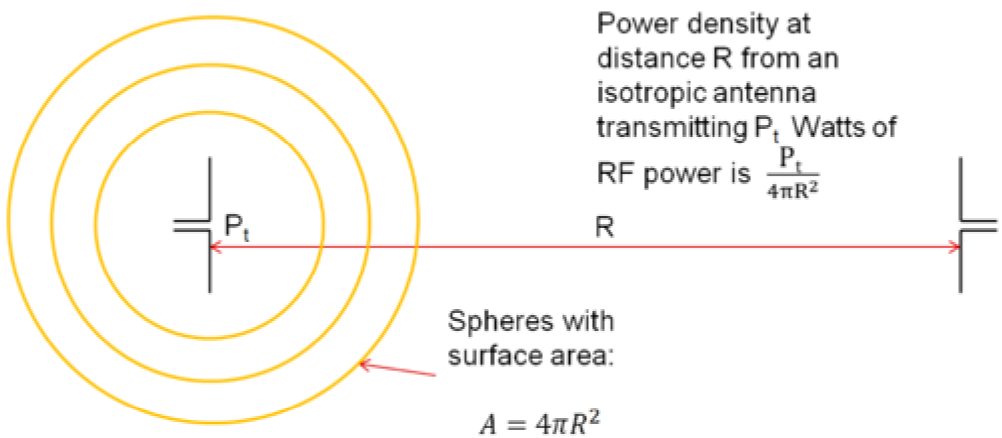


Fig. 4. Conceptual illustration of distance measurement using Received Signal Strength Indicator (RSSI) ("RTLS Toolbox").

A picture of the whole wayfinder prototype system is shown in Fig.5. The hardware of the wayfinder is constructed with three TI Bluetooth 5.1 development kits (LaunchPad CC26x2R1). For evaluation, a rotating camera tripod is used to incrementally sweep the controlled target bearing angle. To prevent the metal in the tripod from creating reflections and interfering with the signal transmission and reception, the Bluetooth circuit boards were elevated 10cm above the metal piece using 3D printed plastic mounting brackets. The 3D design of the brackets was completed using the TinkerCAD and 3D printed with PLA (polylactic acid).

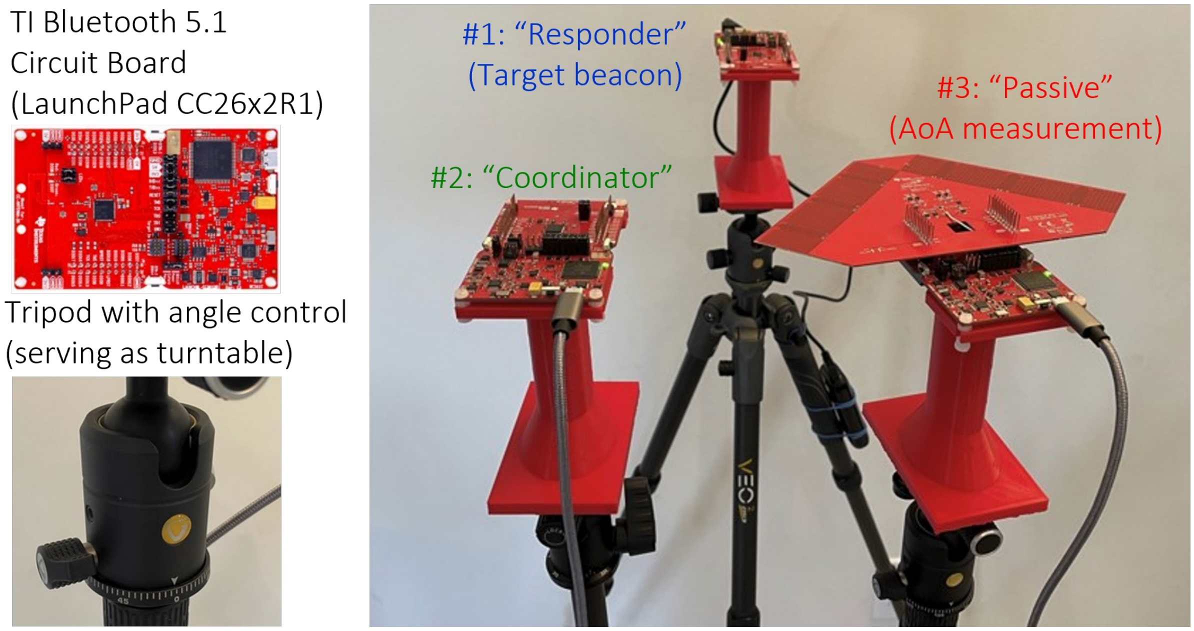


Fig. 5. The whole wayfinder prototype system with three TI Bluetooth 5.1 LaunchPad CC26x2R1 boards, rotating camera tripods, and 3D printed mounting brackets.

Figure 6 zooms in on the wayfinder prototype system components: antenna arrays and 3D printed mounting brackets. The antennas on the TI BOOSTXL-AoA are dipole antennas built-in to the Printed Circuit Board (PCB) with etched traces. The ¼ λ dipole antennas are 30.8 mm long for operating at the center of the 2.4GHz Bluetooth Low Energy frequency spectrum. The phase centers are spaced 35mm apart to meet the < ½ λ spacing requirement to make the maximum phase difference less than 180° to avoid ambiguity in phase measurement. The phase difference range will always fall from 0° to 180° (or –180°) (“Bluetooth® Angle of Arrival Antenna Design.”). The AoA function requires two 30.8mm dipole antennas with 35mm spacing for the 2.4GHz frequency band, which will occupy only a 66mm wide edge that fits most smart phone’s form factor.

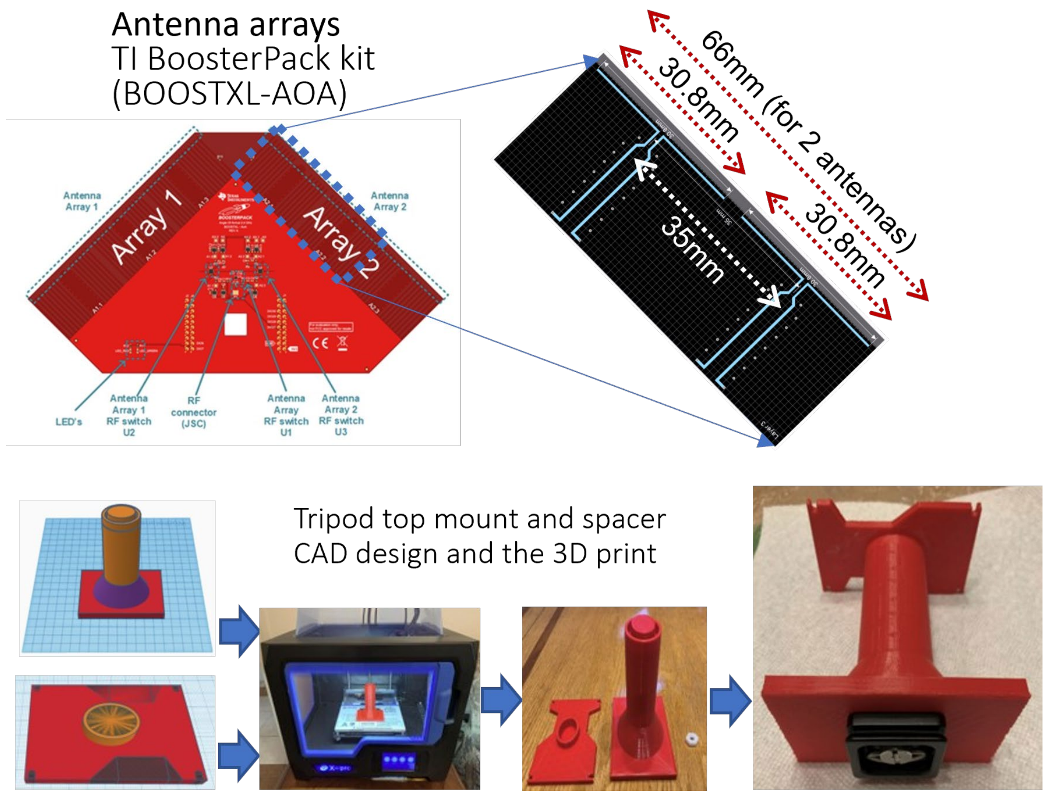


Fig. 6. Close-ups of the wayfinder prototype system components: antenna arrays and 3D printed mounting brackets. AoA function requires at least two antennas occupying a 66mm wide edge that fits within most smart phones' form factor for integration.

The software driver for Real Time Location Service (RTLS) is installed to configure the 3 LaunchPads in different modes ("RTLS Toolbox"). Board #1 is configured as the RTLS "Responder" to act as the target beacon. The Responder responds upon request to transmit the sine wave required for AoA measurement and is placed at the target site such as a bathroom. The "Responder" is battery powered and stand-alone at the target site, not connected to a PC.

Board #2 is configured as RTLS "Coordinator" responsible for discovering available responders in range and establishing connections to them. It sends a request to the selected target responder asking it to transmit the sine wave to Board #3 for AoA measurement.

The most complicated unit is Board #3 with an antenna array installed and software configured as RTLS “Passive”. It passively receives signals from the “Responder” from multiple antennas and calculates the AoA from phase difference. A TI AoA BoosterPack with 2 orthogonal linear antenna arrays is used. Each array consists of 3 antennas (“Bluetooth® Angle of Arrival Antenna Design.”).

The Python program, run on a laptop PC, was written to communicate with the “Coordinator” and “Passive”. The laptop USB port numbers are used to address its instructions and read/write data to the “Master/Coordinator” and “Slave/Passive” respectively. Fig.7 shows an example of the captured AoA waveform in a complex IQ data format that was post processed in Excel to demonstrate that the sampled signal captured had the correct sinusoidal shape. Because of the IQ format, the phase difference can be calculated by multiplying by the conjugate instead of a direct linear subtraction of the phases.

$$e^{i(\phi_x)} * e^{i(-\phi_y)} = e^{i(\phi_x - \phi_y)}$$

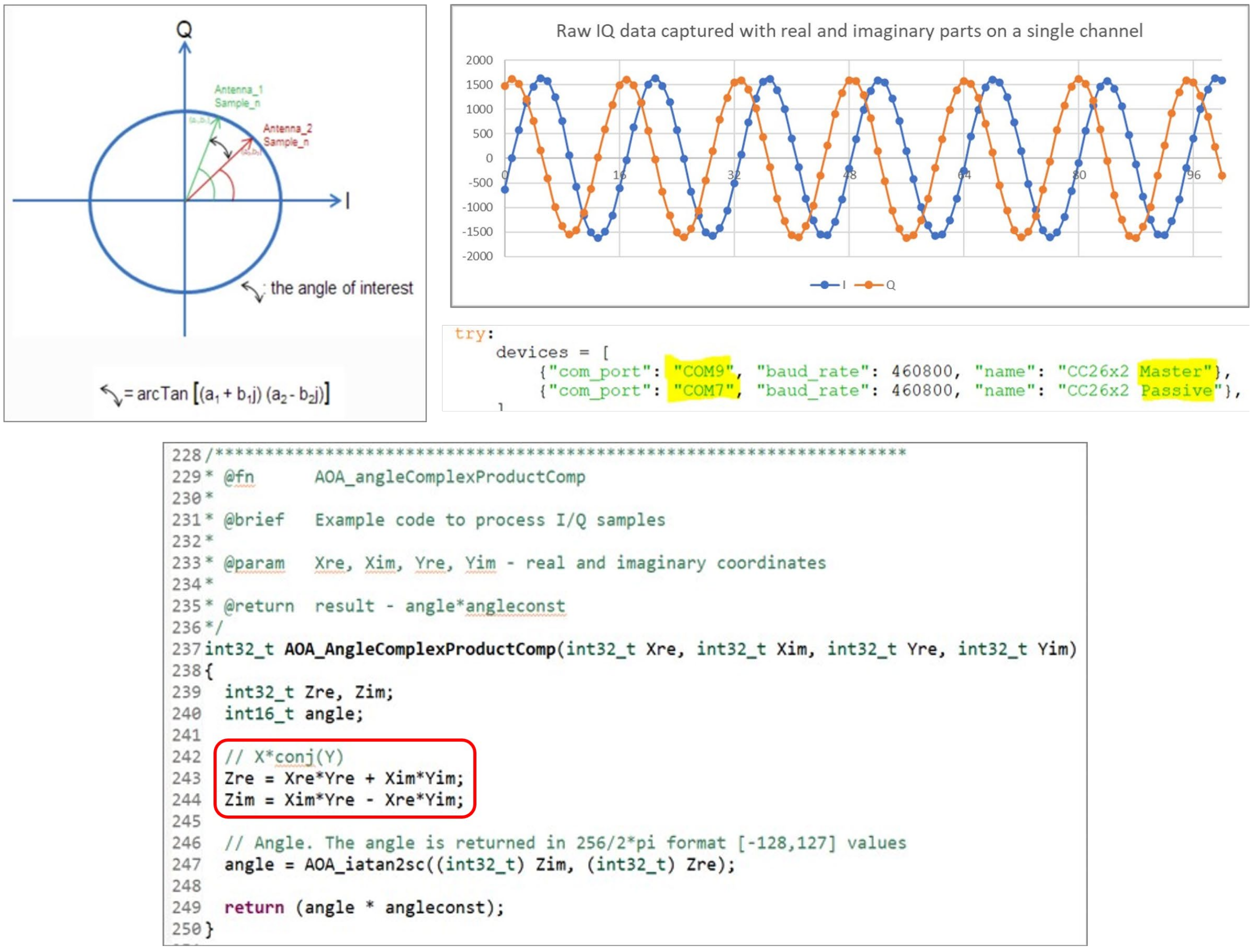
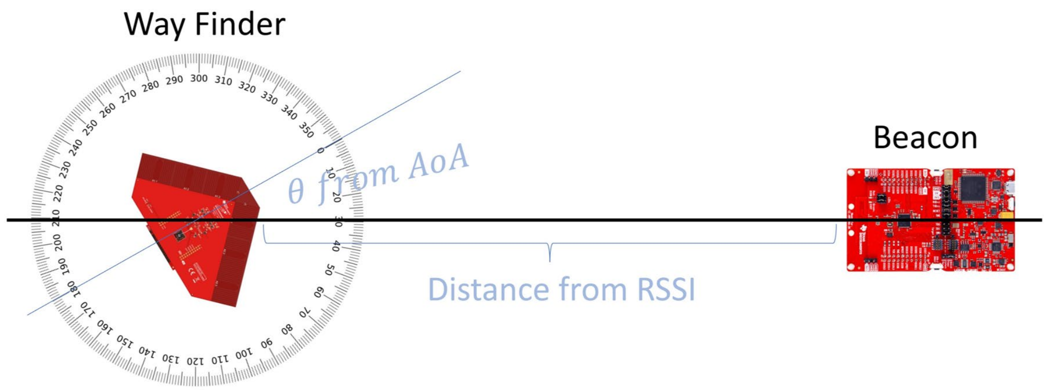


Fig. 7. An example of the captured AoA waveform in a complex IQ Data format, which contains the phase information. Phase difference can be extracted by multiplying by the conjugate instead of linear subtraction.

Wayfinder Accuracy Evaluation Test Results

As shown in Fig.8, to evaluate the target angle measurement accuracy of the prototype system, a 5° increment scale was attached to the rotating camera tripod to gauge the angle sweep. As shown in Fig.8, the antenna array mounted on top of the “Passive” unit is rotated instead of moving the actual target beacon “Responder”, allowing the angle and distance control to be more precise, mitigating the error in moving the hardware.




```
    def _message_receiver(self):
        avg = 0
        n = 0
■ ■ ■
# BJC added code for extract AoA angle data ONLY, and simplify the onscreen data parsing
#-----------------------------------------------------------------------------------------
                if msg.command in ["RTLS_CMD_AOA_RESULT_ANGLE",
                                   "RTLS_CMD_AOA_RESULT_RAW",
                                   "RTLS_CMD_AOA_RESULT_PAIR_ANGLES"] and msg.type == "AsyncReq" \
                                   and sending_node.name != "CC26x2 Master": #if master is the name it is removed
                    self.aoa_results_queue.put(
                        # "name": sending_node.name, # got rid of extra data
                        # "type": str(msg.command),  # got rid of extra data
                        # "identifier": identifier,  # got rid of extra data
                        # "payload": msg.payload
                         msg.payload #this doesn't actually print, commented out in main
                    )

                    avg += msg.payload["angle"]
                    n += 1
                    print(str(msg.payload["angle"]) + "  Average: " + str(round((avg / n),2)))

                    #setup for first use v
                    try:
                        f = open(filename, "r")
                    except FileNotFoundError:
                        f = open(filename, "w")
                        f.write("target, angle, rssi, antenna, channel\n")
                    finally:
                        f.close
                    #set up for first use ^

                    with open(filename, "a") as f:
                        f.write(str(target_angle) + ", " + str(msg.payload["angle"]) + ", " + str(msg.payload["rssi"]) \
                                + ", " + str(msg.payload["antenna"]) + ", " + str(msg.payload["channel"]) + "\n")
#-----------------------------------------------------------------------------------------
```

Create a class to customize message receiver

Filter out non AoA related message content

Running Average AoA from payload

Open text file for data storage

Store only relevant data in file

Fig. 8. The wayfinder accuracy evaluation setup and sample code in Python for data collection and processing.

The accuracy evaluation varies the angle in 22.5° increments (16 settings) at four target distances (1m, 5m, 10m, 15m), recording the AoA and RSSI values and performing an average of 20 consecutive measurements for noise reduction. This results in a very large set of data with >4000 points from the multiple permutations of conditions. Therefore, it is necessary to automate the data collection and processing for the experiment.

The data collection code was built in Python by leveraging the existing TI Development Kit driver debug code reporting the commands, statuses, and data through a message receiver class defined in Python. The added wayfinder Python code extracts and displays only the desired AoA and RSSI data from the message streams and performs an average of 20 consecutive data points collected for each measurement. The Python code also provides an option to store the data in a text file for post processing.

As shown in Fig.9, the Bluetooth AoA wayfinder successfully detected the target within 15 meters with <10° error except a larger error at -90°. The worst-case error is at -90° on the left side of the graph when Antenna Array #1 is orthogonal to the target and blocked by Array #2. A slight AoA accuracy reduction with increasing distance was observed, but it is still adequate for the wayfinder to provide walking directions.

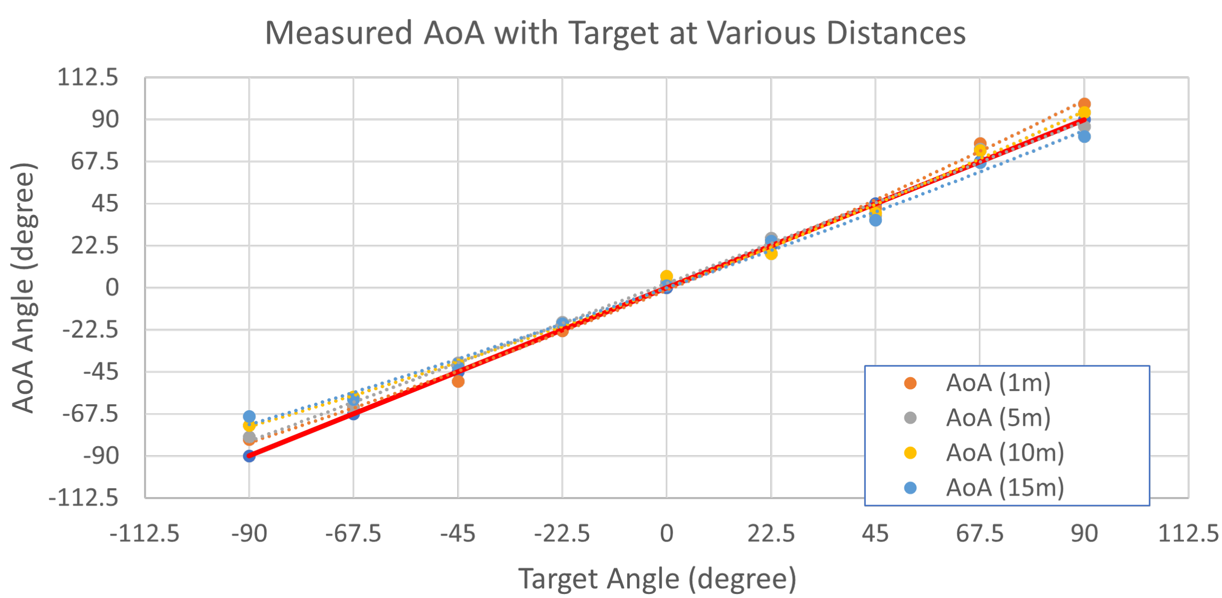


Fig. 9. The AoA measurement accuracy with Target in front (-90° to 90° range) at various target distance (1m, 5m, 10m and 15m).

However, when the target is behind the antenna (90° to 270°), the phase difference measured converts to an angle as if the target is in front of the antenna. As explained in Fig.10, this front-and-back ambiguity is observed for AoA using only a linear antenna array. The ambiguity condition can be removed by using two orthogonal antenna arrays or blocking the signal from a target when it is behind the user. As shown in Fig.11, when Array #1 reports an AoA of 53°, it could be at either 53° or 180°-53°=127°. However, the Array #2's ambiguity reports either 53° or - 53°. So, the combined info from orthogonal linear arrays can determine a unique solution of 53°. Python code to enable Array #2 has been demonstrated. The Python code that can combine the two orthogonal measurements is still in development to reduce the deadtime involving in a required system reset when switching between the two antenna arrays.

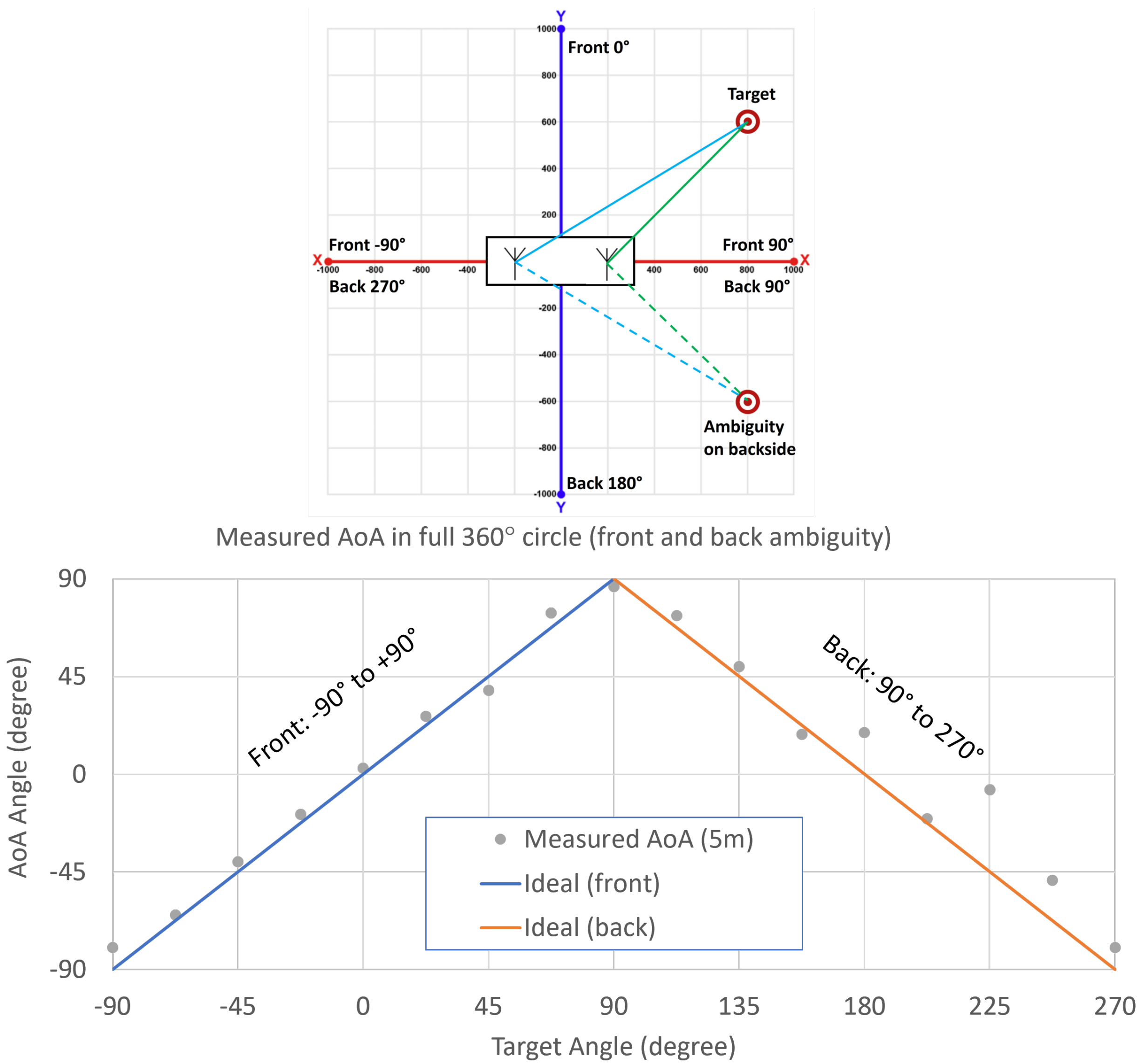


Fig. 10. The front and back ambiguity is observed for AoA using only a linear antenna array.

```
## Start angle of arrival feature
if aoa:
    if rtlsUtil.is_aoa_supported(all_nodes):
        aoa_params = {
            "aoa_run_mode": "AOA_MODE_ANGLE",  ## AOA_MODE_ANGLE, AOA_MODE_PAIR_ANGLES, AOA_MODE_RAW
            "aoa_cc26x2": {
                "aoa_slot_durations": 1,
                "aoa_sample_rate": 1,
                "aoa_sample_size": 1,
                "aoa_sampling_control": int('0x20', 16),
                ## bit 0   - 0x00 - default filtering, 0x01 - RAW_RF no filtering,
                ## bit 4,5 - default: 0x10 - ONLY_ANT_1, optional: 0x20 - ONLY_ANT_2
                "aoa_sampling_enable": 1,
                "aoa_pattern_len": 3,
                "aoa_ant_pattern": [3, 4, 5]
            }
        }
        rtlsUtil.aoa_set_params(aoa_params)
        print("AOA Params Set")
```

Switch to Antenna Array #2

Define antenna #3, #4, #5 as a list to use with Array #2

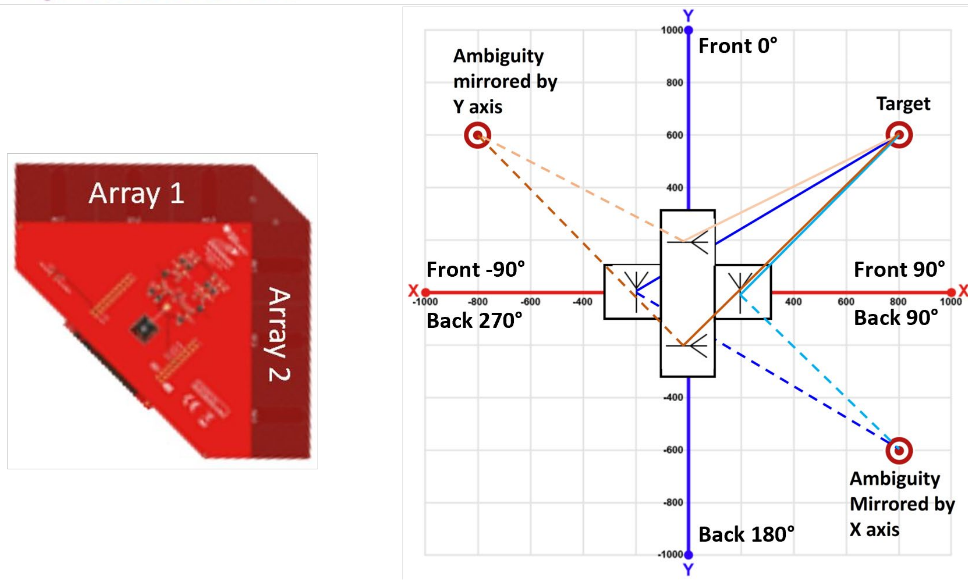


Fig. 11. Proposed AoA Enhancement with Orthogonal Antenna Arrays to avoid the front-and-back ambiguity.

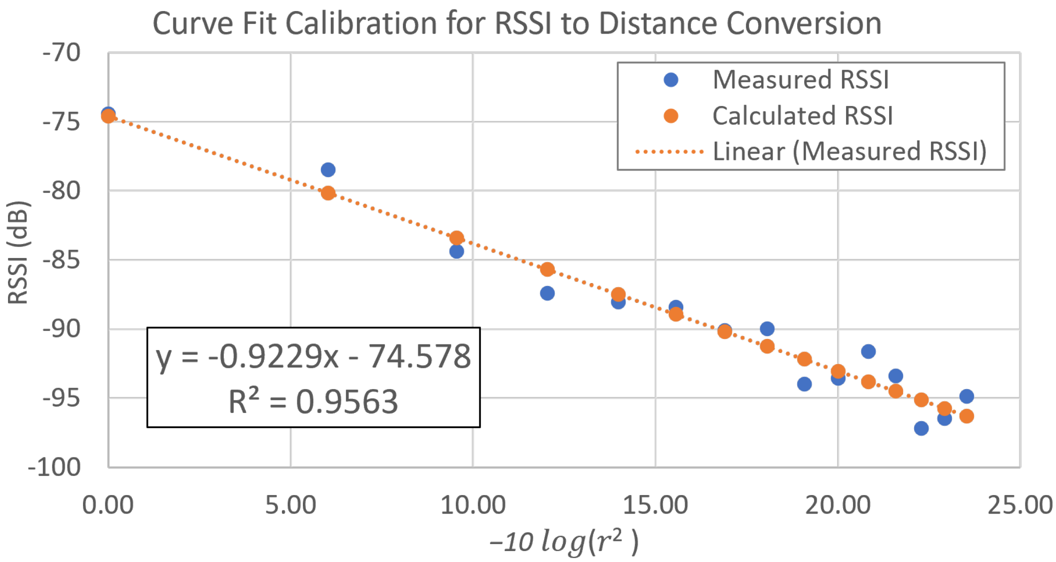
To evaluate the wayfinder's distance sensing accuracy, the beacon is placed from 1 to 15 meters away in 1-meter increments. The recorded RSSI values are averaged over 20 measurements. Then a curve fit is used to calibrate the RSSI-to-distance equation as shown in Fig.12. Since the RSSI is defined in dB of the ratio of power at distance ***r*** over power at reference distance of ***d***, the *RSSI v.s. log(r^2)* relation is expected to be linear:

$$RSSI = 10\,log\left(\frac{P_r}{P_d}\right) = 10\log(d^2) - 10\,log(r^2)$$

Line fit is used to calibrate the linear equation, and derived a refined distance equation:

$$RSSI \approx -74.578 - 0.9229 \cdot 10\, log(r^2)$$

$$r = 10^{-\left(\frac{RSSI+74.578}{0.9229\cdot 10\cdot 2}\right)}$$



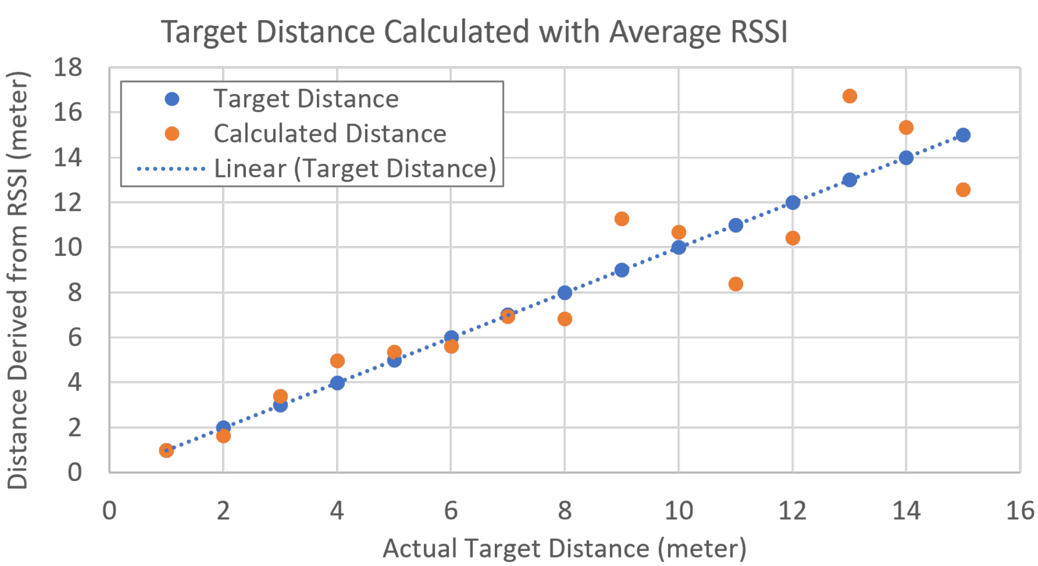


Fig. 12. RSSI target distance measurement calibration through line fit, and the 2m accuracy achieved within the 10m range and 4m accuracy in the full range.

For a more intuitive and graphical display of the target info, the distance r and angle θ are displayed in a radar format using a polar-like coordinate system. Forward direction is considered as 0°, facing up on the radar as shown in Fig.13. Proper scaling is performed to fit the plot and to label the distance at the tip of the green arrow.

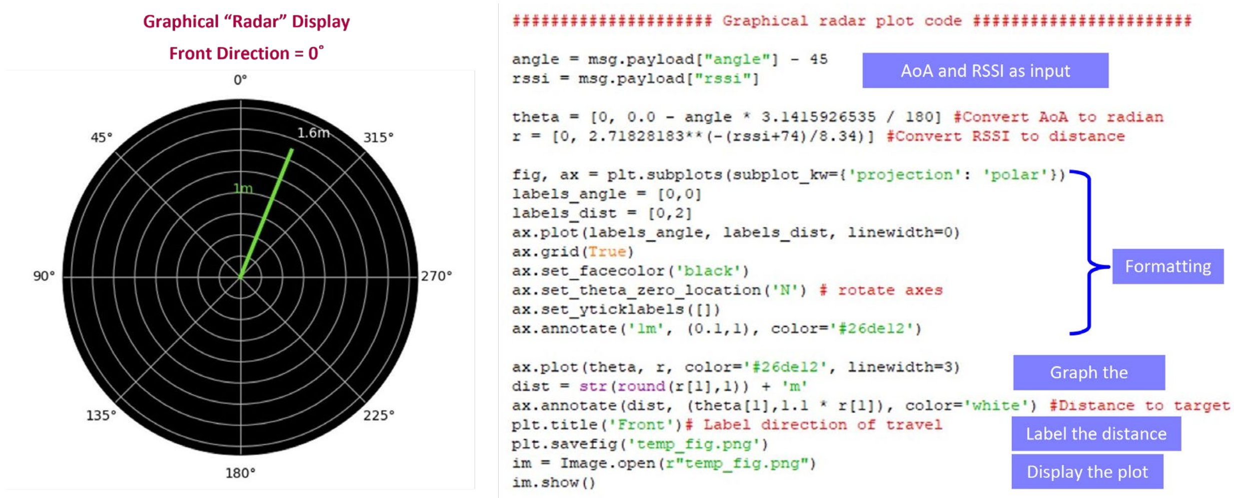


Fig. 13. The graphical display shows both direction and distance to the target in a radar-like format.

Conclusion and Future Research

This project introduced an innovative idea which reverses the standard whole premise mesh IPS configuration and moved the AoA capability to the user device to create a low-cost alternative solution to help visually impaired people navigate public facilities. This fully functional prototype used beacons to mark points of interest, and the AoA based point-to-point wayfinder worked as a smart compass to lead the user to the selected destination.

By introducing a simple linear dipole antenna array and Python code developed on top of the generic drivers for Texas Instruments Bluetooth 5.1 development board, this wayfinder successfully detected both target direction and distance. For a target in the front 180°, the system achieved a 10° accuracy over a 20 data point average. The target distance measured using RSSI also achieved a 2m accuracy within a 10 meters range with a 20 data point average, and a 4m accuracy for the full range. To showcase the real-world applications, a Python program was developed to graphically display a directional arrow and the distance to the target in an intuitive radar format. With a small additional investment, future smart phones and tablets can also integrate multiple Bluetooth antennas and support the new AoA feature. With multiple antennas built in, this proposed wayfinder can even be deployed as a new smart phone app.

Future Research

Some areas for potential improvement in the wayfinder system have been identified, such as the AoA front-and-back ambiguity associated with the linear antenna array that occurs when the human body does not block signal from behind. This ambiguity can be removed by using two orthogonal antenna arrays. A power amplifier can be used to extend the 15-meter range up to 84 meters if the maximum permitted 100mW transmission for Bluetooth is used.

With the target being a wearable assistive device, an integrated gyroscope will be needed to dynamically detect the antenna array angle. The angle and distance info in the radar display can be delivered as voice alerts using text-to-speech for those who are blind or have severely compromised vision.

Works Cited

Andersson, Philip, and Linus Persson. Evaluation of Bluetooth 5.1 as an indoor positioning system. *KTH Royal Institute of Technology, MS thesis*. www.diva-portal.org/smash/get/diva2:1468130/FULLTEXT01.pdf

Armas, Vlad De. “Using angle of arrival for direction finding with Bluetooth 5.1.” *Navigine*, 5 Apr. 2023, https://navigine.com/blog/using-angle-of-arrival-for-direction-finding-with-bluetooth/

“Bluetooth® Angle of Arrival (AoA) antenna design.” *Texas Instruments*, www.ti.com/lit/an/tida029a/tida029a.pdf. Data set.

Cominelli, Marco, Paul Patras, and Francesco Gringoli. “Dead on arrival: An empirical study of the Bluetooth 5.1 positioning system.” *WiNTECH*, https://doi.org/10.1145/3349623.3355475

Girolami, Michele, Paolo Barsocchi, Francesco Furfari, Davide La Rosa, and Fabio Mavilia. “Evaluation of angle of arrival in indoor environments with Bluetooth 5.1 direction finding.” *WiMob*, https://doi.org/10.1109/wimob55322.2022.9941619

Lee, C. K. M, C. M. Ip, Taezoon Park, and S. Y. Chung. “A Bluetooth location-based indoor positioning system for asset tracking in warehouse.” *IEEM*, 3 Feb. 2020, https://doi.org/10.1109/ieem44572.2019.8978639

Mavilia, Fabio, Paolo Barsocchi, Francesco Furfari, and Michele Girolami. “Evaluating the impact of anchors deployment for an AoA-based indoor localization system.” *WONS*, https://doi.org/10.23919/wons57325.2023.10061949

Pierre, Bonzom. “Indoor location: which technology for which purpose?” *ELA Innovation*, 18 July 2019, https://elainnovation.com/en/indoor-location-which-technology-for-which-purpose/

“RTLS Toolbox.” *Texas Instruments*, https://software-dl.ti.com/simplelink/esd/simplelink_cc13x2_26x2_sdk/3.30.00.03/exports/docs/ble5stack/ble_user_guide/html/ble-stack-5.x-guide/localization-index-cc13x2_26x2.html

Solovev, Andrey, and Anna Petrova. “Bluetooth indoor positioning systems: Implementation features & alternatives.” *Integra Sources*, 6 Apr. 2021, www.integrasources.com/blog/bluetooth-indoor-positioning-systems/

THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

Design and Optimization of Assistive Ultrasonic Echolocation

Santani Teng[1], Ian Reynolds[2]
Smith-Kettlewell Eye Research Institute[1], Independent Researcher[2]
santani@ski.org, idreyn@gmail.com

Abstract

Here we introduce "Robin", a prototype wearable device for human echolocation drawing inspiration from echolocating bats, established practices and techniques of unassisted human oral echolocation, and other biomimetic assistive technology. Robin extends this work by considering ergonomic and industrial design concerns across its design and directly addressing electrical, computational, and cost constraints imposed by a fully self-contained, wearable platform. We take a first step in addressing this need with a large-scale online study using simple virtual auditory display to assess the utility of Robin-like devices across key design parameters. We discuss how the time-stretching step of ultrasonic echolocation imposes constraints on acoustic frequency ranges of emitted and presented sound pulses, and distorts binaural cues used for localization in the azimuthal plane. We incorporate these constraints and some proposed remedies for these distortions into echoacoustic simulations mimicking Robin's auditory output. The stimuli were presented as a localization task to online study participants, who discriminated stimulus positions most reliably within 20° of the median plane when maximally slowed and compensated. We discuss the broader approach of empirically sampling parameter spaces, and how alternate or complementary sensing technologies might be used in future mobility aids that similarly leverage biosonar-derived, context-adaptive design principles.

Keywords

Assistive technology, mobility, echolocation, ultrasound, perception, user testing.

Introduction

Active echolocation involves perceiving the environment by analyzing reflections of self-generated sounds. Echolocation using tongue clicks and similar cues is well-documented in blind and visually impaired (BVI) humans (Kolarik, Cirstea, et al.; Thaler and Goodale); early-blind practitioners sometimes discover the technique naturally (Mccarty and Worchel). This method allows BVI travelers to perceive the presence and relative positions of objects (Teng et al.; Rice); the general geometry of their surroundings (Flanagin et al.), and orientation relative to nearby features (Dodsworth et al.; Rosenblum et al.), using active sensory sampling. The perceptual and neural mechanisms of expert echolocation remain subjects of active research; however, expert oral pulse echolocation has been associated with vision-like brain activity (Lore Thaler et al.; L. Thaler et al.), improved auditory spatial cognition (Vercillo et al.), and proficient obstacle-avoidance navigation (Kolarik, Scarfe, et al.). Regular echolocation use is tentatively associated with higher quality of life (Thaler; Norman et al.) in its practitioners, suggesting a generalized benefit for multiple aspects of daily living. However, it remains a rarely used method, limited by incomplete understanding of its perceptual mechanisms, a steep learning curve, possible social reticence, and suboptimal sensory cues.

Assistive artificial sonar systems have a closely shared history with research in biological echolocation, offering the possibility of narrowing the gap between the potential and realized benefits of echolocation. Devices emitting audible or ultrasonic pulses for their operators' perceptual benefit have been tested since at least World War II (Supa et al.; Twersky). Since then, electronic travel aids utilizing ultrasonic reflections have aimed to facilitate navigation, including some with explicit biomimetic design basis (Kay, "A Sonar Aid to Enhance Spatial Perception of the Blind: Engineering Design and Evaluation"; Kay, "Auditory Perception of

Objects by Blind Persons, Using a Bioacoustic High Resolution Air Sonar"; Ifukube et al.; Mihajlik et al.; Waters and Abulula; Kuc and Kuc).

For example, artificial "time-stretching" of ultrasonic echoes into an audible frequency range while preserving temporal structure has shown benefits for shape discrimination compared to non-stretched frequency-matched sounds (Ifukube et al.; Sumiya et al.). This facilitation of ultrasonic echolocation has been explained by increased salience of echo structure at the lower, audible frequencies (Fujitsuka et al.). Still, sensory substitution devices and electronic travel aids have not seen wide adoption compared to traditional methods used by blind travelers, such as white canes or guide dogs (Loomis et al.). Their limitations may arise from mismatches between signals and human perception, difficulty of training, and lack of ergonomic design (Kristjánsson et al.; Maidenbaum et al.).

Making assisted echolocation more useful may require better matching signals to task-relevant perception while facilitating training via customizable, wearable platforms. As a prototype sensory aid following these principles, we introduce Robin, a wearable artificial echolocation system. Robin combines an unobtrusive form factor and configurable ultrasound emissions with real-time digital signal processing to deliver processed, audible echoes to the user through headphones.

Discussion

An echolocation-based sensory substitution platform

Here we describe "Robin," a head-mounted ultrasonic sonar platform designed to generate ultrasonic echoes and deliver audible informative binaural cues to a human user. We aimed for a compromise between low-latency, long-lasting, durable, highly configurable operation; and unobtrusive, portable, low-cost form factors. The resulting device emphasizes an evolution away from heavy, general-purpose components and toward portability, real-time processing, and eventual field-testing.

Upon user-initiated or automatic triggering, the system generates millisecond-scale ultrasound signals (“chirps”), amplifies them for emission by piezoelectric transducers, records the audio, and processes the echoes for playback to the user with latency on the order of 100ms. A handheld controller or connected web application may be used to modulate the shape, frequency, or duration of the sound for use in different spatial contexts. We describe the hardware and software architecture, along with a recent prototype implementation, in Figure 1. Relevant code and designs will be available as Supplementary Material.

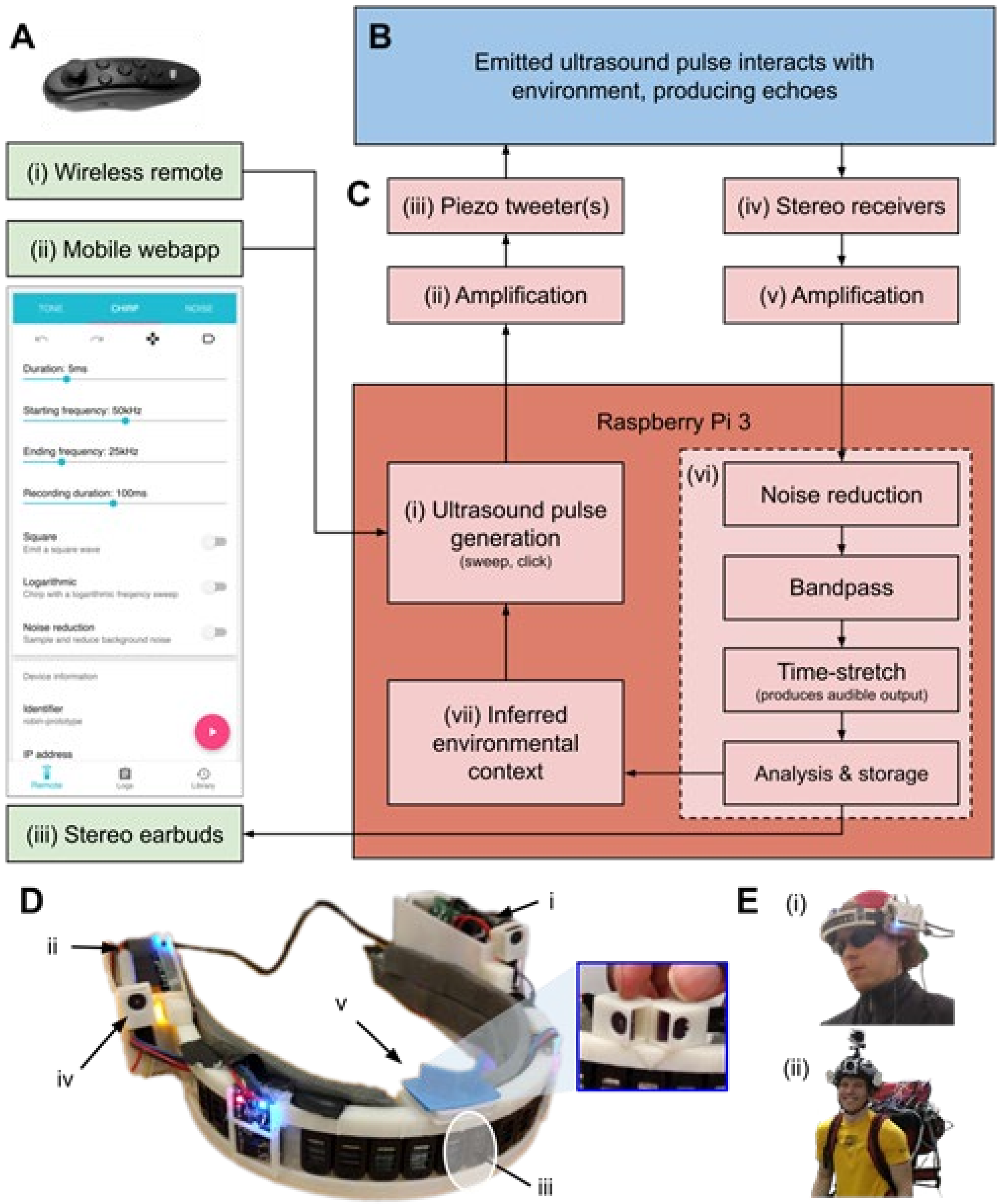


Fig. 1. "Robin" echolocation device. Diagrammatic system overview in which the user, via interface peripherals (A, green, images), interacts with the environment (B, blue) via the Robin transduction/processing assembly (C, pink) centered on the onboard computer housing (red). Representative screenshot of mobile webapp shows customizable pulse properties. Upon user initiation (A*i*,*ii*), the system generates a waveform (C*i*) for amplified pulse emission via single or arrayed tweeters (C*ii,iii*) depending on configuration. The returning echoes are recorded and

amplified by stereo receivers (C*iv,v*), processed for playback/storage (C*vi,* A*iii*), and potentially used to update context-sensitive pulse parameters (C*vii*). An incarnation of the Robin headset is shown in D, with key components labeled as follows: *(i,v–vii)* Raspberry Pi 3 and associated A/D conversion hardware; *(ii)* rechargeable battery and power supply; *(iii)* piezoelectric TAKE-T Batpure tweeter array, with white oval representing alternate single-transducer configuration; *(iv)* one of a pair of stereo microphone ports; *(v)* detachable frontal mount for the microphone enclosures at high compensation factor, detail shown in inset. Wearable form factor is shown in E*(i)*, contrasted with earlier iteration in E*(ii)* (Sohl-Dickstein et al., 2015).

Basic hardware overview

Robin's components and form factor can be divided as follows:

- Enclosure: A 3D-printed partial ring worn around the brow with compartments on each side. This facilitates rapid iteration.
- Onboard computer: Raspberry Pi 3 running Linux-based OS controls timing and signal processing for waveform generation, recording, time-stretching, and playback.
- Audio hardware: Raspberry Pi interfaced with stereo codec and amplifiers.
- Emitter array: 10 ultrasonic piezoelectric tweeters in front ring. An alternate single-emitter configuration was also tested and served as the basis for simulated virtual echoes.
- Receivers: A pair of Knowles ultrasonic microphones with custom ±70V amplifiers capture echoes up to 96kHz.
- Support hardware: Battery, power regulation, headphones for audible playback.

This combines low-cost, compact parts with custom signal processing for a self-contained wearable. The enclosure has several notches for different receiver spacings.

Software and Device Operation

The Raspberry Pi runs Python code to:

- Generate arbitrary ultrasonic waveforms for emission.
- Continually record audio into a rolling buffer.
- After each pulse, begin recording echoes.
- Process sample into audible signals (filter, denoise, time-stretch).
- Play back processed recordings to the user via headphones.

Key parameters like frequency and timing are configurable via Bluetooth remote or web interface served from the device. Some of these include:

- Emission frequency and duration: Robin uses ~20-45 kHz, just above human hearing.
- Audible presentation frequencies: Slowed emulsions are presented at ~1.5-2.9 kHz to leverage sensitivity to binaural cues in this range.
- Slowdown factor S: The factor by which ultrasound is slowed, e.g., S=10 slows a 30 kHz pulse to a 3 kHz audible chirp.
- "Compensation factor" C: The ratio of natural ear spacing to the virtual receiver spacing. C determines the degree to which time-stretching distortion is counteracted. Robin's receivers can be physically moved along its front brim to change this parameter.

The time-stretching operation preserves temporal structure, but distorts spatial cues. The slowdown factor S dilates interaural time differences (ITDs) which normally facilitate localization of lower frequencies. At normal interaural distance (C=1) these ITDs are greatly exaggerated. Full compensation (C=S) restores natural ITDs by placing microphones close together (Figure 2).

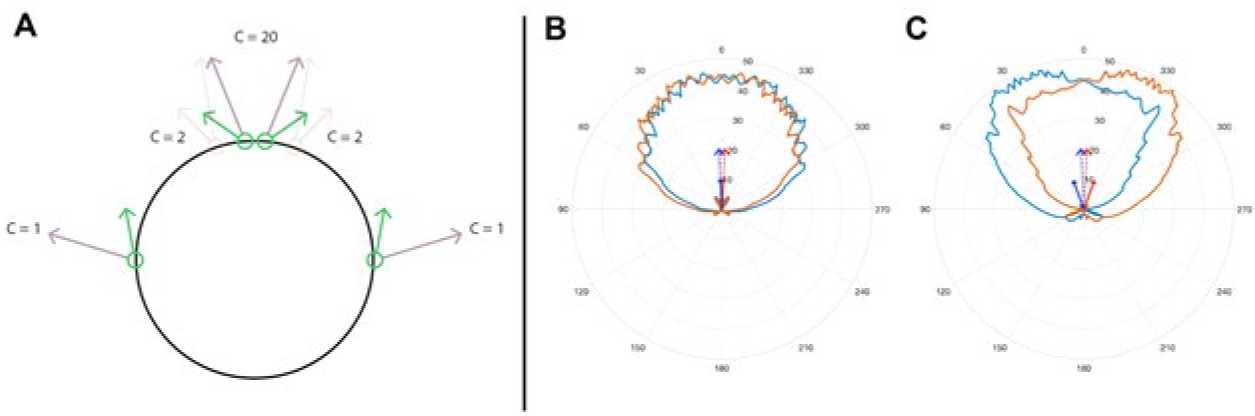


Fig. 2. Illustrative binaural receiver positions and gain patterns at $S = 20$. A. At $C = 1$, receivers are at an ear-like antipodal position, converging quickly toward the front of the head as C approaches S ($S = 20$ in this example), Note the inter-receiver distance is $1/C$ relative to the natural interaural axis. Green arrows indicate the focal axis of each receiver's parabola; magenta arrows indicate the centroid of the induced gain patterns. B shows binaural gain patterns for a receiver at $C = 20$ whose parabolic axis is normal to the head. These closely-spaced receivers have nearly the same gain pattern. C. shows differentiating the gain patterns and recovering a pronounced ILD.

"Robin" as an exemplar of a device type

Preliminary piloting revealed an intuitive link between the slowed temporal structure of echoes and environmental features such as stairs or walls, and relative distances inferable from temporal delay and attenuation. Early user testing feedback suggested a need to closely examine binaural signal characteristics for improved azimuthal localization. To narrow the space of potentially useful device configurations across parameters such as signal design, time-stretch factor, and physical configuration of the device's emitters and receivers, we conducted a simple, large-scale online study of auditory localization using simulated echoacoustic stimuli similar to those generated by the Robin prototype.

We consider a general model of Robin-like wearable devices in which a pair of parabolic receivers are embedded in a circular plastic ring worn on the brow, and the ensonifying pulse is emitted from a point source on the forehead. The focal width and depth of these receivers are fixed, but their position and orientation are allowed to vary. The acoustic pulses are emitted and recorded in the range of 20 to about 60 kHz, consistent with the ranges of other ultrasound-based devices (Sohl-Dickstein et al.; Sumiya et al.; Kay, "Auditory Perception of Objects by Blind Persons, Using a Bioacoustic High Resolution Air Sonar"; Ifukube et al.) and encompassing the first harmonics of major echolocating bat vocalizations (Ulanovsky and Moss). The recordings are then slowed to be human-audible and presented in an 5-AFC azimuth-discrimination task to online study participants.

Our results suggest that closely-spaced binaural receivers which physically compensate for the inherent time-stretching of binaural cues outperform those which provide no compensation, i.e., at the natural interaural distance. This is despite the fact that time-stretched cues are known to convey spatial information (Shinn-Cunningham et al.; Sohl-Dickstein et al.). Stimuli from uncompensated receivers were poorly discriminated within left, front, and right sectors; in particular, the left and right sectors show a strong bias toward the most lateral azimuths. Compensated, closely-spaced receivers recover a higher degree of acuity in the frontal sectors.

Online Localization Study

To assess Robin's utility for spatial localization, we conducted an online experiment in which participants judged the source direction of simulated signals. We modeled the device and a spherical head within a 2D acoustic grid using the k-Wave MATLAB toolbox, which solves wave equations using a pseudospectral time-domain method suited for high-frequency tasks (Treeby and Cox, 2010). The simulation modeled the parabolic shape of Robin's directional microphone by a cutout within the headband portion of the model (Figure 3). We rotated this parabola over a +/- 80° range to generate directional gain patterns. For each combination of slowdown *S* and compensation *C* (see Table 1), receivers were oriented so as to optimally recover natural human interaural level differences (ILDs) using a least-squares criterion relative to reference measurements from a KEMAR mannequin (Gardner and Martin, 1995) (Figure 4). A two-dimensional head-related impulse response (HRIR) was measured at 1˚ azimuthal intervals within the simulation. Experimental stimuli were generated from FM sweep pulses filtered through the HRIRs associated with each source azimuth. This produced audible stereo presentations simulating reflections from different directions (Figure 5).

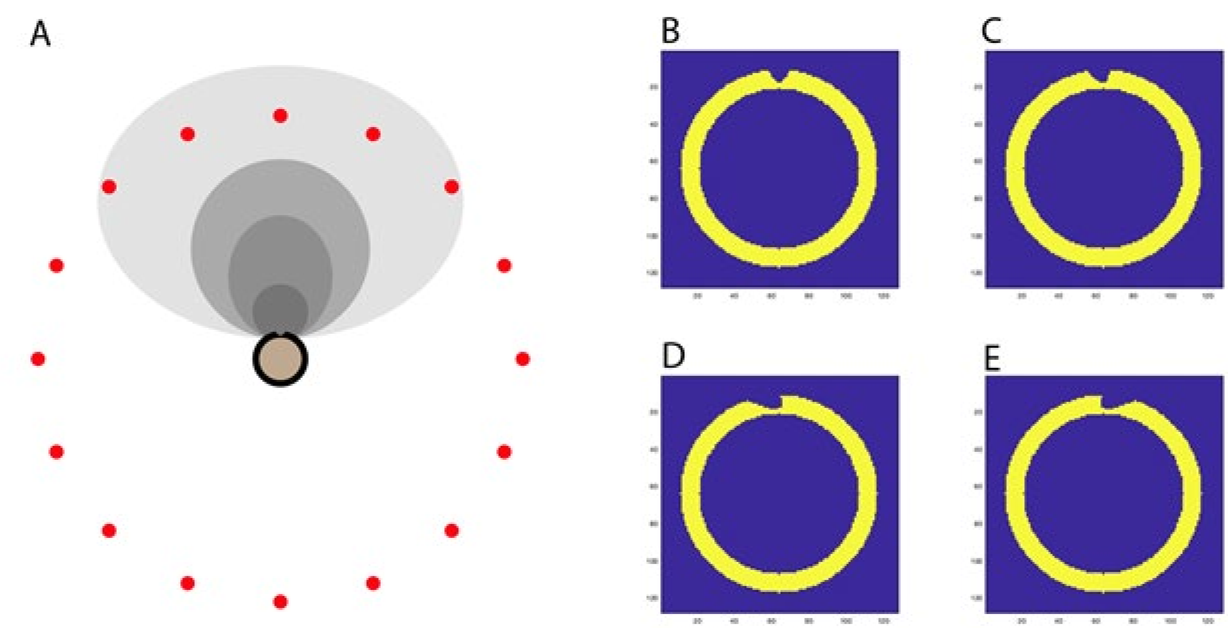


Fig. 3. Simulation overview. (A) a schematic of the acoustic simulation. A circular head (beige) is ringed by a plastic headband (black) with a cutout parabolic notch to create directivity. An acoustic impulse is created at the focus of the notch, and the resulting gain pattern is recorded at sensor points (red) ringing the head in the far field. Not to scale; fewer sensor points shown for clarity. (B-E) plots sampling the space of notches used to record an impulse response, including

one normal to the headband (B) as well as oriented away from the normal at steeper angles (C-E).

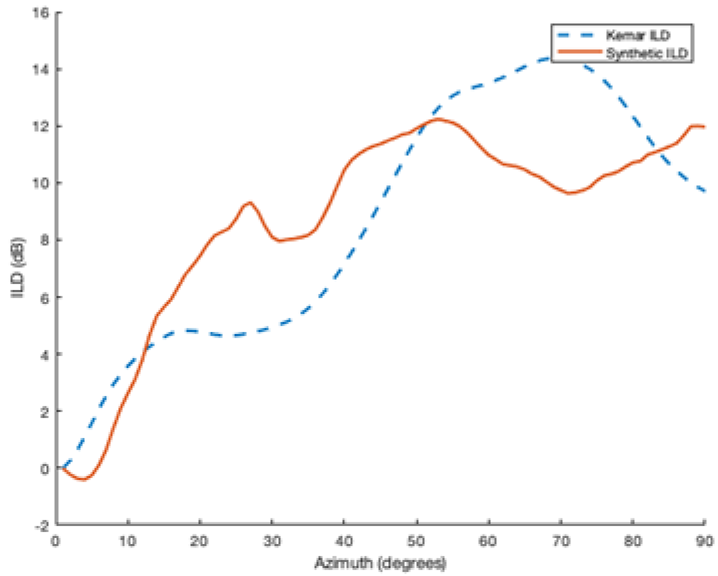


Fig. 4. ILD comparison. The narrowband KEMAR ILD in our presentation frequency range is shown plotted against its closest match among our pool of receiver directivities on the range of azimuths $0 \leq \theta \leq 90$ for compensation factor $C = 12$.

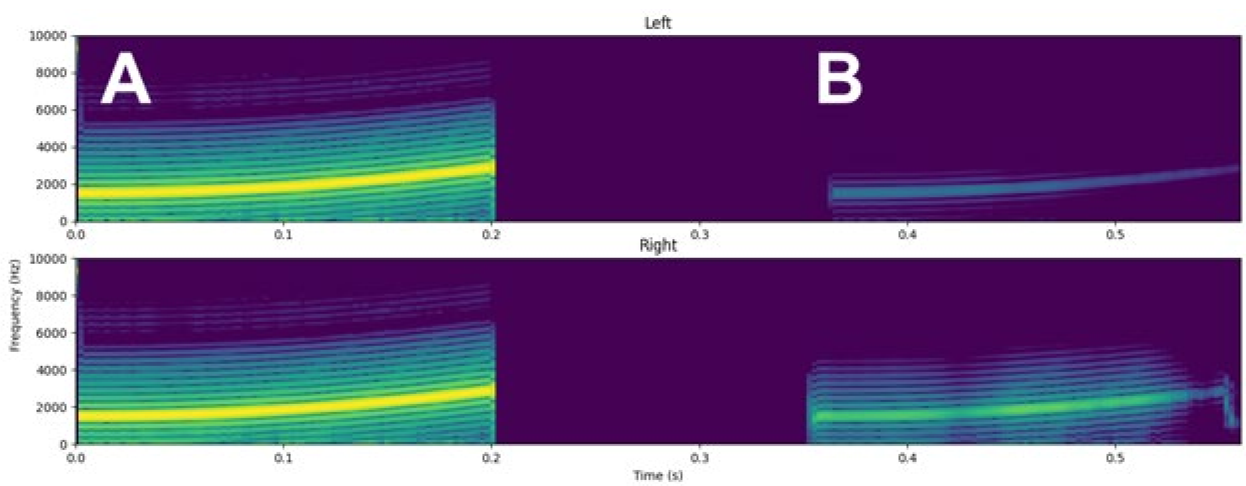


Fig. 5. Example echo stimulus. The spectrogram of a presentation with $S = 20$, $C = 1$, and $\theta = 45^{\circ}$ showing (A) the outbound chirp at $t = 0$ and (B) the inbound echo at $t = t_e$. In this example, the

echo is incident from the right side, so it appears on the right channel (bottom) sooner and with higher amplitudes.

One hundred ten adults with normal hearing and vision were recruited online using the experimental platform Prolific.ac and tested remotely using the Pavlovia platform. All participants provided informed consent in accordance with procedures approved by The Smith-Kettlewell Eye Research Institute's Institutional Review Board. After volume calibration and a headphone check, participants performed the sound localization task as follows, and as shown in Figure 6.

On each trial, participants heard a "chirp" sweep mimicking Robin's emission, followed by a simulated reflection with varying arrival times and intensities at the left and right virtual receivers. The task was to judge from which of 5 lateralized positions in the Left (-60° centered), Center (0°), or Right (+60°) sectors the echo originated. Twenty trials were collected for each of 15 positions across 6 combinations of slowdown S and compensation C factors (Table 1).

To measure performance while controlling for response bias, we computed the sensitivity d' for each pair of positions based on response confusion matrices within each sector (Macmillan and Douglas Creelman). This allowed us to estimate positional acuity beyond simply computing proportions of correct responses. The basic intuition is that proximity matters: a positional error of 1 is more precise than an error of 4. Next, we regressed d' against distance between stimulus positions (Fischer and Whitney) with a modified general linear model that accounts for repeated measures per participant (Bakdash and Marusich). For an observer sensitive to spatial position, more widely separated stimuli are more distinguishable (higher d'), and the regression slope would be more positive.

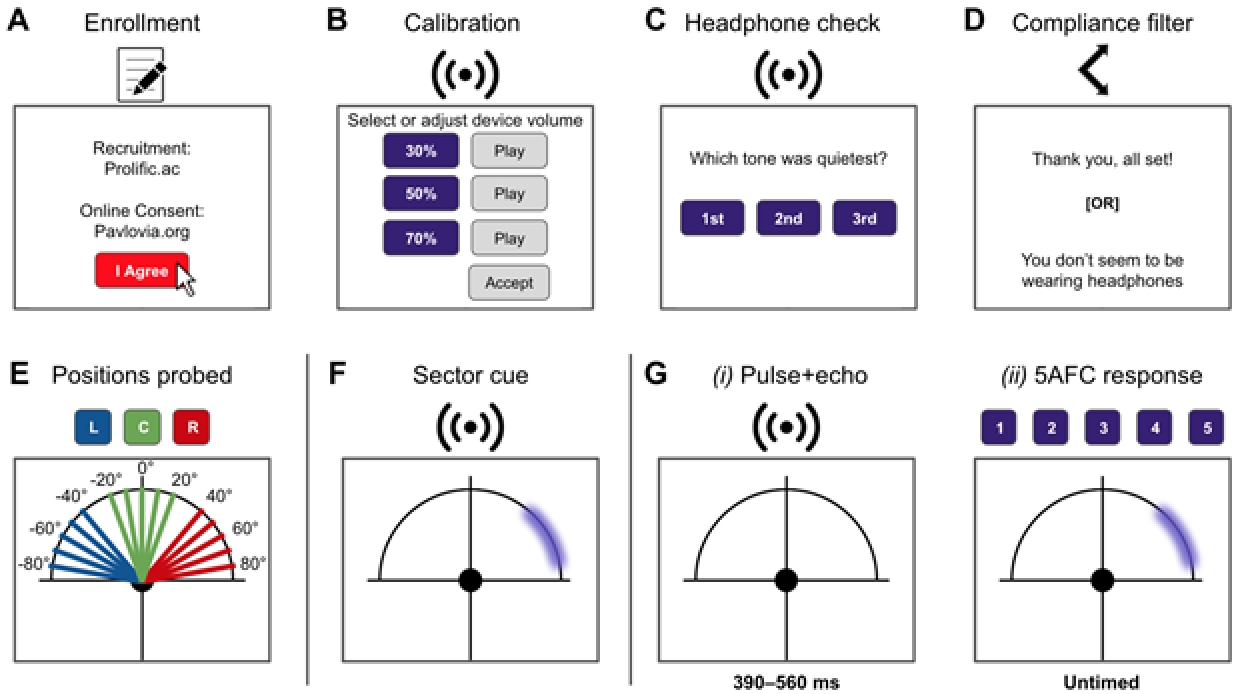


Fig. 6. Overview of online experimental design and procedures. Upon online enrollment and consent (A), participants self-calibrated presentation volume (B), then performed a headphone check task (C) whose result continued or exited the experiment (D). Trials were clustered in blocks corresponding to 40°-wide "sectors" of stimulus azimuths relative to the observer, displayed schematically in (E). During the experiment, each block began by visually displaying and audibly demarcating the sector being probed (F). On each trial (G), a button press triggered stimulus presentation (i), followed by an untimed 5AFC response cue (ii) with the current sector visually highlighted. Responses did not elicit feedback and were followed by the onset of the next trial. See main text for further details.

Behavioral Results

Figure 7 shows response distributions across positions for two conditions. Deviations from the major diagonal reflect greater discrimination difficulty. Formal precision results in Figure 8 confirm reliable position coding only centrally for S=12 and S=20. For S=12, significant discrimination occurred only for C=S (full compensation). Figure 8D shows ~70% accurate lateralization of a reflector ±10° from the midline at S=20, consistent with results with similar stimuli in previous work (Sohl-Dickstein et al.). Overall performance is summarized in Figure 9C, indicating benefits of both greater slowdown and normalized timing cues. Performance resembled that for validation stimuli modeled using a standard HRTF dataset (Fig. 1, Table 1), confirming the difficulty stems from the echoacoustic stimuli.

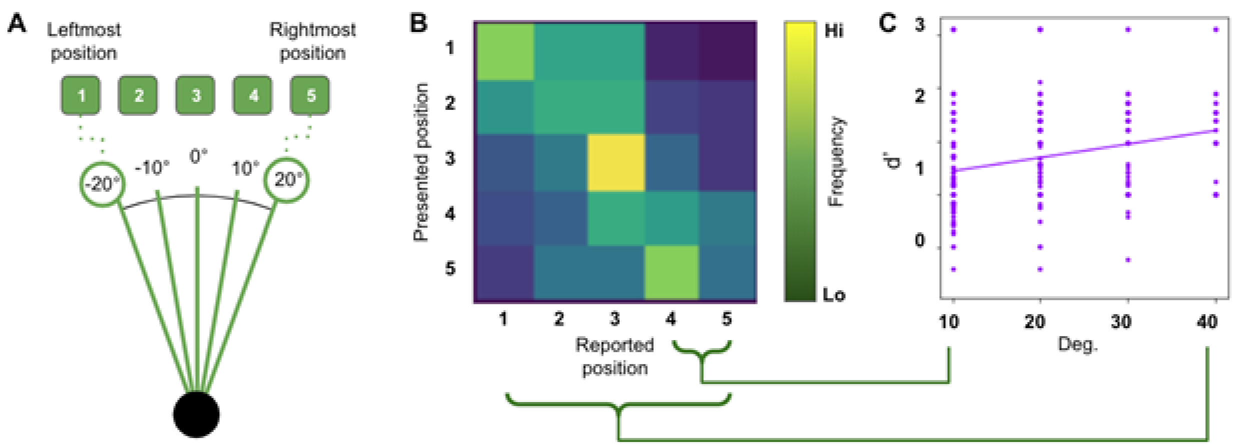


Fig. 7. Computing sensitivity and precision from behavioral responses. A. Participants reported each stimulus position relative to left- and rightmost positions within each sector; azimuths of central sector shown for illustration. B. Example confusion matrix generated from Presented vs. Reported positions to compute d’ as a function of pairwise stimulus separation, ranging from 10° to 40°. C. Estimating spatial precision. Each dot represents d’ for a pairwise stimulus separation. Brackets from panel B indicate two example position pairs (of 10 possible per participant) to compute separation. Slope of linear fit indexes precision of spatial coding across separations and participants. A fit with slope of zero represents the null hypothesis of insensitivity to stimulus position.

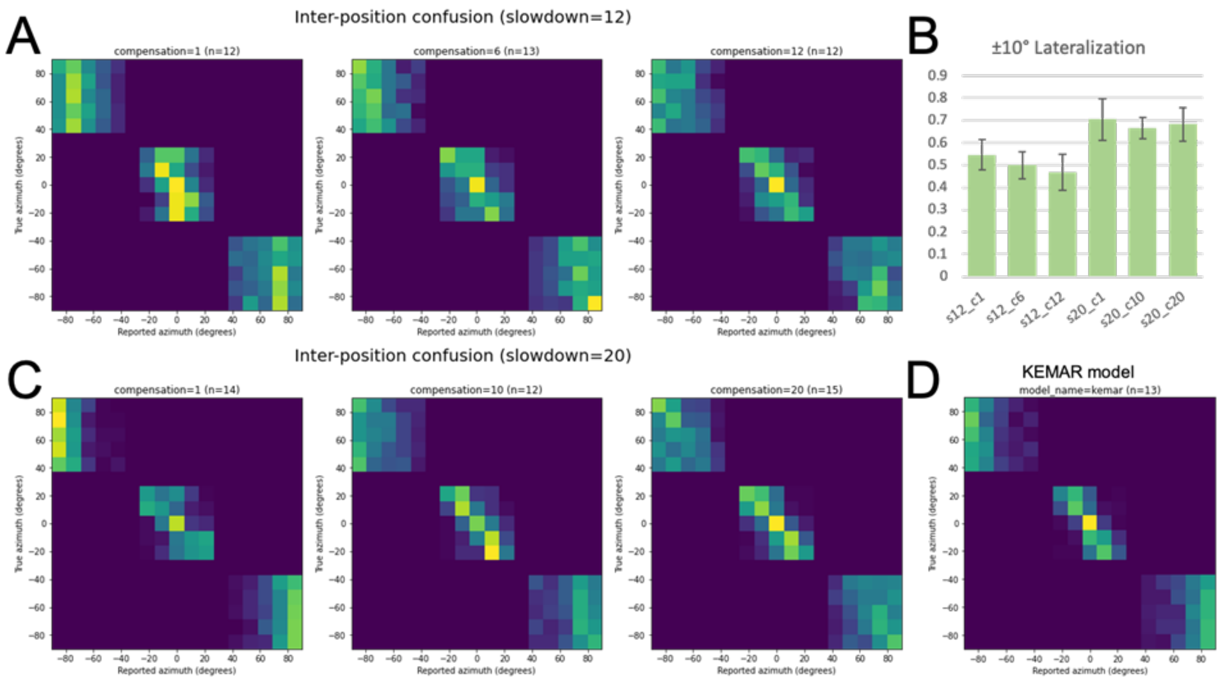


Fig. 8. Behavioral confusion matrices for presented (Y-axis) vs. reported (X-axis) stimulus positions, with azimuths comprising all 3 sectors within each plot. Results separated by no, half, and full compensation in respective columns for slowdown of 12 (row A) and 20 (row C). Color scale reflects normalized frequencies of reported positions for each presented position (sums are equal over each row). Note that trials, and thus responses, were blocked by sector (e.g. left sector stimuli could never be reported in central or right sector). Estimated lateralization performance for ±10° azimuths shown in B for comparison with previous work (see Discussion); category labels indicate slowdown and compensation conditions. Plot in D indicates results for stimuli generated with KEMAR head model with natural interaural distance (S=20, C=1).

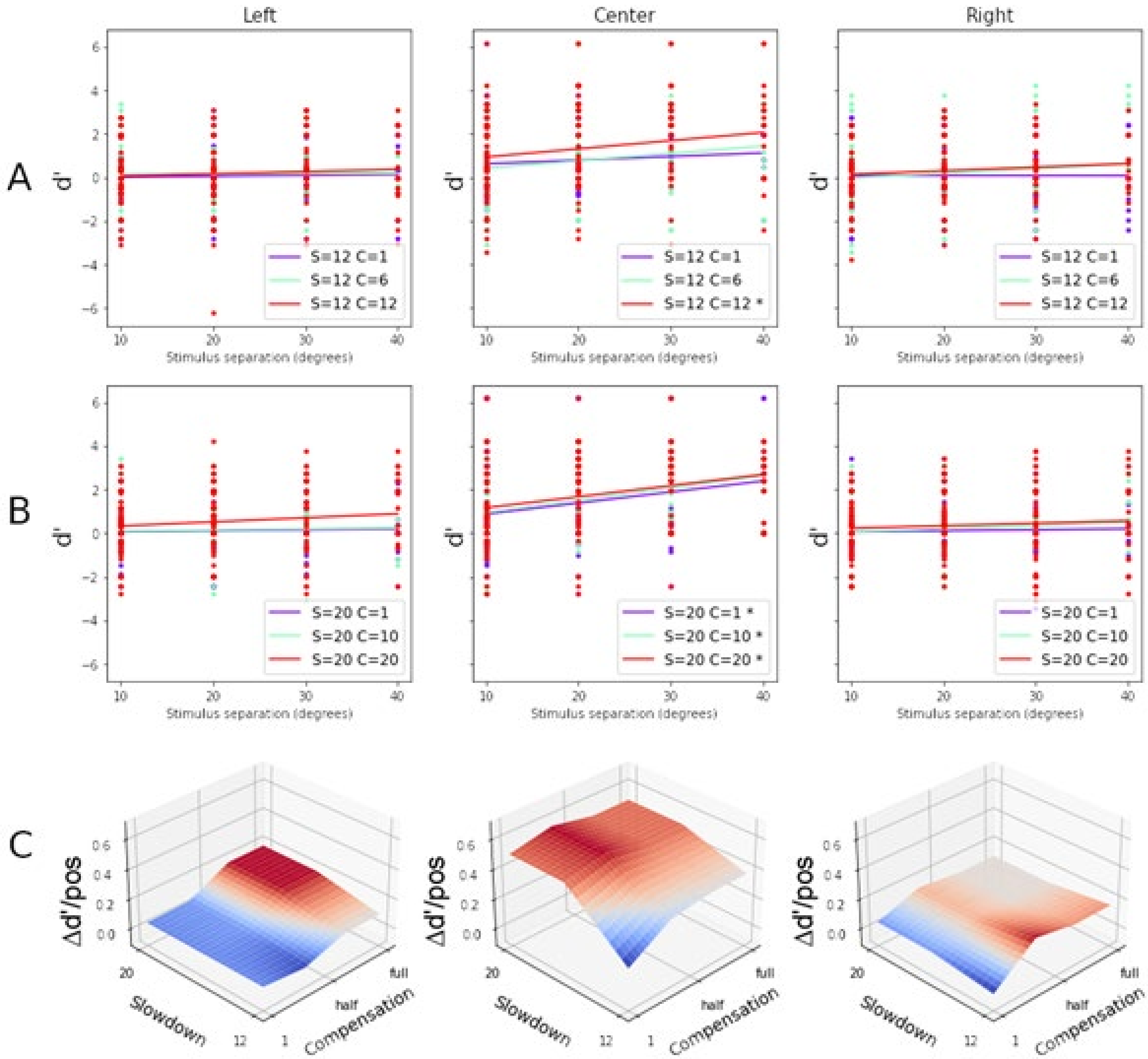


Fig. 9. Behavioral precision. Analyses are divided by slowdown factors of S=12 (A) and S=20 (B) and Left, Center and Right position sectors. Each fit within plots is for a different compensation factor: 1 (no compensation); half (½ of S); or full (equivalent to S). Asterisks in legend indicate significant slopes. C. Interpolated space of precision, plotting slope (Δd' per position shift) on the *z* axis as a function of slowdown and compensation factor in each spatial sector.

Conclusion

Robin is a head-worn device that emits ultrasound pulses into the environment and streams a processed, human-audible version of the recorded echoes back to the wearer. The device combines low-cost, readily available hardware components; a custom hardware/software audio processing pipeline; and design principles inspired by advances in understanding biological echolocation. Its size and configurability make it suitable as both a prototype navigation aid and a research platform for generating and evaluating various acoustic signals as echolocation cues. Ultrasound pulses provide high spatial resolution and may be slowed down (band-shifted to audible frequencies) to produce cues robust to noisy environments and readily interpretable even by novice users. Stepping back, we consider a parameterized family of Robin-like devices and use simulation and online trials to quickly evaluate their presentation of binaural cues encoding the position of echo-reflecting objects, systematically manipulating slowdown and binaural factors to find task-specific optima in the signal parameter space.

Works Cited

Bakdash, Jonathan Z., and Laura R. Marusich. "Repeated measures correlation." *Frontiers in Psychology*, vol. 8, Apr. 2017, p. 456.

Dodsworth, C., et al. "Navigation and perception of spatial layout in virtual echo-acoustic space." *Cognition*, vol. 197, 2020, p. 104185, doi:10.1016/j.cognition.2020.104185

Fischer, Jason, and David Whitney. "Precise discrimination of object position in the human pulvinar." *Human Brain Mapping*, vol. 30, no. 1, Jan. 2009, pp. 101–11.

Flanagin, Virginia L., et al. "Human exploration of enclosed spaces through echolocation." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, vol. 37, no. 6, Feb. 2017, pp. 1614–27.

Fujitsuka, Yumi, et al. "Two-dimensional shape discrimination by sighted people using simulated virtual echoes." *JASA Express Letters*, vol. 1, no. 1, Jan. 2021, p. 011202.

Ifukube, T., et al. "A blind mobility aid modeled after echolocation of bats." *IEEE Transactions on Bio-Medical Engineering*, vol. 38, no. 5, May 1991, pp. 461–65.

Kay, L. "A sonar aid to enhance spatial perception of the blind: Engineering design and evaluation." *Radio and Electronic Engineer*, vol. 44, no. 11, Nov. 1974, pp. 605–27.

Kay, L. "Auditory perception of objects by blind persons, using a bioacoustic high resolution air sonar." *The Journal of the Acoustical Society of America*, vol. 107, no. 6, June 2000, pp. 3266–75.

Kolarik, Andrew J., Silvia Cirstea, et al. "A summary of research investigating echolocation abilities of blind and sighted humans." *Hearing Research*, vol. 310, Apr. 2014, pp. 60–68.

Kolarik, Andrew J., Amy C. Scarfe, et al. “Blindness enhances auditory obstacle circumvention: Assessing echolocation, sensory substitution, and visual-based navigation.” *PloS One*, vol. 12, no. 4, Apr. 2017, p. e0175750.

Kristjánsson, Árni, et al. “Designing sensory-substitution devices: Principles, pitfalls and potential 1.” *Restorative Neurology and Neuroscience*, vol. 34, no. 5, 2016, pp. 769–87.

Kuc, Roman, and Victor Kuc. “Experimental audible sonar to model echolocation by the blind.” *Proceedings of Meetings on Acoustics*, 2016, doi:10.1121/2.0000442.

Loomis, Jack M., et al. “Sensory substitution of vision: Importance of perceptual and cognitive processing.” *Assistive Technology for Blindness and Low Vision*, 2012, pp. 162–91.

Macmillan, Neil A., and C. Douglas Creelman. “Detection Theory: A User’s Guide.” *Psychology Press*, 2004.

Maidenbaum, Shachar, et al. “Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation.” *Neuroscience and Biobehavioral Reviews*, vol. 41, Apr. 2014, pp. 3–15.

Mccarty, Bruce, and Philip Worchel. “Rate of motion and object perception in the blind.” *Journal of Visual Impairment & Blindness*, vol. 48, no. 9, Nov. 1954, pp. 316–22.

Mihajlik, P., et al. “DSP-based ultrasonic navigation aid for the blind.” *IMTC 2001. Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No.01CH 37188)*, vol. 3, 2001, pp. 1535–40 vol.3.

Norman, Liam J., et al. “Human click-based echolocation: Effects of blindness and age, and real-life implications in a 10-week training program.” *PloS One*, vol. 16, no. 6, June 2021, p. e0252330.

Rice, C. E. “Human echo perception.” *Science*, vol. 155, no. 3763, Feb. 1967, pp. 656–64.

Rosenblum, Lawrence D., et al. “Echolocating distance by moving and stationary listeners.” *Ecological Psychology: A Publication of the International Society for Ecological Psychology*, vol. 12, no. 3, July 2000, pp. 181–206.

Shinn-Cunningham, B. G., et al. “Adapting to supernormal auditory localization cues. I. bias and resolution.” *The Journal of the Acoustical Society of America*, vol. 103, no. 6, June 1998, pp. 3656–66.

Sohl-Dickstein, Jascha, et al. “A device for human ultrasonic echolocation.” *IEEE Transactions on Bio-Medical Engineering*, vol. 62, no. 6, June 2015, pp. 1526–34.

Sumiya, Miwa, et al. “Bat-inspired signal design for target discrimination in human echolocation.” *The Journal of the Acoustical Society of America*, vol. 145, no. 4, Apr. 2019, p. 2221.

Supa, Michael, et al. “‘Facial vision’: The perception of obstacles by the blind.” *The American Journal of Psychology*, vol. 57, no. 2, 1944, pp. 133–83.

Teng, Santani, et al. “Ultrafine spatial acuity of blind expert human echolocators.” *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, vol. 216, no. 4, Feb. 2012, pp. 483–88.

Thaler, Lore, et al. “Neural correlates of motion processing through echolocation, source hearing, and vision in blind echolocation experts and sighted echolocation novices.” *Journal of Neurophysiology*, vol. 111, no. 1, Jan. 2014, pp. 112–27.

Thaler, Lore. “Echolocation may have real-life advantages for blind people: An analysis of survey data.” *Frontiers in Physiology*, vol. 4, May 2013, p. 98.

Thaler, Lore, et al. “Neural correlates of natural human echolocation in early and late blind echolocation experts.” *PloS One*, vol. 6, no. 5, May 2011, p. e20162.

Thaler, Lore, and Melvyn A. Goodale. “Echolocation in humans: An overview.” *Wiley Interdisciplinary Reviews. Cognitive Science*, vol. 7, no. 6, Nov. 2016, pp. 382–93.

Twersky, Victor. “On the physical basis of the perception of obstacles by the blind.” *The American Journal of Psychology*, vol. 64, no. 3, 1951, pp. 409–16.

Ulanovsky, Nachum, and Cynthia F. Moss. “What the bat’s voice tells the bat’s brain.” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 25, June 2008, pp. 8491–98.

Vercillo, Tiziana, et al. “Enhanced auditory spatial localization in blind echolocators.” *Neuropsychologia*, vol. 67, Jan. 2015, pp. 35–40.

Waters, Dean A., and Husam H. Abulula. “Using bat-modelled sonar as a navigational tool in virtual environments.” *International Journal of Human-Computer Studies*, vol. 65, no. 10, Oct. 2007, pp. 873–86.

2024
THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

User Experience of Voice Assistants by People with Visual Disabilities

Hyung Nam Kim
North Carolina A&T State University
hnkim@ncat.edu

Abstract

As technology advances, people with visual disabilities increasingly rely on emerging technologies (e.g., Siri, VoiceOver, and Microsoft's Seeing AI application optimized for use with VoiceOver). Those technologies are powered by voice user interfaces that assist users with reading information, controlling systems, and communicating with others. Yet, there is limited research on how people with visual disabilities interact with the voice assistants, focusing on gesture commands, voice commands, and relevant user interfaces. To address the knowledge gap, this study investigated how people with visual disabilities interact with the iPhone's voice assistant features and an assistive technology app accessible through the voice assistant features. This study found that people with visual disabilities had a poor user experience, and design recommendations were provided.

Keywords

Siri, VoiceOver, Seeing AI, user interfaces, usability, accessibility

Introduction

The prevalence of visual disabilities, including visual impairment and blindness, is significant in the United States. Over one million Americans live with blindness, and over 8 million Americans are visually impaired by uncorrected refractive error (Varma et al.). The Centers for Disease Control and Prevention (CDC) has recently released a report that 93 million American adults are at high risk for serious vision loss, and only half of them were likely to visit an eye doctor. The economic cost related to vision issues is estimated to increase to $373 billion by 2050 (CDC).

Today's technological advances have brought highly accessible consumer products to people with visual disabilities in a variety of domains – healthcare, education, rehabilitation training, and Internet access (Kim). For example, a survey study by Crossland et al. found that over 80% of 132 respondents with visual disabilities used a smartphone for phone calls, texting, reading, browsing the Internet, and identifying objects. A longitudinal survey study (WebAIM) was conducted in 2013 and 2018 to assess the technology use among people with visual disabilities. The number of respondents using iOS increased from 43.1% to 64.3%, and the respondents using Android also increased from 18.1% to 23.8%. Nearly 61% of the respondents using iOS and 35.7% of the respondents using Android reported that they used voice assistants (e.g., VoiceOver, Siri) *very* or *somewhat frequently*.

Many users with visual disabilities take advantage of assistive applications (apps) such as Microsoft's Seeing AI (Dockery and Krzystolik). The Seeing AI app (powered by an artificial intelligence technology) can identify people, colors, currency, scenes, objects, and texts, and then audibly describe them for users with visual disabilities. The Seeing AI app is fully accessible with VoiceOver. Yet, when VoiceOver is on, standard touchscreen gestures (for sighted users) will have different effects, and additional gestures will become available to operate the iPhone and apps.

However, little is known about how users with visual disabilities understand and use the voice assistant features. Wong and Tan conducted a single case study to investigate how an individual with visual impairment, named Bill (aged 45 years) learned and used an iPhone. Bill encountered challenges with some apps that are not running well with VoiceOver. When errors occurred, there were no verbal prompts to help Bill to exit from the app. Celusnak, as a blind rehabilitation specialist, acknowledged that VoiceOver is a useful tool for users with visual disabilities in using an iPhone. Yet, Celusnak also argued that the learning curve is steep and very challenging; for example, the gesture command *Split Tap* was suggested to be one of the most difficult concepts for those with visual disabilities to understand and execute properly. Leporini et al. conducted a user study with a large sample (n = 55 participants with blindness) to examine the user experience of VoiceOver and found a range of usability problems; however, they merely relied on a passive approach, i.e., an online survey. There are only a handful of publications on user experience with VoiceOver (Grussenmeyer and Folmer; Park et al.; Smaradottir et al.). There is still a lack of in-depth understanding of what circumstances cause poor user experience associated with voice assistants, especially in terms of finger gesture commands, voice commands, and relevant user interfaces.

To address the knowledge gap, this study conducted observations on people with visual disabilities interacting with VoiceOver and the Seeing AI app accessible through VoiceOver.

Methods

Participants

A convenience sample of eight individuals with visual disabilities participated in this study. The inclusion criteria included 18 years of age or older and visual acuity worse than 20/70 with the best possible correction (World Health Organization). Table 1 shows the detailed characteristics of the participants.

Table 1. Characteristics of the participants

Participant Demographics	**N = 8**
Visual acuity - Between 20/200 and 20/400	5
Visual acuity - Between 20/400 and 20/1200	2
Visual acuity - Less than 20/1200	1
Age	65.71±12.33 (years)
Gender - Male	3
Gender - Female	5
Race/Ethnicity - African American	5
Race/Ethnicity - European American	3
Education - High school or equivalent	3
Education - Associate	4
Education - Masters	1
Do you use a smartphone?	Yes = 7, No = 1
How long have you used the smartphone?	4.90±3.20 (years)
How frequently do you use the smartphone?	Very frequently = 4 Frequently = 1 Very rarely = 2 Never = 1

Materials

An iPhone 12 mini was used in teaching participants how to use VoiceOver and the Seeing AI app, and their interactions were video recorded for further analysis. The tutorial was prepared based on the official User Guide of iPhone VoiceOver (Apple) and Microsoft's Seeing AI app (Microsoft). None of the participants had prior experience with an iPhone and the Seeing AI. User experience was measured with a System Usability Scale (SUS) – a 10-item questionnaire with a five-point Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Procedure

A one-to-one individual tutorial was offered to each participant in which they learned how to use VoiceOver and the Seeing AI app. Afterward, participants completed a System Usability Scale (SUS) questionnaire to measure user experience with VoiceOver and the Seeing AI app. Participants' responses to the SUS questionnaire were analyzed using descriptive statistics. The video recordings of user interactions were analyzed using open, axial, and selective coding.

Results

SUS of VoiceOver

The mean score of the participants' summed responses was 52.5±16.11, which is considered a poor user experience (Bangor et al.). The responses were broken down into positive and negative items (See Figures 1 and 2). The mean score of the five positive items was 3.52±0.52. The participants appreciated it that VoiceOver was equipped with various functions, and they would like to use it frequently. However, their confidence was low.

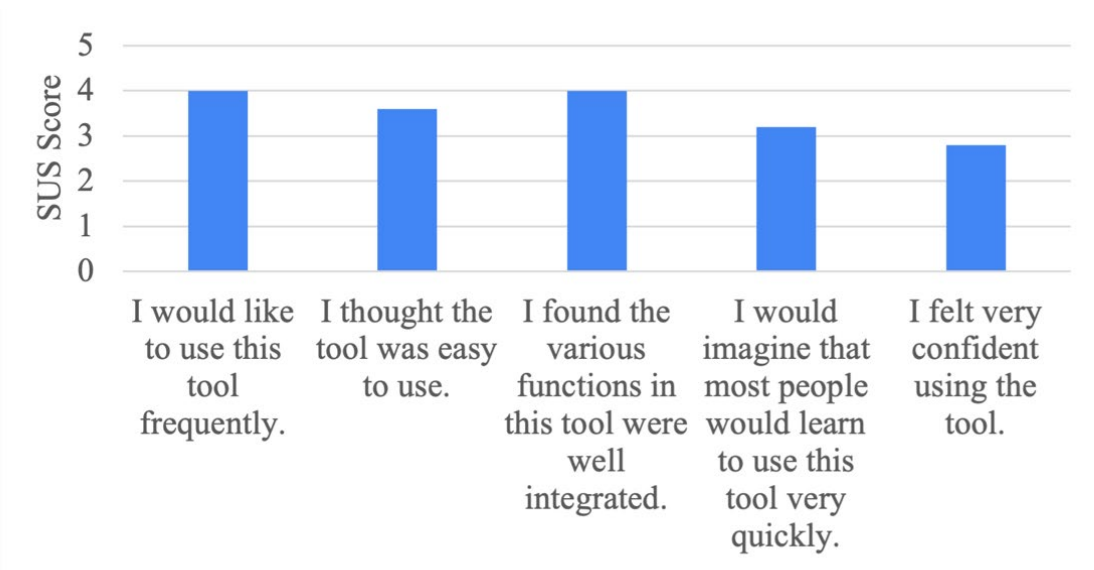


Fig. 1. Mean scores of positive items for VoiceOver.

The mean score of the five negative items was 3.32±1.06. The participants perceived that user interfaces of VoiceOver were well designed to ensure consistency but there were many things for them to learn to operate it properly, such that they would likely rely on a technical person.

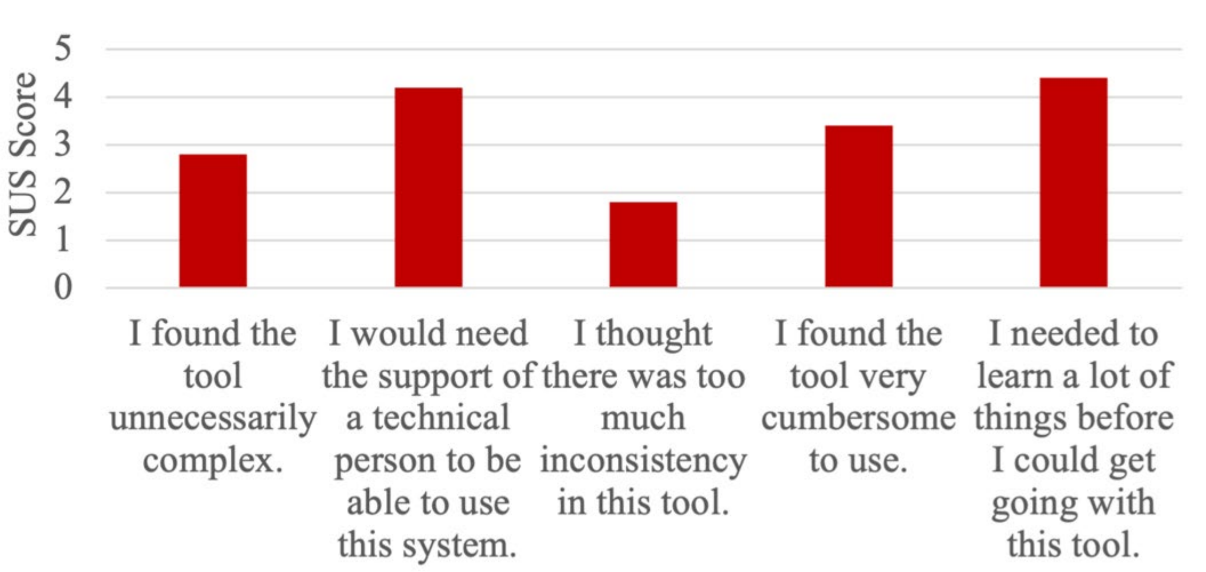


Fig. 2. Mean scores of negative items for VoiceOver.

SUS of the Seeing AI App

The mean score of the participants' summed responses was 74.50±12.17, which is considered a good user experience (Bangor et al.). The responses were broken down into positive and negative items (See Figures 3 and 4). The mean score of the five positive items was 4.24±0.46. The mean score of the five negative items was 2.28±1.18. The participants were satisfied with the Seeing AI app overall.

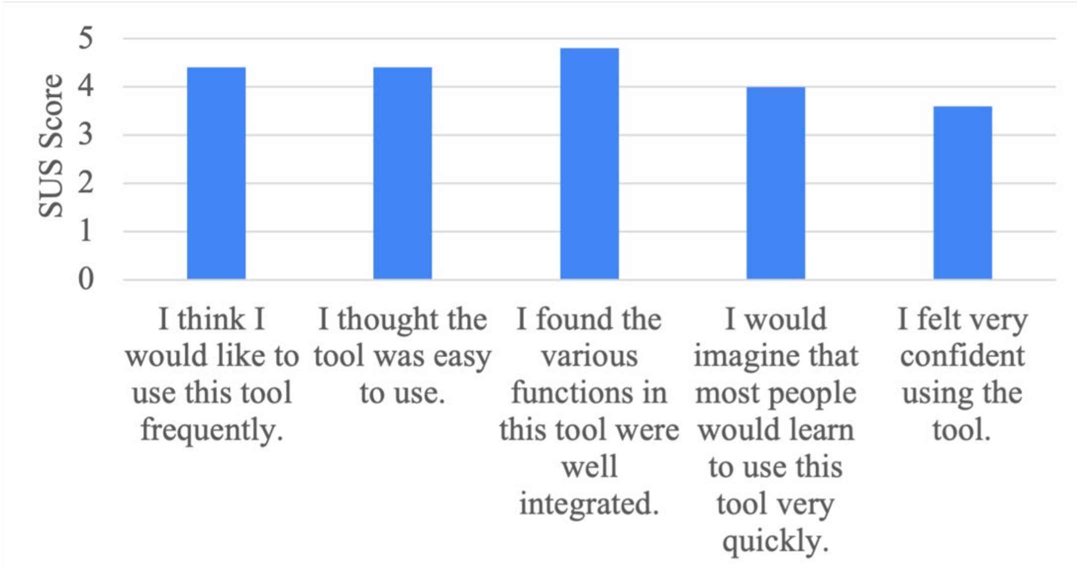


Fig. 3. Mean scores of positive items for Seeing AI.

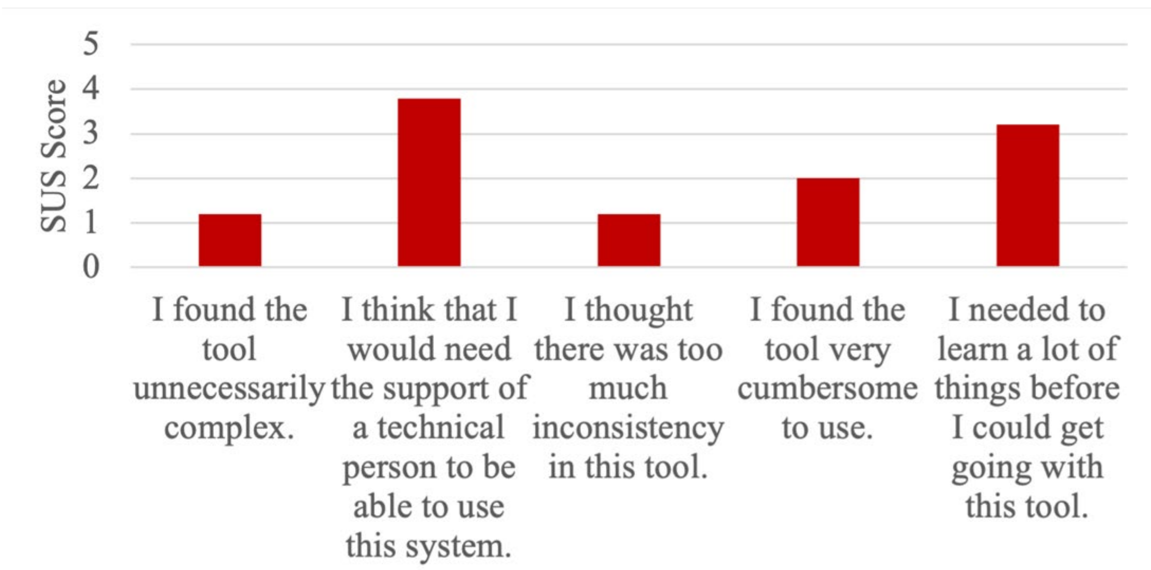


Fig. 4. Mean scores of negative items for Seeing AI.

User Interactions with VoiceOver

The participants were observed to have poor user experience, the cases of which were summarized under three categories: gesture commands, voice commands, and user interfaces (UI) (see Tables 2, 3, and 4). Tables 2 - 4 include adequate recommendations to address each case of poor user experience.

Table 2. Gesture-related incidents and recommendations.

Incidents	**Recommendations**
Scroll down one page: Wrong direction	Users should be given the option to switch the scrolling direction based on their preference.
Scroll down one page: The wrong number of fingers	Users should be given the option to reset the gesture command (e.g., one, two, or three fingers).
Scroll down one page: Fingers apart gradually	The iOS should distinguish the finger gestures for “zoom in” and “scroll down.” Users should be given the option to reset a minimum distance threshold between fingers.
Select, speak an item: Touch but not hold long enough	Once users touch an item, a status message should be provided, e.g., “*You tapped one of four buttons. There are three more*.”
Select the next item: Swipe very slowly	Users should be given the option to change the moving speed for the swiping gesture.
Select an item: Lack of mental models of spatial layouts	Users should be given the option to easily change spatial layouts (e.g., a calendar view for day, week, month, or year).
Quick actions menu: Press-and-hold for too short or long	Users should be given audio feedback. If they press and hold long enough to see the Quick Action menu, a beep sound is generated.
Double tap: Taping on wrong targets	Users should be given voice advice (e.g., “*Double tap on the TIME and DATE texts to schedule an event*”).

Table 3. UI-related incidents and recommendations.

Incidents	**Recommendations**
Affordance: Against Fitts' Law	Fitts' Law should be applied to improve the user interfaces, e.g., a small *plus* sign is currently located at the top right of the screen, which is too far from the place where users enter a new schedule in the calendar. The *plus* sign should be relocated, or there should be an alternative way for users to enter a new schedule more easily.
Affordance: Irremovable notifications	Users should be given an alternative option (e.g., a physical button) to close the Notification Center.
Affordance: Inconsistent direction for a slider	User interfaces should be designed to be consistent with natural hand gestures, e.g., if a slider is designed to move horizontally (left and right), a gesture must be designed to move horizontally instead of vertically (up and down).
Graphic UI: Images/texts vs. clickable buttons	Images/texts and clickable buttons should be easily differed. For example, the VoiceOver should inform users whether interface components (e.g., images, texts, and buttons) are clickable.
Graphic UI: Low color contrast	A strong color contrast should be employed to distinguish between an active button and a disabled button.
Graphic UI: Identical buttons	Different user interface designs should be assigned to buttons for different functions.

Table 4. Voice-related incidents and recommendations

Incidents	**Recommendations**
Message: Unclear message	In case Siri does not understand users' voice commands, users should immediately be allowed to repeat it without saying "Hey, Siri" again.
Wake phrase "*Hey, Siri*": Overwhelmed with the wake phrase by repetition	Users should be given the option to keep having a conversation with Siri for a certain period without repeatedly using the wake phrase.
Voice commands: Unlisted commands	Siri should be able to guess voice commands that are not programmed yet (or users can easily add new commands).

User Interactions with the Seeing AI App

As compared to VoiceOver, the Seeing AI app resulted in a fewer number of poor user experience cases (see Table 5).

Table 5. The app-related incidents and recommendations.

Incidents	**Recommendations**
Camera covered by fingers	The app should be designed to alert users when the camera is blocked (e.g., beep sound).
Camera not aiming at the entire paragraph	Users should receive audible notifications when the entire paragraph is not captured.

Discussion

The participants had a better user experience with the Seeing AI app (mean of overall SUS scores > 74) compared to VoiceOver (< 53). The SUS score of 70 is considered a cut-off point for determining good user experience (Bangor et al.). Given the norm, it can be interpreted that the participants in this study encountered many usability and accessibility problems while using VoiceOver. The observation in this study also confirmed qualitatively that VoiceOver caused more cases of poor user experience. Hence, this study suggested alternative designs to address the poor user experience (see Tables 2, 3, and 4).

The research findings also infer that the Apple's User Guide has limitations. It does not offer alternative formats for users with visual disabilities as it simply consists of texts and images. Furthermore, it does not convey detailed instructions about "dynamic" finger gestures. For example, the participants were not able to learn "how long" they were supposed to press and hold their finger on an app icon in order to have a Quick Action menu (i.e., a pop-up menu with shortcuts for additional actions). Due to such an incomplete user guide, the participants ended up with holding for too long, such that an "X" mark was shown next to the app icon, ready to be deleted. If they accidentally touched the "X" mark, the app would be deleted against their will.

The primary responsibility of a user guide (or manual) is to help users learn how to use a new application; however, a user guide becomes useless if a user cannot understand the instructions as they are poorly written. Regardless of the product quality, users are likely to be unsatisfied due to the lack of understanding about the product. It is well documented that there is a significant relationship between the quality of the user guide and the perceived product quality (Gök et al.). Allwood and Kalén, for example, observed that a user-friendly user guide could help users to spend less time on the tasks and make fewer errors. Byrne tested *Iconic Linkage* (i.e., the use of the same words to present the same information multiple times in a text) to improve the usability of a user guidebook. He revealed that Iconic Linkage contributed to a shorter time to complete a task, improved retention of information, fewer mistakes, and higher user satisfaction with the product. Besides the quality of instructions per se, the medium to deliver the instructions could be another critical factor for good usability (Gök et al.). Alexander compared the effectiveness of print- versus video-based user guides and found that the video format was more likely to result in positive outcomes as users who were given the video format made fewer errors and completed tasks with more accuracy. For users with visual disabilities, alternative formats are recommended such as tactile, haptic, or audio formats for user guides.

The participants also showed a lack of understanding of how voice commands worked. With Siri, users can use many accessibility features, make and receive phone calls, hear notifications, and so forth. The participants seemed to treat Siri as a human-like voice assistant. Hence, they kept talking to Siri without repeating the wake phrase "Hey Siri"; had a long conversation instead of a short voice command; and used voice commands that were not programmed in the Siri system (e.g., *"Hey, Siri. Go to 100.9*", which was a radio channel number). Ghosh et al. also administered the SUS survey with sighted users of Siri, and the mean of overall SUS scores was 54.17, which is greater than that of the participants with visual disabilities in this study. A voice assistant such as Siri is critical for users with visual disabilities to fully use various features of smart technologies, such that more user-friendly voice user interfaces should be provided (e.g., more flexible and natural voice commands).

The participants showed a higher level of satisfaction with the Seeing AI app. As an assistive technology app, the Seeing AI app is also executed based on voice user interfaces as does VoiceOver. Yet, the Seeing AI app has more intuitive user interfaces, i.e., users simply use a built-in camera to capture what they want to identify, and it reads out loud for users, leading to a higher level of user satisfaction. It infers that VoiceOver should be redesigned to be equipped with better user-friendly interfaces and interactions for those with visual disabilities.

Conclusion

Today, many mainstream technologies are accessible to people with visual disabilities via assistive technologies such as Siri, VoiceOver, and Seeing AI app. The assistive technologies can help users with visual disabilities to control a system, obtain information, and communicate with others without barriers, leading to independence in everyday life. However, little is known about the user experience of those applications (Siri, VoiceOver, and Seeing AI). This study conducted in-person observations and found that participants with visual disabilities encountered many poor user experience cases. To address them, this study suggested a set of design recommendations.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1831969.

Works Cited

Alexander, Kara Poe. "The usability of print and online video instructions." *Technical Communication Quarterly* 22.3 (2013): 237-59. Print.

Allwood, Carl Martin, and Tomas Kalén. "Evaluating and improving the usability of a user manual." *Behaviour & Information Technology* 16.1 (1997): 43-57. Print.

Apple. "iPhone user guide." 2023. Web. February 03 2023.

Bangor, Aaron, Philip Kortum, and James Miller. "Determining what individual sus scores mean: Adding an adjective rating scale." *Journal of usability studies* 4.3 (2009): 114-23. Print.

Byrne, Jody. "Evaluating the effect of iconic linkage on the usability of software user guides." *Journal of technical writing and communication* 35.2 (2005): 155-78. Print.

Celusnak, Brian M. "Teaching the iPhone with voiceover accessibility to people with visual impairments." *Journal of Visual Impairment & Blindness* 110.5 (2016): 369-72. Print.

Centers for Disease Control and Prevention. "Fast facts of common eye disorders." 2022. Web. January 27 2022.

Crossland, Michael D, Rui S. Silva, and Macedo, Antonio F. "Smartphone, tablet computer and e-reader use by people with vision impairment." *Ophthalmic and Physiological Optics* 34.5 (2014): 552-57. Print.

Dockery, Dominique, and Magdalena Krzystolik. "The use of mobile applications as low-vision aids: A pilot study." *Rhode Island Medical Journal* 103.8 (2020): 69-72. Print.

Ghosh, Debjyoti, et. al. "Assessing the utility of the system usability scale for evaluating voice-based user interfaces." *Proceedings of the Sixth International Symposium of Chinese CHI*. 2018. Print.

Gök, Osman, Pervin Ersoy, and Gülmüş Börühan. “The effect of user manual quality on customer satisfaction: The mediating effect of perceived product quality.” *Journal of Product & Brand Management* (2019). Print.

Grussenmeyer, William, and Eelke Folmer. “Accessible touchscreen technology for people with visual impairments: A survey.” *ACM Transactions on Accessible Computing (TACCESS)* 9.2 (2017): 1-31. Print.

Kim, Hyung Nam. “User experience of mainstream and assistive technologies for people with visual impairments.” *Technology and Disability* 30.3 (2018): 127-33. Print.

Leporini, B., Buzzi, M. C., & Buzzi, M. “Interacting with mobile devices via Voiceover: Usability and accessibility issues”. *Proceedings of the 24th Australian Computer-Human Interaction Conference*. 2012. ACM. Print.

Microsoft. “Seeing AI in New Languages.” 2023. Web. February 03 2023.

Park, Kyudong, Taedong Goh, and Hyo-Jeong So. “Toward accessible mobile application design developing mobile application accessibility guidelines for people with visual impairment.” *Proceedings of Hci Korea*. 2014. 31-38. Print.

Smaradottir, B. F., Martinez, S. G. and Håland, J. A. “Evaluation of touchscreen assistive technology for visually disabled users.” *IEEE Symposium on Computers and Communications (ISCC)*. 2017. IEEE. Print.

Varma, Rohit, et al. “Visual impairment and blindness in adults in the United States: Demographic and geographic variations from 2015 to 2050.” *JAMA ophthalmology* 134.7 (2016): 802-09. Print.

WebAIM. “Survey of users with low vision #2 results.” 2018. Web. January 30 2023.

Wong, Meng Ee, and Stacey SK Tan. “Teaching the benefits of smart phone technology to blind consumers: Exploring the potential of the iPhone.” *Journal of Visual Impairment & Blindness* 106.10 (2012): 646-50. Print.

World Health Organization. “Change the Definition of Blindness.” 2008. Web. October 4 2022.

THE JOURNAL ON
TECHNOLOGY AND
PERSONS WITH
DISABILITIES

Analysis of a University System's Campus Disability Support Center Websites

Lesley Farmer, Arlene Ramos
California State University Long Beach
lesley.farmer@csulb.edu, arlene.ramos01@student.csulb.edu

Abstract

Post-secondary disability support centers (DCS) have an obligation to provide information that informs students with disabilities and their stakeholders. That that end, this study conducted a content analysis of a major public university systems' DCSs's websites in order to identify patterns and best practices of content matter contained in the campuses' DSC websites and the design of these websites. Findings focused on the DSC's characteristics, organizational decentralization, student services, and website features. The findings provide website models and processes for other DSCs to emulate.

Keywords

Disability support centers, websites, accessibility, universities, California State University, content, web design

Introduction

As students with disabilities transition to post-secondary education, they also need to transition to self-responsibility for requesting accommodations. One of the ways that students locate such disability support services is via the Internet. In that respect, post-secondary disability support centers (DCS) have an obligation to provide information that informs students with disabilities (SWD) and others who support those students such as faculty, staff, relatives, friends and employers. To that end, this study analyzed the websites of a major public university systems' DCSs in order to answer the research question: "What patterns and best practices of content matter exist in the state university system campus's disability support center websites?" The findings provide website models for other DSCs to emulate.

Literature Review

As high school graduates who have disabilities pursue post-secondary education, they have to assume responsibility for self-disclosing their disabilities in order to receive services. One of the ways that high school graduates seek information about institutional disability support is by searching post-secondary educational institutions' websites. Even this task can be daunting as the location of the information varies across institutions and may be difficult to find (Costello-Harris, 2019; Flink, 2019).

Likewise, the usability of the relevant website varies in quality across institutions in terms of interface and guidance (Jackson & Jones, 2014; Quinn et al., 2019; Silvis, Bothma & De Beer, 2019). When SWDs encounter a lack of usability for the above reasons, they may stop engaging with that website and look elsewhere, and may even not obtain disability support services (Gabel, Reid & Pearson, 2017; Meleo-Erwin et al., 2021). Since post-secondary education institution disabilities must comply with state regulations, differences in DSC websites can be attributed in part to state compliance regulations, but even within the same state, unevenness occurs (e.g., Gael, Reid and Pearson; 2017; Ratermann, 2017).

In sum, post-secondary education institution DSC websites constitute a vital means for SWD to locate and use disability support services. Therefore, those websites need to be easily found and navigated, providing relevant information throughout SWD's educational journeys.

Methodology

In order to answer the research question, content analysis of the DSC websites of the California State University system's 23 campuses. This system was chosen because it is the nation's most diverse four-year university system, serving almost a half million students each year. The websites were analyzed by a prior coordinator of one campus's Applied Disabilities Studies Certificate program and one of her student researchers. The two researchers calibrated their assessments to ensure consistency. Based on the literature review and preliminary examinate of sample campus DSC websites, the researchers developed the following variables to use for coding their content analysis: contact information, number and title of DSC staff, advisory group, supervisory organization's office, mission, services, campus enrollment, number of students receiving DSC services, disabilities of students served, types of accommodations offered, types of other interventions, student support groups, mentor programs, resources for students, resources for faculty/staff, links to services outside of DCS that help SWD, training, outreach efforts, public relations publications, events, ADA compliance information, policies and procedures, forms, surveys, data management system, assessments, audits. The websites were analyzed and coded during academic year 2022-2023, and reviewed at the end of the academic year to check for possible updates. Campuses were grouped as small (campus enrollment less than 12,000 students), medium (between 12,000 and 25,000 enrolled students), and large (more than 25,000 enrolled students) to determine if findings were significantly related to campus size. Content analysis of the campuses' DSC websites revealed several clusters of variables.

Findings

Center Organization

DSCs did not have a consistent title across the campuses, although all of the campuses used a combination of the following words: Access/-ible/-ility, Center, Disability, Office, Program, Resources, Services, Student, and Support. Likewise, DSCs were located in different parts of the campuses' organizational structure.

All of the DSCs published their mission statement. Ten DSCs mentioned academic accommodations, and nine mentioned campus collaboration or partnerships. Mission statements varied in terms of detail and scope (e.g., training, legalities, technology, facilities, equity).

All campus websites provided legal ADA compliance information; ten campuses listed information on the DSC website, and the rest listed the information on another campus website Similarly, all DSC websites provided information about their policies and procedures, including guidance for campus accessible events.

DSC operating hours also differed across campuses. During the weekday, DSCs were open between 8 and 13 hours, with a mean and mode of 9 hours. Only one campus was open in the evening, and it had the longest number of open hours. Some campuses had shorter summer hours; no campus DSCs were open on weekends.

The number of DSC staff varied across campuses: from 2 to 19, with a mean and median of 12 personnel. All campuses had a director with that title, but the other personnel held a number of different titles, sometimes performing the same function (although specific duties were not equally detailed on the website).

Eight campus DSCs mentioned their advisory committee, ranging from 10 to 22 members, with an average number of 15 members. Information about these committees varied in detail and currency (e.g., representation, function, accountability).

Four campus DSCs did not appear to have an online data management system. Ten campus DSCs used AIM, and two used Clockwork. One each used Point & Click or My Compass. It seems that two campuses used an in-house product.

Funding was not disclosed on the DSC websites. However, three campus DSC websites provided ways for users to donate to the DSC.

Student Population

Almost 21,000 students received DSC services within the campus system. The number of students served per campus ranged from 114 to 2086 (mean and median of 870, 1.8% to 15% range). Only one campus mentioned student demographics. The staff-to-students ratio ranged from 1:20 to 1:174 (1:78 mean, 1:65 (median). The most frequently cited disability was psychological (29%), and the next highest were learning disability (19%) and ADHD (18%).

Services

All campus DSC websites explained their process for student registration and requesting accommodation. All campuses provided testing, notetaking, facilities/classroom accommodations, course-related accommodations, assistive technology and alternative media, mobility; most offered priority registration. Additional accommodations mentioned: animals by seven DSCs, research/technical assistance by two, and pregnancy support by two. In addition, some DSCs offered interventions for their students (psychological counseling, academic coaching or tutoring, employment or workability programs, conflict resolution service, self-advocacy advice, time management advice).

Support groups were also mentioned. The federally-funded TRIO program for students from disadvantaged backgrounds was noted in 21 campus DSC websites. Five campus DSCs mentioned their support group for students with ASD, two mentioned a co-curricular group, two mentioned a peer mentoring program, and one sponsored a Disability Justice Culture Club.

All campuses provided instructions on using the DSC service. Twelve campus DSC websites did not mention other trainings. A few DSC websites provided additional online training for students (e.g., videos, self-paced tutorials, specialized workshops on careers, post-graduate education, self-advocacy). Several campus DSC websites provided online training for faculty (e.g., reference documents, videos, self-paced tutorials, on-demand DSC presentations). Eight campus DSC websites listed resources associated only with their services. Four campus DSC websites listed only their services or their campus resources, five listed only their services or local community resources. Three listed resources associated with their DSC, the campus, the local community, and national organizations. Most DSC websites included information about voting. One campus DSC each mentioned emergency procedures, publications about disabilities, recreational options, and study abroad opportunities.

Faculty too may need resources to help them support SWD. All campuses provided links to online resources to faculty staff: at the least, about accommodation requests. The amount and depth of information about accommodations and work with SWDs varied significantly across campuses. Some websites used a FAQ approach and others provided links to relevant information; two provided handbooks.

Campuses routinely provide online public relations (PR) materials. Four DSC websites included videos, three posted newsletters, two linked to Instagram, two had web pages for visitors or parents, and one provided a linked brochure. All DSCs included contact information.

However, thirteen DSC websites included no other explicit outreach effort. Two DSC websites provided information targeted to parents, three targeted to prospective students (general, transfer, Latinx), and one targeted guidance counselors. One DSC website included information for campus visitors in general, and two did outreach to the local community. Five campuses held disability awareness or recognition events, and one specifically encouraged community participation.

All campuses are supposed to be audited by the university system's disability office; three campus DSCs have been audited by that office since 2017. Three other campuses had campus internal audits or assessments; one was audited in 2021, one was audited annually until 2016, and one other has been audited annually to the present.

Table 1. CSU System Total Number of Students with Disability and Percent of Students with Each Disability.

Disability	**Total # of students with disability**	**% of students with disability**
Visual Limitation	441	0.02
Communication Disability	123	0.59
Mobility Limitation	1054	5.04
Learning Disability	4065	19.44
Deafness	590	2.82
Attention Deficit / Hyperactivity Disorder (AD/HD)	3856	18.44
Acquired Brain Injury	335	1.60
Psychological or Psychiatric Disability	6065	29.01
Autism Spectrum Disorder	1432	6.85
Temporary Disabilities	340	1.63
Other Functional Limitations	2605	12.46
TOTAL	20,906	8.90

Websites

Each DSC websites' interface was rated, based on the literature reviews' criteria for assessing websites in general and DCS websites (Jackson & Jones, 2014; Quinn et al., 2019; Silvis, Bothma & De Beer, 2019). Accessibility rating based on two website accessibility checkers: https://www.accessibilitychecker.org/ and https://accessible.com/accessscan. Main criteria (labeling, organization, navigation, searching, content) and specific criteria (images, consistency, clutter, controllability, error handling, help) were each rated on a scale of 1 poor to 5 excellent, then totaled and averaged. No website was completely ADA-compliant. Labeling was the most consistently good, and the searching feature was the least satisfactory. Images were the most varied feature.

Table 2a. Campus DSC Website Rating by Criteria (B to M).

	Bakersfield	Channel Islands	Chico	Dominguez Hills	East Bay	Fresno	Fullerton	Humboldt	Long Beach	Los Angeles	Maritime Academy	Monterey Bay
Accessibility	4	3	4	3	3	4.5	4	3	3	4	3	3
Labeling	4	4	5	4	4	4	3	4	3	3	3	3
Organization	3	2	2	2	2	3	2	4	3	3	2	3
Navigation	3	2	2	2	2	3	2	3	3	2	2	3
Search	1	1	1	1	1	1	1	4	2	1	1	1
Content	3	4	4	3	3	3	3	4	4	3	2	4
Subtotal	18	16	18	15	15	18.5	15	22	18	16	13	17
Images	1	2	1	1	1	3	2	3	2	2	1	2
Consistency	3	2	3	2	2	2	2	4	3	2	2	3
Clutter/ Minimal	4	3	2	3	3	3	3	4	3	3	3	4
Controllability	3	2	3	4	3	2	2	3	3	2	2	3
Error prevention/ recovery	3	2	3	3	2	3	3	3	2	2	1	3
Help	2	2	3	3	3	2	3	2	3	2	2	3
Subtotal	16	13	15	16	14	15	15	19	16	13	11	18
TOTAL	34	29	33	31	29	33.5	30	41	34	29	24	35

Table 2b. Campus DSC Website Rating by Criteria (N-S and Campus Averages).

	North ridge	Pom ona	Sacrame nto	San Bernar dino	San Diego	San Fran cisco	San José	San Luis Obispo	San Mar cos	Sono ma	Stani slaus	Campus Averages
Accessibility	3	4	3	4	2	3	3	3	3	3	4	3.15
Labeling	3	4	3	3	3	3	3	4	4	4	3	3.31
Organization	3	3	2	2	3	3	3	4	3	3	3	2.85
Navigation	3	3	2	2	3	3	3	4	4	3	3	2.92
Search	2	1	1	1	1	5	2	1	1	2	2	1.54
Content	3	3	3	3	3	4	3	4	4	3	3	3.23
Subtotal	17	18	14	15	15	20	17	20	19	18	18	17.00
Images	4	4	1	4	3	3	2	1	1	4	2	2.46
Consistency	3	2	2	2	3	3	3	3	3	3	3	2.69
Clutter/ Minimal	3	3	2	3	2	3	3	3	2	3	3	2.85
Controllability	2	2	2	2	2	2	2	2	2	2	2	2.08
Error prevention/ recovery	3	3	3	1	2	3	3	3	3	3	3	2.62
Help	2	2	2	1	2	3	2	2	2	2	2	2.08
Subtotal 2	17	16	12	13	14	17	15	14	13	17	15	14.77
TOTAL	34	34	26	28	29	37	32	34	32	35	33	31.77

Discussion

Organization Decentralization

This university system uses a decentralized governance approach, which is demonstrated by the variety of titles in the program and its personnel. Such differences can be confusing for high school graduates with disabilities who are trying to decide which campus to enter. On the local level, some program titles (e.g., Student Life or Campus Services) might be so generic that the student might have difficulty identifying DSCs (Ratermann, 2017). Some variations appear to be due to high-level administrative decisions and priorities, beyond the scope of the DSC's control (e.g., fit in the campus's organizational structure, operating hours), which can impact their services.

Staffing also reflected campus-specific allocation of resources, the most obvious example being the ratio of staff to SWD served, which was significantly correlated (p=.028) with the size of the student population. While the proportion of students being served by their DSC is not correlated with the campus size, it was somewhat related to the extent of outreach reflected in websites. In most cases, staff were not listed has having specialties; SWDs might be more likely to register for services, particularly upon entering the university if more outreach is evidenced and a specialty point person was identified (Costell-Harris, 2019).

Student Services

In terms of information provided, all campuses listed staff, services, resources, and legal information, and associated policies and procedures to that end. Some accommodations are associated with other campus offerings such as housing, programs, and support offices. A few campus websites noted SWD-focused or campus-side social groups. Which can help SWDs adjust to campus life.

As part of accommodations, DSCs provide training on using their services, but only a few websites included videos and self-paced tutorials, either for students or for faculty. Moreover, few websites advertised topical training such as self-advocacy or post-graduation options.

SWD and other stakeholders typically appreciate having online resources to support students' academic journey. While DSC websites did link to helpful campus resources, only five mentioned community resources, which limits SWDs' preparation for life-long success.

Website Technical Highlights

Confirming Quinn et al. (2019) findings, the system's DSC websites were uneven in terms of their degree of accessibility. Most websites were nominally accessible, although no one was issue-free, and one campus website had several broken links, lack of alt-text images, and inadequate contrast ratio; this finding was disappointing. It was not determined who controlled the website development and maintenance, so the issue might stem from campus technology staffing and protocols. Most website users expect a search tool in order to locate the information they want (Jackson & Jones, 2014; Quinn et al., 2019). However, few of the campuses' DSC websites had an internal search feature, and most help features minimal;' even information pop-up dialogue boxes would help users get clarification or tips. None of the websites modelled best practices of including a variety of relevant images that reflected their clientele, and a few had few or no images as well, which resulted in a less interesting website.

Conclusion

This investigation of one university systems' DSC website provides insights into the kind of information provided, and the quality and relevance of that information, based on their comparisons. No one DSC was ideal, and all were adequate, but some small changes could improve most websites without too much trouble.

For instance, the university systems' decentralized governance led to substantial differences in the website's location, design and content. A system-side website "template" could be developed, with some core information that applies to all campuses (including national resources) and areas for campus-specific information (e.g., housing accommodations). The web interface (e.g., layout, organization, navigation) must be clear and intuitively obvious. In any case, DSCs need to provide basic information about their services, procedures, and resources. They need to address accommodations and training for the entire campus community, not just students. They also need to show how the entire campus supports SWDs through partnerships with various service units on campus, within the community, and ideally nationally.

DSC staff and administrators can also use this study to examine their own DSC structures and services, such as staff to student ratio and percentage of students being served, in order to assess workloads or to expand their services and outreach efforts.

A few limitations of the study should be noted. Only one university system was examined, which limits some generalizability. Website content was only as accurate as the DSC provided (e.g., staff designations, amount of information, currency of websites).

As hinted above, having SWDs share their experiences with the websites would provide important insights, especially if they compared DSC websites within the system, which was beyond the scope of this study. Those DSCs with websites that were mentioned as good practices could be interviewed to determine how they designed and developed them as a means to identify effective planning and implementation factors. Also mentioned as another study is comparing other post-secondary institutions' DSC websites. While not within the scope of this study, having SWDs and other stakeholders share their experiences using the DSC website could reveal important gaps and confusing points as well as useful information, which can then facilitate website areas for improvement.

In the final analysis, DSC websites serve as a communication tool to support SDWs and their stakeholders. At the campus level, the DSC's website also serves as an indicator of that institution's value and support of that DSC, so it is in the interest of top-level administrators to ensure that their DSC's website is the best possible.

Works Cited

Costello-Harris, V. A. (2019). Evidence of inclusion on college websites. *Journal of Postsecondary Education and Disability*, *32*(3), 263-278.

Gabel, S. L., Reid, D. P., & Pearson, H. (2017). Disability, diversity, and higher education. In *Disability as Diversity in Higher Education* (pp. 170-184). Routledge.

Jackson, D. L., & Jones, S. J. (2014). A virtual commitment: Disability services information on public community college websites. *Journal of Postsecondary Education and Disability*, *27*(2), 129-138.

Meleo-Erwin, Z., Kollia, B., Fera, J., Jahren, A., & Basch, C. (2021). Online support information for students with disabilities in colleges and universities during the COVID-19 pandemic. *Disability and Health Journal*, *14*(1), 101013. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7543900/

Quinn, S., Belmonte, A., Davis, E., Gardewine, A., & Madewell, G. (2019). Access [dis] abled. *Disability Studies Quarterly*, *39*(4). https://dsq-sds.org/index.php/dsq/article/view/6603/5466

Raterman, K. (2017). The rhetoric of disability: An analysis of the language of university disability service centers. Master's thesis, Chapman University.

2024
THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

Accessible Informed Consent Process in Interactive ASL Apps

Hannah Benjamin[1], Natnail Tolossa[2], Michaela Brandt[3], Ben Kosa[4], Poorna Kushalnagar[3], Raja Kushalnagar[3]
New Jersey Institute of Technology[1], Rochester Institute of Technology[2], Gallaudet University[3], University of Washington[4]
hb279@njit.edu, ntt7381@rit.edu, michaela.brandt@gallaudet.edu, bkosa2@cs.washington.edu, poorna.kushalnagar@gallaudet.edu, raja.kushalnagar@gallaudet.edu

Abstract

Since informed consent became a mandatory measure in medical research, research participants have greater protections against research-related harm and exploitation. However, the information provided is often in written English. This creates a significant language barrier for deaf and hard of hearing people who use sign language as their primary means of communication. Additionally, hearing researchers, who make up the majority (NSF, 2017), are often less inclined to include deaf individuals in research due to the added work that is necessary to reduce the communication barrier between researchers and deaf participants. To address these issues, our research leveraged machine learning and artificial intelligence-based technology to test the usability of a user-centered and low-resource informed consent app-based toolkit. This toolkit allows researchers to easily provide interactive informed consent content entirely in American Sign Language. Building on the work of Kosa et al. (2023), we found that deaf people considered the app-based informed consent process to be accessible when completed entirely in ASL. This finding indicates the continued development of this technology would increase accessibility for the signing deaf community. This technology could be used by researchers to diversify their samples, improving the quality and broad applicability of the results of their research.

Keywords

Deaf and hard of hearing, machine learning, informed consent, American Sign Language, sign language recognition

Introduction

This research tested technology that aims to reduce this communication barrier to create a more inclusive and accessible research environment for both Deaf individuals and researchers. By using machine learning models and artificial intelligence to capture and recognize sign language, we are able to provide a tablet-based application that uses American Sign Language (ASL) throughout the informed consent process. This application, the ASL Consent App, uses ASL videos to explain the informed consent process and participants are also able to respond using ASL. Before we presented the ASL-Consent App to participants, we made significant changes from the initial iteration of the Kosa et al. (2023) version of the application to improve usability.

To evaluate the overall usability of our ASL Consent App, we conducted two rounds of testing with Deaf and Hard of Hearing participants who used ASL (members of the Deaf community). The first round of participants were primarily senior citizens over the age of 65 whereas the second round of participants were primarily people of color with a diverse age range, but a majority were under 50 years old. Though the feedback from both groups was generally positive, their perception and expectations of the ASL Consent App differed. Senior citizens were sometimes unsure how to interact with the tablet in that they would center their bodies in front of the tablet instead of the camera or they would forget that the machine learning model only recognized a very specific set of signs. The feedback from the second round of participants indicated that they felt the app was easy to learn and easy to use, however some commented that those without experience with technology may require more user training which matched the experience of the senior citizens from the first round.

It was hypothesized that presenting the informed consent process in a culturally and linguistically appropriate manner for the signing deaf community would allow deaf individuals to have more autonomy as research participants as well as assist researchers in reducing the barriers that often prevent the deaf community from participating in research studies. The survey findings indicated this can be achieved through the development of accessible, application-based technology.

Deaf and Hard of Hearing Challenges in Understanding Consent

The Deaf community, comprised of deaf and hard of hearing sign language users, is a small population that is often at risk for marginalization (Sanfacon, Leffers, Miller, Stabbe, DeWindt, Wagner, & Kushalnagar, 2020; Kushalnagar, Reesman, Holcomb, & Ryan, 2019; Kushalnagar & Miller, 2019; NIH, 2022). Since English is often a second language for deaf ASL users, literacy is low and health literacy is also low due to the lack of accessible language (Anderson et. al. 2020). Informed consent in the signing deaf community is not achievable if the information is only provided in written English, a language that many deaf individuals consider their second language (Mckee et al. 2013). While the Deaf community is considered a population that experiences health disparities because of their disability status (Pérez-Stable 2023), they are not necessarily considered a vulnerable population requiring additional protections according to the National Institute of Health. To ensure deaf and hard-of-hearing sign language users are fully able to make decisions regarding informed consent, it is necessary to ensure equitable access to informed consent content in their primary language, ASL.

It is necessary to understand the importance of Community Based Participatory Research (CBPR) because developing assistive technologies for the Deaf community requires collaboration between researchers and Deaf individuals. The goal of CBPR is to both provide procedural information in accessible language as well as make clear that the research being performed is beneficial to the Deaf community (Singleton et al. 2014). Second, investigating Sign Language Recognition (SLR) technologies that are currently in development can provide a foundation to build upon further research (Papastratis et al. 2021). The work of Anderson et al. (2018) focused on how to provide social equality for Deaf participants in qualitative research. Accessible recruitment, sampling, data collection, and data analysis procedures must be utilized to conduct ethical and accurate research with the Deaf community. Data collection should be performed in the participants' primary language to reduce translation bias and increase translation accuracy. The study by Anderson et al. (2020) states that "the deaf community is one of the most understudied in the research community".

Sign Language Recognition

Sign Language Recognition (SLR) refers to the ability of machines to recognize sign language, which allows for Sign Language Translation (SLT): the ability for machines to translate from sign language to a spoken language, like English. Although meaning-to-meaning Sentence-Level SLT is not yet possible, recent breakthroughs in machine learning have made Individual Sign Language Recognition (ISLR) possible (Desai et al., 2023). One potential application of ISLR is what Kosa et al (2023) coins as Sign Language Interactability, which describes allowing users to interact with technology through sign language.

Allowing Deaf and Hard of Hearing participants who use ASL to navigate and sign the Informed Consent process may help make the process more user friendly. A preliminary study with a prototyped ASL Informed Consent Process using Sign Language Interactability in Kosa et al (2023) shows that Sign Language Interaction has great promise in doing this. However, the preliminary study had a limited sample size of 14 participants that did not reflect the diversity of the broader signing deaf community, a subpar user interface for their prototype, and an extremely limited user study procedure that took place over Zoom. Participants were only allowed to watch a recording of the app being used without being able to use it themselves, meaning responses from participants didn’t fully reflect the usability of the app as their experience was indirect. One of the aims of this paper is to re-evaluate the usability of Sign Language Interactability in the ASL Informed Consent Process with an improved user interface.

Discussion

Development

In developing our ASL Informed Consent App, we based the core design on previous work done by our team (Kosa et al, 2023), but have iteratively made improvements to the user interface, features, and backend of the app. In the previous study, participants gave feedback that having the ASL informed consent process on an iPhone screen was too small. We incorporated this feedback into our current design by developing our ASL Consent App for the iPad. Participants also suggested the addition of an English transcript to supplement the ASL videos, which is consistent with previous work that evaluated how users prefer to view sign language videos (Willis et al., 2019).

Using the designs that these previous works developed, we incorporated the Multimodal Visual Languages User Interface (M3UI) into the design for our app, which found that users prefer to view sign language content alongside an English transcript that automatically highlights the English text in sync with the sign language that is being shown in the video. Participants in Kosa et al. also gave feedback regarding improvement in navigation feedback (e.g. breadcrumbs) and how it wasn’t obvious how to use the novel ASL Interactability feature. We addressed the lack of navigation feedback by adding a sequential navigation bar that overviews every section in the informed consent process and shows what sections the participant has completed, which section they are on, and how many more they have left to do before they are done.

Participants sometimes had to ask for help during the middle of the ASL informed consent process because they weren't sure what to do next at the end of a section or how to navigate in ASL due to it being a novel concept. We addressed this issue in our current iteration of the ASL Consent App by adding an onboarding process that demonstrates how ASL Interactability works in the app. The previous iteration of the ASL Consent App in Kosa et al. had a separate section for signing the digital informed consent form that required the user to record themselves signing their full name, which would be stored in a secure database that could be viewed anytime as proof of signature. In their user study, Kosa et al. received feedback that participants found signing their full name felt clunky, and so at the end of our ASL Informed Consent process, the user only needs to sign "CONSENT" in ASL (which equates to "I Consent" in English) to provide their signature.

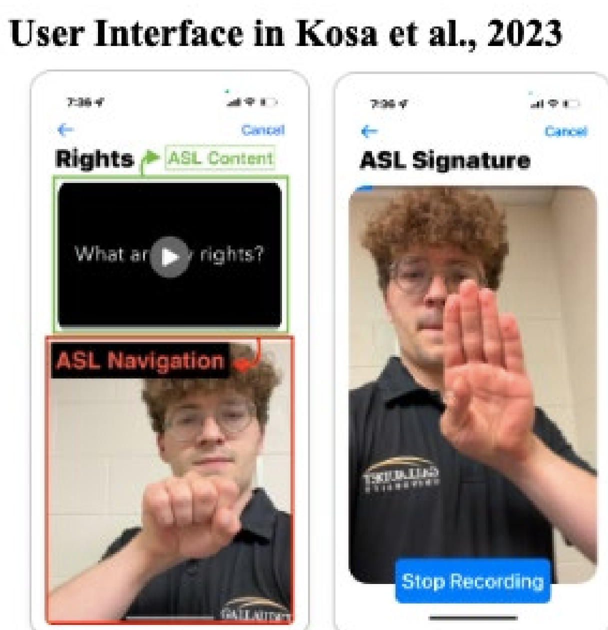


Fig. 1. ASL Informed Consent User Interface in Kosa et al., 2023.

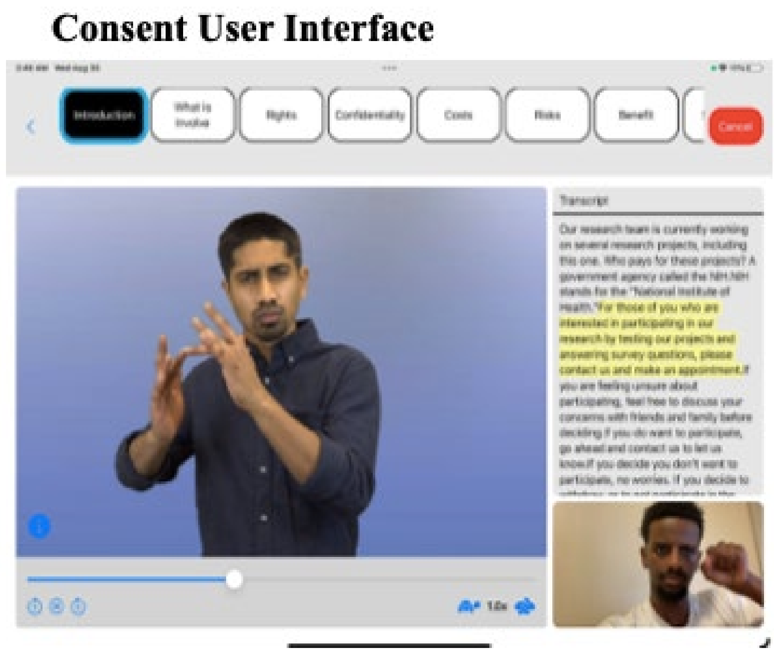


Fig. 2. Current ASL Informed Consent User Interface.

Method and Evaluation

All participants were deaf or hard-of-hearing fluent users of American Sign Language. The first round of participants was recruited from the Deaf Seniors of America (DSA) conference in Hollywood, Florida. The second round of participants were recruited from the National Deaf People of Color (NDPC) conference at Gallaudet University in Washington, D.C. Each participant worked with two iPads. We performed user testing using iPads that had the ASL consent app preloaded onto them. One iPad was used to collect informed consent, demographic information, and user feedback via a Qualtrics survey. The other iPad had the ASL consent app with which participants would interact with. Participants were given brief user training on how to use the application before user testing began. Participants were trained to avoid extraneous hand movements while using the app and to sign clearly when engaging with the receptive camera. Additionally, researchers found it necessary to explain to participants that while this application used artificial intelligence, they should consider the AI to be in its language-learning infancy.

This explanation alleviated concerns about the way artificial intelligence was being used and encouraged the participants to sign intentionally and clearly. Once briefed, participants would interact with the ASL consent app. After the participants completed testing the app, they then took the System Usability Survey (SUS) to collect feedback. The SUS consists of 10 questions that are answered using a Likert scale.

We learned from our experience with participants at the DSA that it may be helpful to include a video prompt for each section of the app to prompt the user to move forward or backward through the app by signing “Yes” or “Back”. We added this video prompting feature for the second round of testing at the NDPC conference. In addition to this, we added a text box and video recording option to the SUS survey so that participants could explain why they gave their rating for each SUS question. This allowed us to collect both quantitative and qualitative user feedback.

Results

Data was collected from a total of 34 participants, 14 participants from the DSA conference and 20 from the NDPC conference. Demographic data such as age and education level were collected. DSA participant ages ranged from 35-74 and NDPC participant ages ranged from 23-61 for an overall age range of 23-74. Education levels were sorted into the categories of high school or less, college graduate, and postgraduate. A table of demographic information for the DSA participants and NDPC participants is provided in Figure 3 and Figure 4, respectively.

Table 1. DSA Participant Demographic Data.

Participant	**Age Group**	**Highest Degree Completed**
P1	65-74	College Graduate
P2	65-74	High School or Less
P3	65-74	High School or Less

Participant	Age Group	Highest Degree Completed
P4	65-74	High School or Less
P5	65-74	College Graduate
P6	50-64	College Graduate
P7	65-74	College Graduate
P8	65-74	College Graduate
P9	65-74	College Graduate
P10	50-64	College Graduate
P11	65-74	College Graduate
P12	35-49	College Graduate
P13	65-74	College Graduate
P14	65-74	College Graduate

Table 2. NDPC Participant Demographic Data.

Participant	Age Group	Highest Degree Completed
P1	50-64	Postgraduate
P2	35-49	College Graduate
P3	35-49	High School or Less
P4	35-49	Postgraduate
P5	35-49	Postgraduate
P6	18-34	College Graduate
P7	18-34	College Graduate
P8	18-34	Postgraduate
P9	18-34	College Graduate
P10	50-64	College Graduate
P11	50-64	Postgraduate
P12	35-49	Postgraduate
P13	18-34	College Graduate
P14	50-64	High School or Less

Participant	Age Group	Highest Degree Completed
P15	35-49	College Graduate
P16	18-34	College Graduate
P17	18-34	Postgraduate
P18	35-49	College Graduate
P19	35-49	College Graduate
P20	50-64	College Graduate

Discussion

Quantitative Analysis

The System Usability Scale (SUS) was used to evaluate the usability of the ASL Consent App. It is important to understand that the SUS can only determine if a system is usable or not usable. It cannot determine why the system is usable or not usable. This can only be discovered through open ended questions and qualitative analysis. The SUS consists of 10 questions that are answered using a Likert scale rating from 1 to 5. A rating of 1 means “strongly disagree” and a rating of 5 means “strongly agree”. The overall SUS score is calculated using the formula by Brooke (1995) and scores range from 0-100. According to Bangor et al. (2008), the average SUS score is 70.14 and systems that score below 70 “should be considered candidates for increased scrutiny and continued improvement”. SUS scores above 70 are considered “passable” and “truly superior” systems have scores above 90 (Bangor et al. 2008).

The average SUS score across participants recruited from the DSA was 71.96 (s = 16.67). The average SUS score across participants recruited from NDPC was 71.25 (s = 17.65). The average SUS score across all participants from both the DSA and NDPC was 71.54 (s = 17.02). Given that the ASL Consent App is in its first iteration of user testing, an average SUS score of 71.54 is a sufficient indicator that our prototype is worthy of further development. Since the development of the ASL Consent App will involve many iterative versions that will be continually tested, the SUS scores of each iteration can be used as one metric to evaluate the progression of the system (Bangor et al. 2008). A table of the SUS scores for the DSA participants and NDPC participants are provided in Figure 5 and Figure 6, respectively.

Table 3. DSA Participants' SUS Scores.

Participant	SUS Score
P1	67.5
P2	65
P3	97.5
P4	50
P5	47.5
P6	80
P7	90
P8	82.5
P9	60
P10	50
P11	62.5
P12	87.5
P13	90
P14	77.5

Table 4. NDPC Participants' SUS Scores.

Participant	SUS Score
P1	95
P2	75
P3	90
P4	100
P5	72.5

Participant	SUS Score
P6	42.5
P7	75
P8	65
P9	32.5
P10	57.5
P11	95
P12	70
P13	87.5
P14	70
P15	85
P16	70
P17	77.5
P18	55
P19	50
P20	60

A one-way ANOVA test was used to determine whether there was a significant difference between the average SUS scores of different age groups. The chart in Figure 7 provides a visualization of the average scores between age groups. The age groups were divided into the categories of 18-34, 35-49, 50-64, and 65-74. The one-way ANOVA test revealed a p-value of 0.616 which indicates that there is not a significant difference between the average SUS scores of each age group. A one-way ANOVA test was also used to determine whether there was a significant difference between the average SUS scores for different education levels. The sample was divided into three groups according to the highest level of education completed by each participant. These groups were categorized as “high school or less”, “college graduate”, and “postgraduate”. The chart in Figure 8 provides a visualization of the average scores between education levels. The one-way ANOVA test revealed a p-value of 0.136 which indicates that there is not a significant difference between the average SUS scores across education levels. A t-test was conducted to determine whether there was a significant difference between the scores of participants recruited from the DSA and participants recruited from the NDPC conference. The chart in Figure 9 provides a visualization of the average scores between participants from DSA and participants from NDPC. The t-test resulted in a p-value of 0.908 which indicates that there is not a significant difference between the two groups.

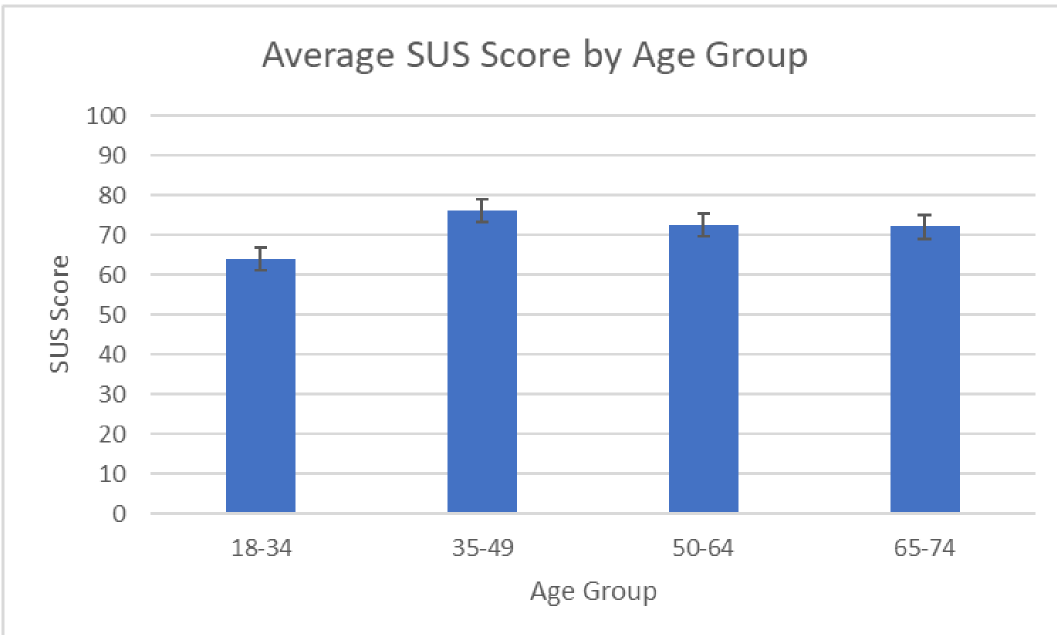


Fig. 3. Average SUS Score by Age Group.

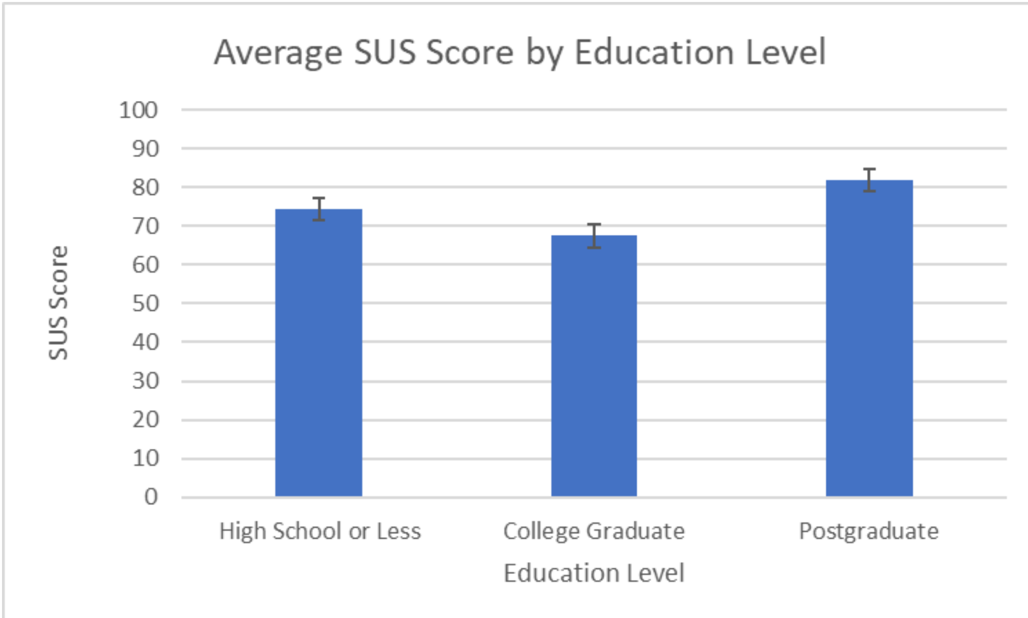


Fig. 4. Average SUS Score by Education Level.

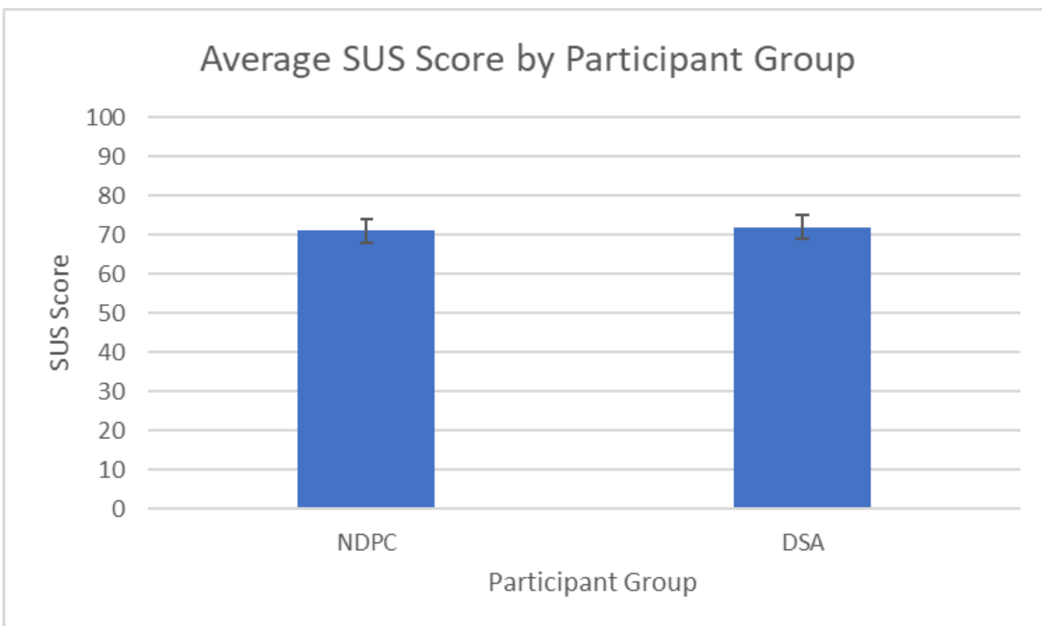


Fig. 5. Average SUS Score by Participant Group.

Qualitative Analysis (Thematic Analysis)

Participants from the NDPC conference were given the opportunity to provide feedback by typing their comments into a textbox or recording a video of themselves signing. Five participants utilized the textbox for feedback and participant feedback was generally positive. Four participants described the ASL Consent App as “easy to use”, “simple”, and “user-friendly”. One participant noted that “ASL instruction prior to use may help for those who are not tech savvy” which emphasizes the need for adequate user training. Developing a consistent system for user training may increase the overall usability of the ASL Consent App.

The participants at the DSA and NDPC conferences both commented that the positioning of the iPad camera made interacting with the application awkward because the body positioning was not intuitive. One NDPC participant said the camera being positioned on the side of the iPad when it was in landscape mode “made it a little bit awkward as I had to shift my body sideways to be centered in the viewfinder. If I didn’t do that and looked at myself, it looked like I was in the wrong position.”

The qualitative analysis of the open-ended feedback from participant video recordings at NDPC revealed four themes, seen in figure 9: the responsiveness, expansion, ease, and accessibility of the application. As predicted, the overall impression of the technology was positive and encouraged further development, with one user saying, “I am impressed that the app itself is very responsive, it is quick to catch whatever sign you throw at it.” 23% of respondents mentioned wanting to see this technology expanded and used in the real world. One participant mentioned how they could imagine this making other healthcare documentation, like living wills and Do Not Resuscitate paperwork, more accessible for DHH people who often do not understand these documents. The most common comment was on ease, with 50% of the feedback including this theme. Every participant who mentioned ease found the system clear, intuitive, and easy to use. One participant noted “because it’s smooth, easy… [and] interactive, yes, I quickly agree I would use this technology in the future.” Participants describing the system as easy to use indicates a likelihood to engage with this technology as a means to increase their access to healthcare information. 18% of participants commented explicitly on accessibility, all of them supporting the theme that this system will make documented healthcare information more accessible to DHH sign language users.

Table 5. Percentage of Themes Appearing in Participant Qualitative Feedback.

Theme	Percentage Present in Participant Responses
Responsiveness	23%
Expansion	23%
Ease	50%
Accessibility	18%

Conclusion

Our overall average SUS score of 71.5 indicates that our app development is moving in the right direction. The first round of participants were primarily senior citizens aged 65 years or older and this age group is generally less technologically literate than younger age groups. We followed the logic that if senior citizens were able to navigate the ASL consent app, then younger populations would likely be able to use our app easily. The participants recruited from the NDPC conference had a more diverse age range, with a majority of participants being under 50 years old. Overall, we were satisfied with the SUS feedback as a sign to continue this research.

Future Work and Limitations

There were a few limitations in the technology of our ASL consent app. One of the limitations is that the machine learning model used for sign language recognition was developed using healthy people as a model. Therefore, the app does not accommodate conditions which impact manual dexterity, such as arthritis. Some participants signed “Yes” using a handshape that had their thumb out rather than tucked in which would result in incorrect sign language recognition. In the future, the machine learning model needs to use more diverse datasets. Also, many participants used an ASL classifier for “Oh I see” (“OIC”) which could be considered as a sign to include in the machine learning model.

A prompting feature at the end of each video that asks users if they are ready to move to the next section would be helpful for users to understand when they should interact with the app. This feature would create a more conversational feel to the app. Additionally, the application could be modified to visually indicate when the camera is “looking” at the user, i.e., ready to receive prompts in ASL. This visual indication could take the form of animated eyes that open when the machine is receptive to ASL and closed when it is not.

As the application learns more language, this technology could be expanded and applied for industry interests beyond informed consent documents in research settings. This technology could theoretically be leveraged to provide sign language interactability for a variety of applications, making documents and apps accessible to sign language users in any context where paperwork is deployed.

Capturing sign language is necessary for creating datasets that can be used to train machine learning algorithms. These algorithms can then be used to recognize and translate sign language. The data collected from sign language capturing and translation can also be used to produce avatars or videos that use AI to generate a “human” animation that looks more realistic. These technologies can all be applied to fields such as healthcare, education, and general communication (Papastratis et al. 2021).

Acknowledgements

This work was funded by NSF REU Site Grant #2150429 awarded to Dr. Raja Kushalnagar, PI. This work was also supported by a grant from the National Institutes of Health (G08LM013797), awarded to Dr. Poorna Kushalnagar at the Center for Deaf Health Equity at Gallaudet University.

Works Cited

Anderson, M. L., et al. (2020). Deaf ACCESS: Adapting consent through community engagement and state-of-the-art simulation. *Journal of Deaf Studies and Deaf Education*, *25*(1), 115–125.

Anderson, M. L., et al. (2018). Deaf qualitative health research: Leveraging technology to conduct linguistically and sociopolitically appropriate methods of inquiry. Qualitative Health Research, 28(11), 1813–1824.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24(6), 574–594. https://doi.org/10.1080/10447310802205776

Brooke, J. (1995). SUS - A quick and dirty usability scale. https://www.researchgate.net/publication/228593520

Bragg, Danielle, et al. Sign language recognition, generation, and translation: An interdisciplinary perspective. *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*. 2019.

Desai, Aashaka, et al. ASL citizen: A community-sourced dataset for advancing isolated sign language recognition." arXiv preprint arXiv:2304.05934 (2023).

Gurbuz, S. Z., et al. (2020). A linguistic perspective on radar micro-doppler analysis of American Sign Language. *IEEE International Radar Conference (RADAR)*, Washington, DC, USA, 2020, pp. 232-237, doi:10.1109/RADAR42522.2020.9114818

James, T. G., et al (2022). They're not willing to accommodate deaf patients: Communication experiences of Deaf American Sign Language users in the emergency department. qualitative health research. *Qual Health Res*. 2022 *Jan;32(1)*:48-63. [doi: 10.1177/10497323211046238](doi: 10.1177/10497323211046238)

Kosa, B. S., et al (2023). ASL consent in the digital informed consent process. *Journal on Technology and Persons with Disabilities*. CSUN Assistive Technology Conference. 288-305.

Kushalnagar P, Reesman J, Holcomb T, Ryan C. Prevalence of anxiety or depression diagnosis in Deaf adults. *J Deaf Stud Deaf Educ*. 2019 Oct 1;24(4):378-385. https://doi.org/10.1093/deafed/enz017

Kushalnagar, P. & Miller, C. (2019). Health disparities among mid-to-older Deaf LGBTQ adults compared to mid-to-older Deaf non-LGBTQ adults in the United States. *Health Equity*. https://doi.org/10.1089/heq.2019.0009

Kushalnagar, P., Paludneviciene, R., & Kushalnagar, R. (2019). Video remote interpreting technology in health care: Cross-sectional study of Deaf patients' experiences. *JMIR Rehabilitation and Assistive Technologies*, *6(1)*, e13233.

Kushalnagar, P., Harris, R., Paludneviciene, R., & Hoglind, T. (2017). Health information national trends survey in American Sign Language (HINTS-ASL): Protocol for the cultural adaptation and linguistic validation of a national survey. *JMIR Research Protocols*, *6*(9), e172.

McKee, M., Schlehofer, D., & Thew, D. (2013). Ethical issues in conducting research with deaf populations. *American Journal of Public Health, 103(12)*, 2174–2178.

McKee, M., et al. (2012). Engaging the Deaf American sign language community: lessons from a community-based participatory research center. *Progress in Community Health Partnerships: Research, Education, and Action, 6(3)*, 321–329.

National Institute on Minority Health and Health Disparities. Health disparities experienced among persons living with disabilities. *National Institute of Health*, 2022. https://www.nimhd.nih.gov/funding/approved-concepts/2022/health-disparities-among-persons-with-disabilities.html

National Science Foundation, National Center for Science and Engineering Statistics. Doctorate recipients from U.S. universities: 2017." *Special Report NSF* 19-301. 2017 https://ncses.nsf.gov/pubs/nsf19301/

Papastratis, I., Chatzikonstantinou, C., Konstantinidis, D., Dimitropoulos, K., & Daras, P. (2021). Artificial intelligence technologies for Sign Language. *Sensors*, 21(17). https://doi.org/10.3390/s21175843

Pérez-Stable, E., Otto Valdez, R. (2023). Announcement of Decision to Designate People with Disabilities as a Population with Health Disparities. *National Institute of Health*. https://nimhd.nih.gov/about/directors-corner/messages/health-disparities-population-designation.html

Sanfacon, K., et al (2020). Cross-sectional analysis of medical conditions in the U.S. deaf transgender community. *Transgender Health*. https://10.1089/trgh.2020.0028

Singleton, J. L., Jones, G., & Hanumantha, S. (2014). Toward ethical research practice with deaf participants. *Journal of Empirical Research on Human Research Ethics: JERHRE, 9(3)*, 59–66.

Willis, Athena, et al. Multimodal visual languages user interface, M3UI. *The Journal on Technology and Persons with Disabilities* (2019): 172.

THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

Accessible Terminal Application for Visually Impaired Users Utilizing Screen Readers

Takahiro Miura[1], Hiroki Fujii[2], Ryuki Yamazaki[2], Delgerbayar Erdenesambuu[2], Masaki Matsuo[2], Masatsugu Sakajiri[2], Junji Onishi[2]
National Institute of Advanced Industrial Science and Technology (AIST)[1], Tsukuba University of Technology[2]
miura-t@aist.go.jp, k213202@cc.k.tsukuba-tech.ac.jp, ryuki@solab.work, deegii@solab.work, matsuo@cs.k.tsukuba-tech.ac.jp, sakajiri@cs.k.tsukuba-tech.ac.jp, ohnishi@cs.k.tsukuba-tech.ac.jp

Abstract

Visually impaired computer users employ screen readers to navigate their machines, receiving auditory output for visualized content. When utilizing a terminal application for program execution, they encounter challenges in parsing large volumes of character strings sequentially, impeding efficient information retrieval. This study aims to address this issue by introducing an accessible terminal application tailored for the visually impaired, termed *WEBBY Term*. Distinguished by its capability to output the preceding operation's results to a supplementary window, this feature facilitates the streamlined retrieval and examination of essential information. To gauge usability, we conducted a comparative analysis between WEBBY Term and PowerShell, the conventional terminal on Windows, among visually impaired individuals with programming experience. Our findings indicate that the developed terminal surpasses PowerShell in terms of usability for this user demographic.

Keyword

Visually impaired, terminal application, WEBBY Term, usability, screen readers

Introduction

Visually impaired individuals facing challenges in visually checking screens rely on screen readers to operate personal computers. These tools provide synthetic voice output of textual information displayed on the computer screen, enabling users to navigate and operate the system through auditory cues. In the context of programming training courses, visually impaired students using screen readers encounter the additional challenge of learning unique operations within a programming environment tailored for screen reader conditions.

The increasing complexity of integrated development environments (IDEs) poses accessibility challenges in web development. Kearney-Volpe et al. identified five issue categories related to accessibility in web development, encompassing visual information without accessible equivalents, orientation, navigation, lack of support, and knowledge and use of supportive technologies (Kearney-Volpe 8:32). Mountapmbeme further highlighted challenges in code navigation, comprehension, editing, and other accessibility issues in user interfaces (Mountapmbeme 7:26). Visually impaired programmers encounter difficulties in debugging, locating specific information, utilizing indentation-based language, understanding scope levels, accessing autocomplete features, and discerning relationships between code entities (Albusays 100).

Researchers have sought solutions by developing plugins for coding helpers in various IDEs such as Eclipse (Baker 3052) and Visual Studio (Potluri 681:11). Ehtesham-Ul-Haque et al. proposed a *grid-coding* style for web-based editors, implementing a special editor that expresses codes in a spreadsheet-like view (Ehtesham-Ul-Haque 44:21). To create visual media for inclusive and audio games, Matsuo et al. introduced *AudibleMapper*, facilitating the development of 2D maps without visual information (Matsuo 544, Matsuo 191, Matsuo 70).

Efforts have also been made to facilitate training of visually impaired students in programming languages. Kane et al. conducted a programming workshop for blind students, involving the creation of Ruby programs for data analysis and accessible tactile visualizations using 3D printing (Kane 252). Onishi et al. developed and evaluated a remote communication support system for visually and hearing–visually impaired students during programming lectures and exercises (Onishi 176).

Despite these advancements, visually impaired programmers face challenges when developing web applications that require debugging on a web server. This necessitates terminal operations using network protocols such as SSH (secure shell protocol). However, existing terminal manipulations are not optimized for screen readers, often containing superfluous information. Visually impaired users must learn various additional commands, creating a higher usage barrier compared to that for sighted users.

The objective of this study is to develop an accessible terminal application tailored for visually impaired computer users. Usability evaluations of the developed terminal and conventional counterparts were conducted among visually impaired individuals with programming experience.

Overview of the System

Development Environment

The accessible terminal application was developed using the following technologies:

1. Windows 11 Pro for 64-bit CPUs
2. Node.js 16.13.1
3. Electron 13.1.1
4. xterm 4.19
5. Compatible screen readers

- PC-Talker (PC-Talker)
- NVDA (NVDA)

The terminal application introduced in this study is designated as *WEBBY Term*, an acronym for *Web-based Educational-support for Blind persons orchestrated by Bayer and Yamazaki*, who are both the authors with total blindness and spearheaded the creation of this terminal. The series of developments and preliminary evaluations leading up to the assessment presented in this paper adhered to the inclusive research scheme outlined by Nind (Nind 2014) to reflect the opinions of visually impaired developers in a sequential manner.

Implemented Functions

Our system, developed using web technology, offers users an interface familiar to standard web browsing conventions. To enhance usability and efficiency, the following functions have been integrated:

- Temporary Storage of Last Standard Output: The last standard output is temporarily preserved in the complementary window.
- Notification with Sound Effect: Users receive a notification accompanied by a sound effect upon completion of the preceding command execution, signaling the transition to the next command state.
- Screen Scrolling Functionality: The system enables users to scroll through the screen, reading the previous output line by line, facilitated by the up/down directional keys.

Figure 1 illustrates the scenario in which PowerShell is invoked using WEBBY Term. In Figure 2, the last standard output is presented in the complementary window. Specifically, the output of the *ls* command is presented, and the name strings of the “Music” directory are selected using keyboard commands exclusively.

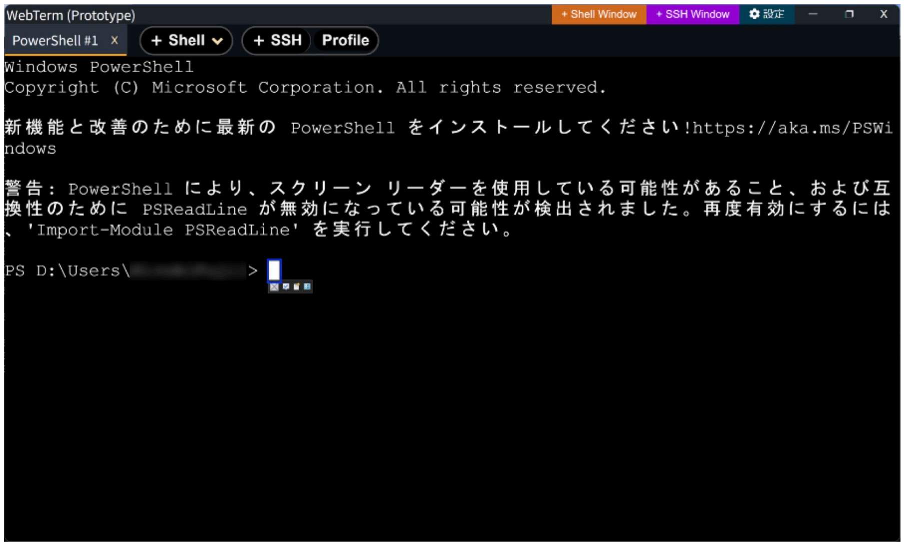


Fig. 1. Screenshot of the window with the WEBBY Term.

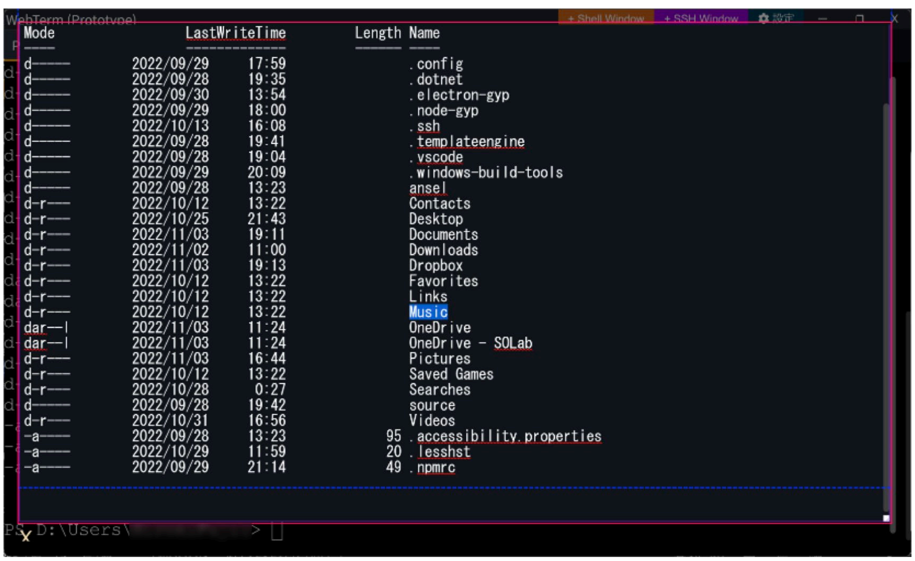


Fig. 2. Complementary window to temporarily store and display the output of WEBBY Term.

Evaluation

Participants

The evaluation of the system was performed by five visually impaired students, comprising three men and two women, in their teens to twenties, who were pursuing studies in computer science. Four participants were congenitally visually impaired, with two being totally blind and three having low vision. All participants reported daily use of personal computers and smartphones equipped with screen readers, and possessed programming experience. Over the last two weeks, they dedicated a minimum of 5 hours per day to computer-related activities. All participants were regular users of screen readers, with four employing NVDA, two using PC-talker, and one relying on JAWS for their computing needs.

Procedure

To assess the efficacy of our developed system, we conducted a comparative experiment pitting our terminal application against the PowerShell system, one of the most widely used multifunctional terminals in the Windows environment. Participants were instructed to navigate a terminal and leverage its standard output for tasks while concurrently engaging in surveys on a website. The participants seamlessly switched between the terminal, web browser, and text editor windows as required. Task procedures were executed using both our terminal system and PowerShell. Subsequently, participants responded to a questionnaire that included usability questions measured through UMUX-Lite (Lewis, 2102). Additionally, participants provided free-form feedback detailing their experiences with the two terminal applications.

Specific Task Procedures

The participants performed the following steps on both WEBBY Term and PowerShell.

1. Clone a specified Git repository: Participants were informed in advance about the time required for repository cloning.
2. Simultaneously perform a specified web search operation while cloning: If cloning is completed, proceed to steps 3 and 4.
3. Navigate to the repository directory created in step 1.
4. Execute the specified web search operation
5. Execute the *ls* command: Hereafter, use the standard output of the *ls* command to inspect file information.

6. Confirm the existence of the specified file.

7. Conduct the specified web search.

8. Document the file information from step 6.

9. Reconfirm the existence of the specified file.

10. Document the file information from step 9.

Results and Discussion

Usability Score

Figure 3 illustrates the SUS scores estimated based on UMUX-Lite responses. Despite the small sample size, our terminal application, WEBBY Term, outperformed PowerShell significantly ($t(4) = -3.57$, $p = 0.02 < 0.05$, $d = 1.26$ (Large)). The medians of estimated SUS scores were 50.0 for PowerShell and 77.0 for WEBBY Term. Applying Bangor's criteria (Bangor 594), the median score for PowerShell fell marginally low in the acceptability range and almost OK in the adjective ratings. In contrast, the WEBBY Term scored acceptable in the acceptability range and received good in the adjective ratings.

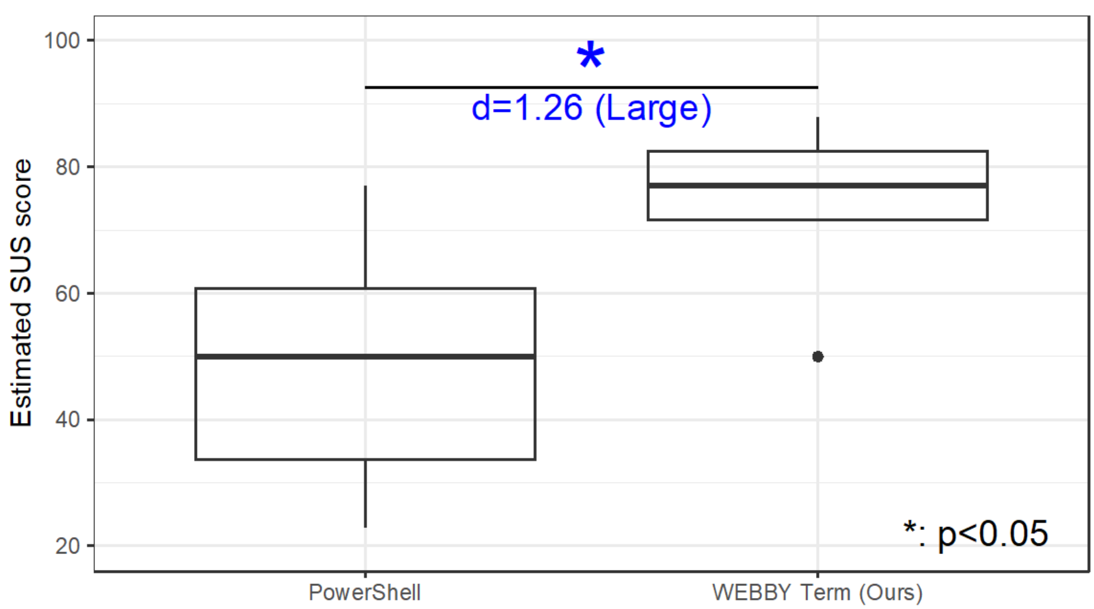


Fig. 3. Evaluated usability scores of the forms using UMUX-Lite.

Preference on Visual Condition

Despite the limited participant size, we conducted an analysis of variance (ANOVA) to explore the impact of two factors—visual impairment status and terminal type—on the estimated SUS scores. Given the challenge of ensuring normality in SUS scores, we applied an aligned rank transform (ART) (Wobbrock 146). The results revealed significant effects for both visual impairment status ($F(1,1) = 6.52$, $p = 0.043 < 0.05$, $\eta^2 = 0.26$ (Large)) and terminal type ($F(1,1) = 6.87$, $p = 0.039 < 0.05$, $\eta^2 = 0.36$ (Large)). However, no significant interactions were observed between these two factors ($F(1,1) = 1.80$, $p = 0.22 > 0.10$, $\eta^2 = 0.08$ (Middle)).

The mean SUS scores for blind and low vision participants were 33.7 pts and 68.9 pts, respectively, when using PowerShell, whereas for WEBBY Term, these scores were 77.1 pts and 79.8 pts, respectively. Generally, WEBBY Term demonstrated higher usability scores, with visually impaired individuals tending to rate it particularly favorably. Even among low vision users, the evaluation of WEBBY Term surpassed that of PowerShell. This outcome may be attributed to the fact that blind programmers specifically designed and developed WEBBY Term for screen reader conditions. Consequently, blind participants, adept with screen readers, might have appraised WEBBY Term more positively. However, these results are constrained by the small participant pool. A more extensive evaluation involving a larger and diverse group of visually impaired programmers across various skill levels is imperative for a comprehensive understanding of the strengths and weaknesses of the proposed terminal.

Feedback Comments

Feedback comments for PowerShell aligned with issues commonly reported in conventional terminal applications. Participants expressed challenges in navigating back from another window to the terminal and confirming task completion. Difficulties in checking details without output redirection for long standard outputs were also noted, consistent with findings from prior studies (Albusays 100, Kearney-Volpe 8:32).

In contrast, feedback for WEBBY Term highlighted distinctions in basic operations and acknowledgment of useful features present in other terminals. Participants appreciated the ease of cursor movement, the ability to check standard output contents using standard keystrokes, and the function of temporarily saving and displaying in the complementary window. However, one user suggested the need for a feature to enable users to acclimate to switching between focus and browse modes when using NVDA as a screen reader, given the web-like operations employed by WEBBY Term. Therefore, enhancing user familiarity with basic operations is identified as an area for improvement in the WEBBY Term.

Conclusion

In addressing programming barriers faced by visually impaired programmers, we conceived and assessed WEBBY Term—an accessible terminal application tailored for the visually impaired. The key achievements of this study are as follows:

- Efficient Information Retrieval: The terminal's distinctive feature involves outputting the result of the preceding operation to the complementary window, enhancing the efficiency of information retrieval and browsing.
- Usability Comparison with a Conventional Terminal: The usability of WEBBY Term was systematically compared with that of PowerShell among visually impaired individuals with programming experience. Results indicated the superior usability for WEBBY Term over PowerShell, evident in both programmers with total blindness and low vision.

Our future endeavors involve enhancing the stability of WEBBY Term for broader use and facilitating its availability to the public. Additionally, we aim to implement this accessible terminal application in the realm of program education, specifically catering to students with visual impairments.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers JP21H00885, JP21K18483, JP21K18484, JP23H03903, and JP23K17582. We would also like to thank all those who cooperated in the experiment.

Works Cited

Aboubakar Mountapmbeme, Obianuju Okafor, Stephanie Ludi. 2022. Addressing accessibility barriers in programming for people with visual impairments: A literature review. *ACM Trans. Access.* Comput. 15, 1, Article 7:1-26. https://doi.org/10.1145/3507469

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction, 24(6)*, 574-594.

Catherine M. Baker, Lauren R. Milne, Richard E. Ladner. 2015. StructJumper: A tool to help blind programmers navigate and understand the structure of code. *In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3043–3052. https://doi.org/10.1145/2702123.2702589

Claire Kearney-Volpe, Amy Hurst. 2021. Accessible web development: Opportunities to improve the education and practice of web development with a screen reader. *ACM Trans. Access. Comput. 14(2)*, 8:1-32. https://doi.org/10.1145/3458024

Junji Onishi, Takahiro Miura, Takeshi Okamoto, Masaki Matsuo, Masatsugu Sakajiri. Online communication assistant system for Deafblind person, Future perspectives of AT. *eAccessibility and eInclusion*, pp:171-176, 2020.

Khaled Albusays, Stephanie Ludi, and Matt Huenerfauth. 2017. Interviews and observation of blind software developers at work to understand code navigation challenges. *In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17)*. Association for Computing Machinery, New York, NY, USA, 91–100. https://doi.org/10.1145/3132525.3132550

Masaki Matsuo, Takahiro Miura, Masatsugu Sakajiri, Junji Onishi, Tsukasa Ono. Audible mapper & ShadowRine: Development of map editor using only sound in accessible game for blind users, and accessible action RPG for visually impaired gamers. *Lecture Notes in Computer Science*, 9758:537-544, 2016.

Masaki Matsuo, Takahiro Miura, Junji Onishi, Masatsugu Sakajiri, Tsukasa Ono. Experience report of blind gamer to develop the accessible action RPG. *Journal on Technology and Persons with Disabilities*, 5:172-191, 2017.

Masaki Matsuo, Takahiro Miura, Ken-ichiro Yabu, Atsushi Katagiri, Masatsugu Sakajiri, Junji Onishi, Takeshi Kurata, and Tohru Ifukube. 2021. Inclusive action game presenting real-time multimodal presentations for sighted and blind persons. *In Proceedings of the 2021 International Conference on Multimodal Interaction* (ICMI '21). Association for Computing Machinery, New York, NY, USA, 62–70. https://doi.org/10.1145/3462244.3479912

Md Ehtesham-Ul-Haque, Syed Mostofa Monsur, and Syed Masum Billah. 2022. Grid-coding: An accessible, efficient, and structured coding paradigm for blind and low-vision programmers. *In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (UIST '22). Association for Computing Machinery, New York, NY, USA, Article 44, 1–21. https://doi.org/10.1145/3526113.3545620

Lewis, James R., et al. UMUX-LITE: When there's no time for the SUS. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 2099-2102, https://doi.org/10.1145/2470654.2481287.

Nind, M., What is inclusive research? *Bloomsbury Publishing*, 2014. https://doi.org/10.5040/9781849668149

NVDA, https://www.nvaccess.org/about-nvda/ (Accessed: 2023.09.05).

PC-Talker, https://www.aok-net.com/screenreader/ (Accessed: 2023.09.05) (in Japanese).

Shaun K. Kane and Jeffrey P. Bigham. 2014. Tracking @stemxcomet: teaching programming to blind students via 3D printing, crisis management, and twitter. *In Proceedings of the 45th ACM technical symposium on Computer science education* (SIGCSE '14). Association for Computing Machinery, New York, NY, USA, 247–252. https://doi.org/10.1145/2538862.2538975

Wobbrock, J.O.; Findlater, L.; Gergle, D.; Higgins, J.J. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. *In Proceedings of the SIGCHI conference on human factors in computing systems*, Vancouver, BC, Canada, 7–12 May 2011; pp. 143–146.

Venkatesh Potluri, Priyan Vaithilingam, Suresh Iyengar, Y. Vidya, Manohar Swaminathan, and Gopal Srinivasa. 2018. CodeTalk: Improving programming environment accessibility for visually impaired developers. *In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18). Association for Computing Machinery, New York, NY, USA, Paper 618, 1–11. https://doi.org/10.1145/3173574.3174192

iSinkwe: An Application that Synchronises Text and Audio for Enhanced Reading

The Journal on Technology and Persons with Disabilities

Rynhardt Kruger, Avashna Govender, Willem van der Walt, Ilana Wilken
Council for Scientific and Industrial Research
rkruger@csir.co.za, agovender1@csir.co.za, wvdwalt@csir.co.za, iwilken@csir.co.za

Abstract

We present iSinkwe, a system to produce synchronised accessible EPUB3 books of text and audio. With iSinkwe, users are able to synchronise EPUB3 publications with human-narrated or computer-generated speech, via an accessible web interface. Documents in other formats can also be converted to EPUB3. Developed specifically to address reading barriers experienced by users with print disabilities such as dyslexia and visual impairment. iSinkwe is also of particular importance for regions with low literacy such as South Africa. This paper describes the motivation and context for its creation, the components that make up iSinkwe, a discussion on the relevance it has for the accessibility community, and how users can interact with the system. A usability study was performed on a previous iteration of iSinkwe, with mixed results. We report on the lessons we learned, and subsequent improvements to the system. Finally, we describe future work planned to extend its functionality.

Keywords

Accessible books, literacy, reading, synchronised highlighting

Introduction

The development of accessible book formats like DAISY and EPUB 3 has revolutionised the way in which print-disabled readers access information, by extending the traditional audio book with sophisticated navigational capabilities (Engelen; Garrish). When text is included in these books, the reader also retains access to the symbolic form of the content, which allows interrogation with assistive technologies that output to text-to-speech (TTS) or a Braille display. However, combined text and audio EPUB 3 books are difficult to produce, requiring careful synchronisation of the source text and audio. Although TTS can be used to generate audio from text, it is inadequate when the text contains complex dialogue or words with uncommon pronunciations (Kuligowska et al. 234). In a multilingual context such as South Africa, TTS also struggles to accurately reproduce the code switching that occurs in local documents, that is, alternating between different languages in the same sentence (Setati et al. 128).

Discussion

Motivation and Context

With the advances in voice computing technologies, it is now possible to develop digital books augmented with speech. However, most of these developments have focused primarily on widely used languages such as English, German, and French. The reality is that little or no development of language technologies in most minority and under-resourced languages of the world has been done, especially those spoken in Sub-Saharan Africa. Yet these languages serve an equal purpose in the socio-economic development of communities where they are spoken. In such communities, literacy in English is typically lacking, and therefore human language technology solutions that only cater for languages such as English are not helpful. These language barriers are further aggravated by a digital/connectivity divide resulting in communities being denied access to information in their home languages. The language and digital divides further exacerbate low literacy levels by isolating communities.

To achieve an equitable society, i.e., one where information exchange benefits all members of society, it is crucial that we develop solutions that tear down the language barriers that exist in such communities. By offering human language technologies in local languages, we can ensure that people who speak the minority languages can also enjoy literature, educational content, and other forms of information in an accessible format. South Africa, in particular, has a diverse linguistic landscape and ten of the twelve official languages can be classified as under-resourced. Therefore, developing solutions that support these languages is crucial for alleviating language barriers that exist in local communities in this country.

From an educational standpoint South Africa has many regions that struggle with literacy challenges. One contributing factor is learners having limited access to educational content in their home languages. Providing solutions that provide access to educational content in their home languages is a valuable tool that promotes literacy. In addition to language barriers, many learners with print disabilities such as dyslexia and visual impairment, as well as learners with low literacy also struggle with reading barriers that most human language technology solutions don't cater for.

Proposed Solution

As a solution to the above-mentioned challenges, we present iSinkwe, a system that augments EPUB 3 publications by automated synchronisation of text with audio. By using iSinkwe, existing human-recorded audio can be synchronised with the text, to word, sentence and paragraph level. This facilitates the creation of multimodal accessible books which provides the benefits of symbolic text combined with the rich expressiveness of audiobooks produced by human voice artists (Knox 127). iSinkwe can also utilise TTS to synthesise audio, for parts of the book for which no pre-recorded audio is available. The TTS functionality is capable of auto-language switching, based on the language tags in the HTML.

The audio from either natural sources or TTS is embedded inside the publication using EPUB 3 media overlays. This embedding is particularly suited for the South African context, where data connectivity is sporadic, and users may not have access to high quality voices provided as cloud services. However, since both the text and audio are contained within the EPUB 3, users retain the capability to explore the text of the book with their chosen screen reader (King 265). iSinkwe is also capable of generating older DAISY 2.02 publications using the DAISY Pipeline2, to be played on older devices that users may still retain.

Interaction

iSinkwe consists of three components. iSinkwe Convert is a converter, to allow conversion of alternative document formats into EPUB 3. iSinkwe synchronize synchronizes the text of EPUB 3 documents with audio. Finally, iSinkwe Read is an EPUB 3 reader which supports media overlays.

The web user interfaces for iSinkwe Convert and iSinkwe Synchronise is presented in Figure 1 and Figure 2. In the iSinkwe Convert web user interface, the user needs to upload their document, select an EPUB 3 layout and enter the title and author of their desired EPUB 3 document. In the iSinkwe Synchronise web user interface, the user is expected to upload their EPUB 3 document, select the language of the EPUB 3 book, select the type of audio they desire (human-narrated or computer-generated) and select the voice characteristics which includes the gender and specific speaker of the voice. In the case where users selected human-narrated audio as the audio type, they are also expected to upload a corresponding audio file for each section of the book.

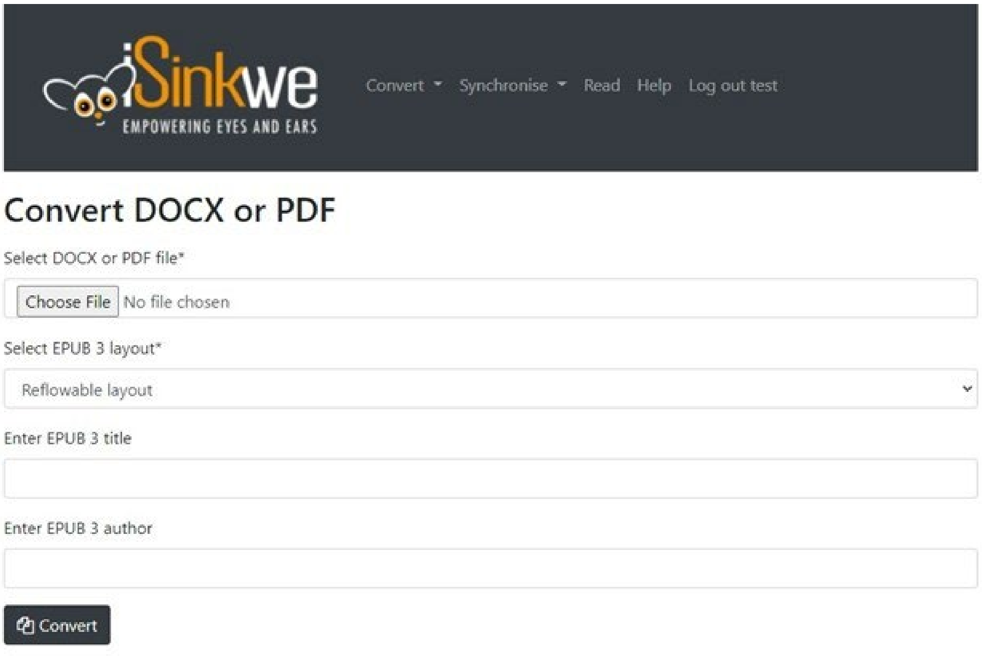


Fig. 1. Screenshot of iSinkwe Convert.

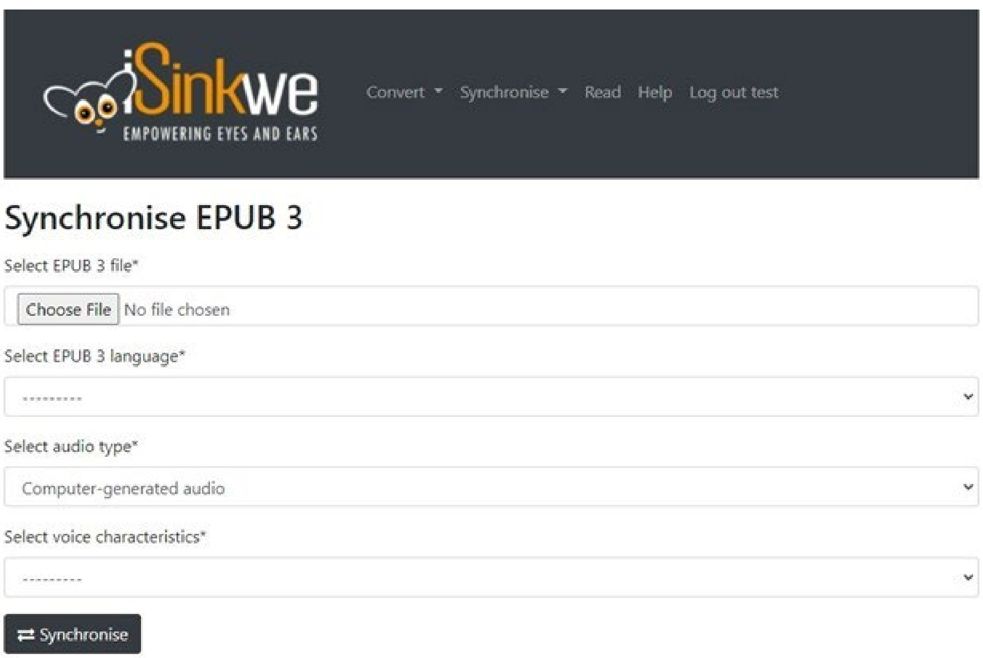


Fig. 2. Screenshot of iSinkwe Synchronise.

The mobile application user interface for iSinkwe Read is illustrated in Figure 3. In this function, the synchronised EPUB 3 created in iSinkwe Synchronise is given as input to iSinkwe Read, which can be downloaded by the user to a mobile device. In the figure, a screenshot of iSinkwe Read is presented which demonstrates the highlighted sentence feature. When a sentence is highlighted, the sentence is simultaneously read out loud using either human-narrated speech or computer-generated speech. The settings allow the user to choose whether they prefer highlighting on word, sentence or paragraph level (highlighted in black, grey and green in Figure 3 respectively).

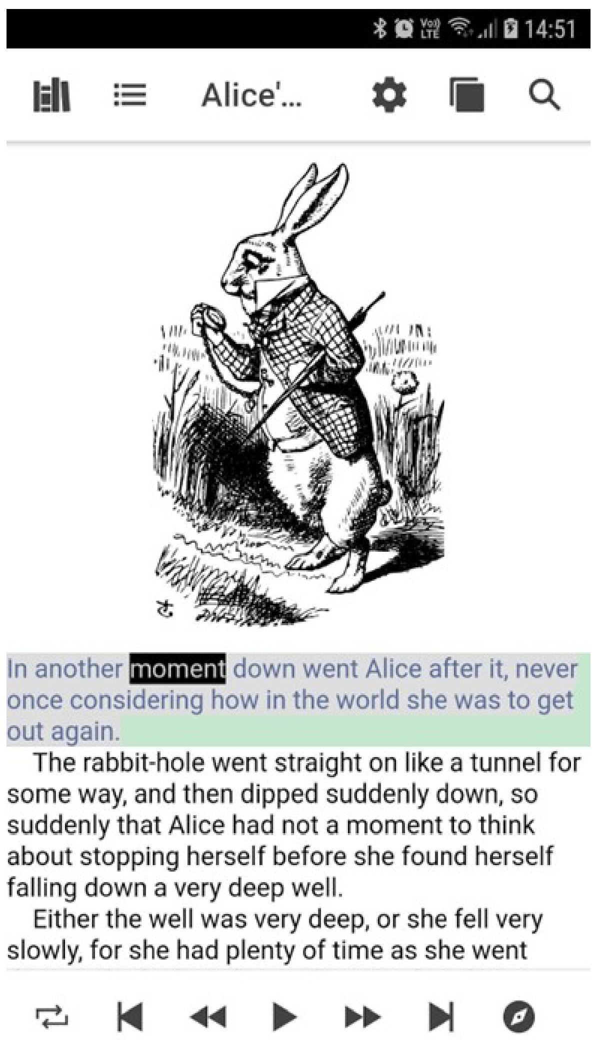


Fig. 3. Screenshot of iSinkwe Read.

Technical Description

iSinkwe does the automatic alignment by first extracting the XHTML files from the input EPUB file, then adding three levels of span tags with ID attributes to each file (paragraph, sentence and word level). If a pre-existing audio file exists for a given XHTML file or sub-section of such a file, alignment is done immediately. If no pre-existing audio is available, the text is first synthesized using text to speech to create an audio file. Using a dynamic time warping aligner (Müller 69), audio offsets for the text are calculated. The output of the aligner is a SMIL file which is then added to the EPUB document according to the media overlay specification in the EPUB 3 standard.

Dynamic time warping requires us to synthesize the text to an audio file and compare that with the audio file provided as input. Put another way, audio is synthesized even if pre-existing human audio is available, however, the user never hears this audio as it is only used internally to do the alignment. Dynamic time warping (DTW) is language dependant. A TTS engine that supports the language or languages in the book is therefore required. For South African languages, iSinkwe utilises a TTS engine also developed by our research group (Louw et al.), commercialised under the name Qfrency (CSIR). However, alignment in any other language can be supported for which a suitable TTS engine is available.

Evaluation

A previous version of iSinkwe was evaluated by a small group of diverse users. In total, ten users participated in the evaluation, comprising five blind or visually impaired users, three sighted educators for disabled learners, and two sighted members from the publishing industry. Because Covid-19 was still a realistic threat in South Africa at the time this work was performed, we decided to conduct the evaluation virtually. However, users were also offered the option of an in-person evaluation by a team member.

To evaluate iSinkwe, users were asked to perform a number of tasks with pre-selected documents as well as with their own documents. A questionnaire with qualitative and quantitative questions was administered after the trials to gather feedback on the participants' experience with each component of the system (iSinkwe Convert, iSinkwe Synchronise, and iSinkwe Read). Questions included what users liked the most about the component, what users liked the least, whether they experienced any issues, and what they would change. Users were also asked to rate how likely they were to recommend the component on a scale from one to ten, where ten is the most likely.

Figure four depicts the possibility that users would recommend iSinkwe Convert to someone they know. Some of the reasons that users gave for recommending iSinkwe Convert include "it's a very functional product that can save blind people a lot of trouble once they have a text document they want to read", and "it is so user-friendly and accessible, even people with limited computer literacy would be able to use it". One user who gave a score of one stated that they do not currently see a need for EPUB3 documents. Another stated that they also require the conversion of charts and other graphical content. Our current improvements to the system focus on the conversion of graphical mathematical content, and we hope to extend this functionality to charts and diagrams in the future.

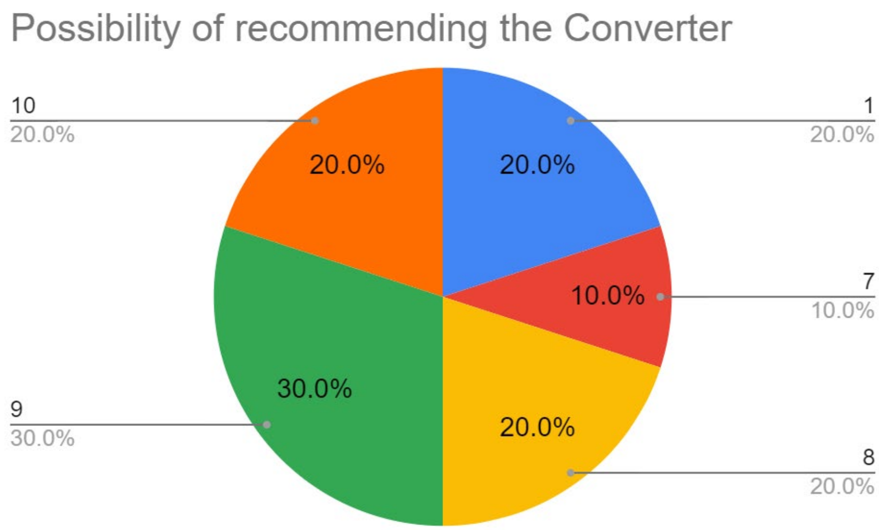


Fig. 4. The possibility of users recommending iSinkwe Convert to their friends, colleagues or family members.

Figure five depicts the possibility that users would recommend iSinkwe Synchronise to someone they know. Some of the aspects that users appreciated about iSinkwe Synchronise include “The clearly marked buttons and combo boxes.”, and “Having the freedom to choose the voices and type of augmentation.”. Some of the reasons why users scored iSinkwe Synchronise highly include “Although there are some teething troubles, it is very accessible and a brilliant tool for text-to-speech on documents which saves time.”, and “It’s a really functional product.”. The users who gave a score of one gave as reasons that the synchronisation takes too long, and that they found the interface rather complex with many steps to follow. Unfortunately, the speed of synchronisation is dependent on the dynamic time warping algorithm, although users are able to start the process and log off from the interface while the synchronisation takes place. Users are notified via email when the synchronisation is complete.

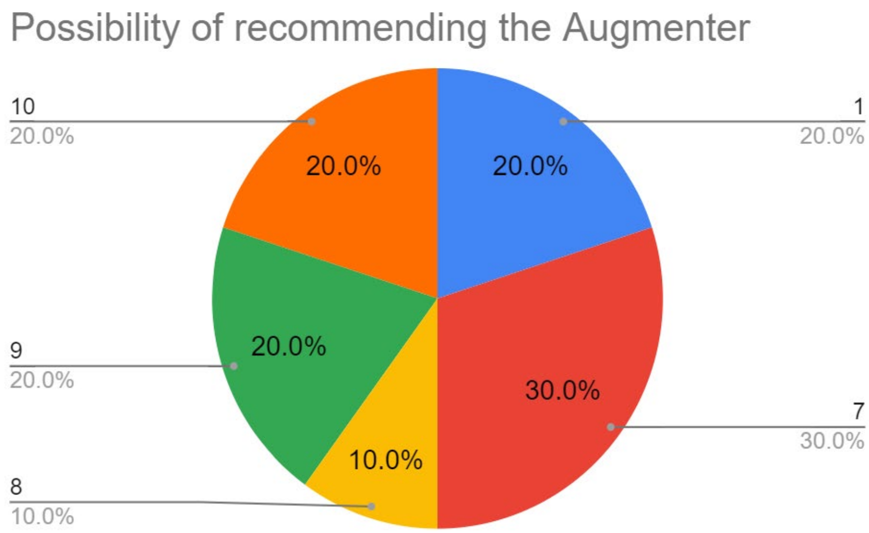


Fig. 5. The possibility of users recommending iSinkwe Synchronise to their friends, colleagues or family members.

Figure six depicts the possibility of users recommending iSinkwe Read to someone they know. Two users indicated that they will not likely recommend the Reader to other users (they gave a score of 1), since it “does not have any of the capabilities of the screen reader that we are already using at our school”. One user gave a rating of 3, stating the Reader still needs work. The users who gave a score of 9 said the speech is clear and the controls are labelled clearly, but the navigation is confusing to new users, and also that the Reader is “very user friendly and efficient (once you find and get your book imported)”.

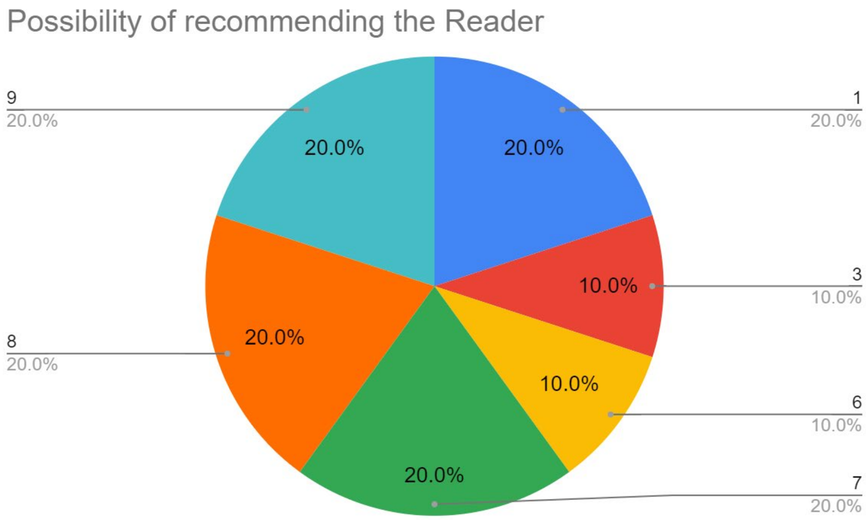


Fig. 6. The possibility of users recommending iSinkwe Read to their friends, colleagues or family members.

Overall, the blind and visually impaired users had a mostly positive experience with iSinkwe Convert and iSinkwe Synchronise, although they suggested some improvements from the web content accessibility guidelines (WCAG), which were subsequently integrated into the system. However, their experience with iSinkwe Read was mixed, with many stating that they already have ebook readers that they prefer. We therefore continue to focus on standard conforming EPUB 3 books that can be read with any reader that supports media overlays.

The members of the publishing industry were specifically interested in iSinkwe Convert and iSinkwe Synchronise, and therefore mainly used iSinkwe Read to verify the resulting documents. They had a mostly positive experience with iSinkwe and commented that iSinkwe has the potential for opening a world of reading to those who cannot or who struggle to read. In contrast, the educators had the least positive experience of the three user groups. At the time, iSinkwe did not fit their use case, since they required software that allows learners to read and write on the same document. Their experience can also be attributed to the high workload of educators in South Africa, especially educators for disabled learners, resulting in their limited time to explore new technologies. We have subsequently created an instruction video, to assist users in familiarizing themselves with the system.

The participants also had to complete a Systems Usability Scale questionnaire which assisted with determining the usability of iSinkwe (Brooke 189). iSinkwe obtained a SUS score of 62, which measures as a D-grade on the scale. This means iSinkwe was rated as between OK/Fair and Good. A score of 68 is deemed average. However, significant development has since taken place and a new user evaluation is needed to calculate an updated score for iSinkwe.

Future Work

Current development focuses on an iSinkwe component for automatically recognising inaccessible mathematical expressions from bitmap images using optical character recognition capable of interpreting mathematical information. By utilizing the MathJax rendering library (Cervone et al.), these expressions are augmented with MathML which the user can explore, as well as a textual description to be synchronised using the existing audio synchronisation functionality. In future work, we aim to extend this image recognition functionality to also recognise other technical information types like diagrams that occur in textbooks (Emerson and Anderson 20) and develop a method by which users may explore these objects.

Conclusion

This paper described iSinkwe, a solution for augmenting books and other electronic documents with human and/or computer-generated speech. iSinkwe accomplishes this by utilising the EPUB 3 document format with media overlays.

iSinkwe was evaluated by users, teachers and publishers, and the feedback received during the evaluation was incorporated into subsequent versions of iSinkwe as far as it was possible and feasible. Improvements are also made when necessary to ensure a good user experience for those who use iSinkwe.

iSinkwe's core functionality is to create multimodal digital books and therefore plays a key role in accessibility to individuals with visual impairments and other print disabilities. It provides an encouraging solution for those who have reading barriers and in addition supports users who prefer to read in their home languages that are resource scarce. By synchronising text with audio, it adds an additional mode of reading to existing documents, thereby offering a multi-modal solution that allows users to read in a more interactive and engaging way which has many advantages as discussed in (Rumsey 191). Therefore, especially in the low-literacy context, iSinkwe contributes towards turning non-readers into readers, makes content available to a wider range of users and promotes a more interactive and engaging approach to literacy development.

Works Cited

Brooke, John. "Sus: A "quick and dirty' usability." *Usability evaluation in industry* 189.3 (1996): 189-194.

Cervone, Davide, Peter Krautzberger, and Volker Sorge. "Towards universal rendering in MathJax." *Proceedings of the 13th International Web for All Conference*. 2016.

CSIR. "Qfrency". *Council for Scientific and Industrial Research*. www.qfrency.com. Accessed 4 Oct. 2023.

Emerson, Robert Wall, and Anderson, Dawn. "What mathematical images are in a typical mathematics textbook? Implications for students with visual impairments." *Journal of Visual Impairment & Blindness* 112.1 (2018): 20-32.

Engelen, Jan. "E-books and audiobooks: what about their accessibility?" *International Conference on Computers for Handicapped Persons*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

Garrish, Matt. "What is EPUB 3?" *O'Reilly Media, Inc*., 2011.

King, Alasdair. "Screenreaders, magnifiers, and other ways of using computers." *Assistive Technology for Blindness and Low Vision* (2018): 265-288.

Knox, Sara. "Hearing hardy, talking Tolstoy: The audiobook narrator's voice and reader experience." *Audiobooks, literature, and sound studies*. Routledge, 2011. 127-142.

Kuligowska, Karolina, Paweł Kisielewicz, and Aleksandra Włodarz. "Speech synthesis systems: Disadvantages and limitations." *International Journal of Engineering & Technology* (UAE) 7 (2018): 234-239.

Louw, Johannes A., et al. "The speech text-to-speech entry for the Blizzard Challenge 2016." *Blizzard Challenge 2016*. Cupertino, United States of America. 16 September 2016.

Müller, Meinard. "Dynamic time warping." *Information retrieval for music and motion* (2007): 69-84.

Rumsey, Francis. "Audio in multimodal applications." *Journal of the Audio Engineering Society* 58.3 (2010): 191-195.

Setati, Mamokgethi, et al. "Incomplete journeys: Code-switching and other language practices in mathematics, science and English language classrooms in South Africa." *Language and education* 16.2 (2002): l 128-149.

"Web content accessibility guidelines (WCAG) 2.1". *World Wide Web Consortium*. www.w3.org/TR/WCAG21/. Accessed 5 Dec. 2023.

The Journal on Technology and Persons with Disabilities

Accessible Web Development for Underrepresented Disabilities

Cesar Coronel, Jonna Alonso, Carsen Potter
Southern Utah University
cesarcoronel@suu.edu, jonnalynnalonso@suu.edu, carsenpotter@suu.edu

Abstract

This paper explores how to improve inclusivity for individuals with invisible underrepresented cognitive disabilities in Information and communication technology (ICT), specifically web design. To this end, we focused on the following topics: the Web Content Accessibility Guidelines (WCAG) 2.0 A and AA standards web requirements designed to accommodate those with underrepresented cognitive disabilities, the challenges of accommodating underrepresented disabilities in ICT, practical strategies for accommodating those with underrepresented cognitive disabilities in ICT, and an examination of accessibility overlays as tools to help accommodate those with underrepresented cognitive disabilities. We argue that it is imperative for all organizations to strive to improve inclusivity for those with underrepresented cognitive disabilities. By applying the WCAG A and AA standards in partnership with these practical guidelines discussed within this paper; organizations can create a more welcoming, supportive, and inclusive environment for those affected by these underrepresented cognitive disabilities as well as an overall improved experience for all users. Research on this topic was done in partnership with the Forest Inventory and Analysis (FIA) under the 19-CS-11242305-075 agreement, to better understand and help meet the needs of the disabled community.

Keywords

Cognitive and learning disabilities, design development, and web.

Introduction

Since the passing of Section 508 of the Rehabilitation Act of 1973, there has been a national push for organizations to implement accessibility standards for electronic and information technology (EIT). This also involves making ICT content accessible to individuals, employees, and members of the public alike, with disabilities that may impact their ability to use and interact with digital content in a way that mimics the experiences of individuals without disabilities. However, these standards tend to focus on physical impairments, such as visual and mobility impairments, which are well-accommodated by Section 508 guidelines. Certain underrepresented cognitive disabilities, such as attention-deficit/hyperactivity disorder (ADHD) and anxiety, are not always as well-understood or recognized, and, as such, are often underrepresented. As over 26% of the population struggle with these types of invisible cognitive disabilities ("Mental Health Disorder Statistics"), this can make it difficult for people with underrepresented disabilities to access and use digital content that is considered 'accessible' or 'compliant' under the current Section 508 guidelines.

In recent years, the scope of accessibility considerations has expanded to include not only physical disabilities but also cognitive impairments. Web Content Accessibility Guidelines (WCAG) 2.0 A and AA standards can help through the inclusion of guidelines that address some of the invisible underrepresented cognitive disabilities such as dyslexia, anxiety, and attention deficit hyperactivity disorder (ADHD). As we progress, we will be referring to WCAG 2.0 A and AA standards as simply WCAG. For the scope of this paper, we focused on the experiences of people with underrepresented disabilities in a broader sense, as many are affected by similar difficulties and experiences. Underrepresented cognitive disabilities can impact how people process information (Kc). For example, they can affect people's perception, memory, language, attention, problem solving, and comprehension (Pichiliani et al.). Underrepresented cognitive disabilities can present unique challenges for web users, affecting users' ability to understand, navigate, and interact with websites. Thus, common issues faced by individuals with underrepresented cognitive disabilities include difficulty in reading, understanding complex instructions, and processing substantial amounts of information ("Inclusive Teaching: Intellectual Disability"). With proper training, web developers can play a crucial role in ensuring that websites and web applications adhere to WCAG standards while meeting the needs of people with underrepresented cognitive disabilities. This paper delves into the critical aspects of WCAG compliance, provides in-depth insights into cognitive accessibility challenges, and offers practical solutions for web developers to create an inclusive online environment. Additionally, this paper discusses practical strategies that web developers can employ to create web content that conforms to current WCAG standards, with a specific focus on addressing underrepresented cognitive disabilities.

Discussion

Underrepresented Cognitive Disabilities and Web Accessibility for WCAG Compliance

Users who struggle with unclear directions, complicated wording, unclear symbols, and user interfaces may find themselves struggling to comprehend necessary information (Seeman-Horwitz et al.). Research shows people with disabilities attributed to mental conditions were particularly susceptible to these types of comprehension difficulties (Druss et al.).

Previous testing was done on many of the most public popular websites, such as Amazon and Wikipedia to ensure WCAG accessibility. These popular sites were tested by Trusted Tester certified research assistants using the Accessible Name & Description Inspector (ANDI) tool and were found to be on average more than 80% compliant with current WCAG guidelines (Potter).

However, none of the websites tested used any additional tools such as accessibility overlays, to make the website more accommodating to individuals with underrepresented cognitive disabilities such as dyslexia, anxiety, and/or attention deficit ADHD). This lack of web resources for those with an underrepresented cognitive disability exacerbates the difficulties faced by those who struggle with these types of issues.

Practical Strategies

Ensuring WCAG compliance with a focus on underrepresented cognitive disabilities is an essential step toward creating an inclusive and accessible online environment. By implementing the practical strategies outlined in this paper, web developers can contribute to the development of websites and web applications that cater to the needs of all users regardless of ability. Web accessibility is not merely a legal requirement for government websites but a moral obligation. It is incumbent upon organizations to embrace these strategies to make the internet a more inclusive space for everyone.

Organizations have a key role in changing the culture of accessibility. Requiring developers to become intimately familiar with WCAG and use those guidelines as a detailed framework for making web content accessible, organizations can establish a culture of accessibility and inclusivity. Organizations can also provide developers with training and resources on underrepresented disabilities, create an accessibility team or committee to oversee accessibility efforts, acquire a willingness to make reasonable accommodations, respond timely to accessibility feedback, and foster a culture of empathy, inclusion, and/or acceptance. Organizations can also require that developers practice the principles of universal design. Universal design offers numerous benefits, including increased accessibility and usability for a wider range of users. According to the Web Accessibility Initiative (WAI) of the World Wide

Web Consortium (W3C), universal design ensures that websites are usable by people with cognitive impairments, thus promoting inclusivity and equal access to information and services ("Introduction to Web Accessibility"). Furthermore, the incorporation of universal design principles can lead to enhanced user experience, resulting in higher user engagement and satisfaction (Henry, Shawn Lawton. "Just Ask: Integrating Accessibility Throughout Design"). Therefore, developers would greatly benefit from an example-rich guide on how to approach web development with underrepresented cognitive disabilities in mind. The guide can include the following practical strategies and examples: Plain Language and Clear Instructions, Simplify Navigation, Visual Design Considerations, Progressive Disclosure, Multimedia Accessibility, User-Centered Testing, and Error Handling.

Organizations can start by ensuring developers use plain and concise language to convey information. This includes providing clear and straightforward instructions for tasks and processes while avoiding jargon, acronyms, and complex sentence structures. Simplifying navigation involves implementing a clear and consistent website structure with a logical hierarchy. Some ways to achieve this include using intuitive menus and navigation labels as well as minimizing distractions and unnecessary animations that can confuse users. Similarly, considerations for visual design, such as having a clean and clutter-free design with ample white space; using legible fonts, appropriate font sizes, and high-contrast color schemes; and offering options for customizing font styles and text spacing, ensures users can easily navigate the website and access necessary information. Organizations can also capitalize on a progressive disclosure of elements and information to avoid users having to process substantial amounts of information at once. This may include presenting material in a step-by-step manner, revealing details gradually and allowing the user more control over the information they consume at any given time. Websites may also utilize accordion tabs and expandable sections to manage complex content. For multimedia, websites could provide users transcripts and control of playback speed for audio and video content and captions and subtitles for multimedia presentations. In terms of accessibility testing, having a user-center approach can prove beneficial, such as involving individuals with underrepresented cognitive disabilities to identify and address issues during usability testing and accessibility evaluations as well as gathering feedback and iterating on the web design and functionality. If users encounter errors while interacting with a website, having clear and concise errors messages, providing suggestions or assistance for correcting errors, avoiding punitive measures for errors, and allowing users to easily recover from mistakes will increase accessibility and inclusion for users with underrepresented cognitive disabilities.

Organizations can provide developers with the universal design principles to follow that would allow for better inclusivity, while helping to ensure WCAG standards are met, such as the use of Semantic HTML. Semantic HTML provides a solid foundation for accessibility. Developers can utilize HTML5 elements like `<header>`, `<nav>`, `<main>`, `<section>`, and `<footer>` to structure content. Semantic HTML helps screen readers and other assistive technologies understand the content's hierarchy and purpose. Implement keyboard navigation by ensuring all interactive elements, such as buttons, links, and forms, are fully navigable using a keyboard. Avoid relying solely on mouse-driven interactions. Use the `tabindex` attribute to control the keyboard navigation order. Aim to provide alternative text for images. Every image can include descriptive alternative text (alt text). Alt text conveys the image's content and function to users who cannot see it. Use concise, meaningful descriptions and create accessible forms. Forms are a common feature of web applications. Ensure that forms are accessible by using proper labels, grouping related form fields, and providing clear error messages. ARIA (Accessible Rich Internet Applications) attributes can enhance form accessibility. Organizations should conduct in-house testing with assistive technologies, such as regularly test websites with popular assistive technologies like screen readers (e.g., JAWS, NVDA, etc.) and keyboard-only navigation. This hands-on testing will help identify and fix accessibility issues. Provide closed captions for multimedia. If the website includes multimedia content, ensure that it includes closed captions for videos and transcripts for audio. This benefits users with hearing impairments and those who may prefer reading content. Finally, organizations can achieve greater inclusivity for those with underrepresented cognitive disabilities by ensuring web development occurs with a responsive design in mind. Responsive design is not only crucial for different screen sizes but also for various assistive technologies. Ensure that the site's layout and content adapt well to different devices and orientations.

Use of Accessibility Overlays

Accessibility Overlays are Javascript-powered addons that rest on the ‘edge’ of a website or right on top of the final layer of information being presented for viewing. Being ‘on the edge’ allows page overlays to execute a variety of functions between the website and the visitor (“How to Use Accessibility Widgets in Order to Improve Usability”). It is important to note that while accessibility overlays can offer valuable support, they cannot replace the fundamental principles of web accessibility. Developers should strive to create content with semantic HTML and adhere to the WCAG to ensure that websites are as accessible as possible from the outset.

Additionally, overlays can be used as supplementary tools to further enhance accessibility for users with underrepresented cognitive disabilities and need to be chosen and configured with care to avoid potential usability issues or conflicts with assistive technologies. When implementing these accessibility widgets or tools, it is crucial to ensure that they are designed and configured with user customization in mind. These accessibility overlays should aim to help a user's ability to operate the content by facilitating how people with underrepresented cognitive disabilities process information. The specific disabilities that are helped with overlays include issues with content perception, limiting cognitive load, attention to detail, and overall comprehension. Organizations can have developers perform usability testing with individuals who have underrepresented cognitive disabilities to gather feedback and make improvements to the tools' effectiveness in addressing accessibility issues.

Technical Examples for Accessibility Overlays

Text-to-Speech (TTS) Widget. A TTS widget provides an audio narration of the website's content, making it accessible to users with underrepresented cognitive disabilities who may have difficulty reading or comprehending text. The widget can be integrated into the website with a “Listen” button that allows users to have the content read aloud.

```html
<!-- HTML -->
<button id="tts-button">Listen</button>

<!-- JavaScript to implement TTS functionality -->
<script>
  const ttsButton = document.getElementById('tts-button');
  ttsButton.addEventListener('click', () => {
    // Implement text-to-speech here
  });
</script>
```

Fig. 1. Text-to-Speech (TTS) Widget Code Example.

Simplified Reading Mode. This widget offers a simplified reading mode that removes distractions, adjusts fonts, and provides a cleaner layout to improve readability for users with underrepresented cognitive disabilities. Users can activate the “Simplified Reading Mode” from the widget, and the page's content is presented in a simplified format.

```
<!-- HTML -->
<button id="simplify-button">Simplify</button>

<!-- JavaScript to toggle simplified mode -->
<script>
  const simplifyButton = document.getElementById('simplify-button');
  simplifyButton.addEventListener('click', () => {
    // Toggle simplified mode
  });
</script>
```

Fig. 2. Simplified Reading Mode Code Example.

Interactive Tutorial Overlay. This tool provides step-by-step interactive tutorials with visual cues and simple language to guide users through complex tasks, making it helpful for individuals with underrepresented cognitive disabilities. Users can access the overlay from a “Guided Tutorial” button, which then provides step-by-step instructions with visuals.

```
<!-- HTML -->
<button id="tutorial-button">Guided Tutorial</button>

<!-- JavaScript to activate tutorials -->
<script>
  const tutorialButton = document.getElementById('tutorial-button');
  tutorialButton.addEventListener('click', () => {
    // Start interactive tutorial
  });
</script>
```

Fig. 3. Interactive Tutorial Overlay Code Example.

Read Aloud for Form Fields: This widget reads out form field labels and instructions to assist users with underrepresented cognitive disabilities in completing online forms accurately. When a user focuses on an input field, the widget automatically reads the associated label and instructions.

```
<!-- HTML -->
<form>
  <label for="name">Name:</label>
  <input type="text" id="name" name="name" required>
</form>

<!-- JavaScript to enable automatic reading -->
<script>
  const formInputs = document.querySelectorAll('input[type="text"]');
  formInputs.forEach((input) => {
    input.addEventListener('focus', () => {
      // Implement automatic reading of labels here
    });
  });
</script>
```

Fig. 4. Read Aloud for Form Fields Code Example.

Focus Enhancement Tool. The focus enhancement tool highlights interactive elements (e.g., buttons, links) and provides audio feedback when users navigate with the keyboard, aiding users with underrepresented cognitive disabilities in understanding the page's interactive elements. When enabled, the tool highlights focused elements with a border and provides an audible cue when navigating through them.

```
<!-- HTML -->
<button id="focus-button">Enable Focus Enhancement</button>

<!-- JavaScript to implement focus enhancement -->
<script>
  const focusButton = document.getElementById('focus-button');
  focusButton.addEventListener('click', () => {
    // Implement focus enhancement
  });
</script>
```

Fig. 5. Focus Enhancement Tool Code Example.

Content Simplification Button. This button simplifies the content on a page, removing complex language and providing concise summaries to improve comprehension for users with underrepresented cognitive disabilities. Users can click the "Simplify Content" button to toggle between the original content and the simplified version.

```
<!-- HTML -->
<button id="simplify-content-button">Simplify Content</button>

<!-- JavaScript to toggle content simplification -->
<script>
  const simplifyContentButton = document.getElementById('simplify-content-
  simplifyContentButton.addEventListener('click', () => {
    // Toggle content simplification
  });
</script>
```

Fig. 6. Content Simplification Button Code Example.

Customizable Font and Color Palette. This tool allows users to customize font styles, sizes, and color schemes according to their preferences, making content more accessible for those with underrepresented cognitive disabilities, such as individuals with visual processing disorders. Users can access a settings menu to adjust font styles, sizes, and colors.

```
<!-- HTML -->
<button id="customize-button">Customize</button>

<!-- JavaScript to open customization options -->
<script>
  const customizeButton = document.getElementById('customize-button');
  customizeButton.addEventListener('click', () => {
    // Open customization menu
  });
</script>
```

Fig. 7. Customizable Font and Color Palette Code Example.

Progressive Disclosure Panels. These panels allow users to reveal additional content incrementally, reducing cognitive load and improving focus on the main content. Users can click on expandable sections or tabs to reveal more detailed information.

```
<!-- HTML -->
<button class="expand-button">Expand Section</button>
<div class="expanded-content" style="display: none;">
  <!-- Content to be revealed -->
</div>

<!-- JavaScript to toggle expansion -->
<script>
  const expandButton = document.querySelector('.expand-button');
  const expandedContent = document.querySelector('.expanded-content');
  expandButton.addEventListener('click', () => {
    expandedContent.style.display = 'block';
  });
</script>
```

Fig. 8. Progressive Disclosure Panels Code Example.

Content Highlighter and Notetaking Tool. This tool enables users to highlight and take notes on web content, supporting users with underrepresented cognitive disabilities in organizing and retaining information. Users can select text to highlight and add notes, which are saved for reference.

```
<!-- HTML -->
<button id="highlight-button">Highlight</button>
<button id="note-button">Take Note</button>

<!-- JavaScript to handle highlighting and notetaking -->
<script>
  const highlightButton = document.getElementById('highlight-button');
  const noteButton = document.getElementById('note-button');

  highlightButton.addEventListener('click', () => {
    // Implement text highlighting
  });

  noteButton.addEventListener('click', () => {
    // Implement notetaking
  });
</script>
```

Fig. 9. Content Highlighter and Notetaking Tool Code Example.

Text Magnifier and Simplifier. This tool provides a magnifier for text, allowing users to enlarge and simplify text on-demand for easier reading. Users can hover over text to magnify and simplify it using a tooltip.

```
<!-- HTML -->
<span class="magnify-simplify" title="Click to magnify/simplify">This is s

<!-- JavaScript to handle tooltip-based magnification and simplification -
<script>
  const magnifySimplifyElements = document.querySelectorAll('.magnify-simpl
  magnifySimplifyElements.forEach((element) => {
    element.addEventListener('click', () => {
      // Implement magnification and simplification
    });
  });
</script>
```

Fig. 10. Text Magnifier and Simplifier Code Example.

Conclusion

Ensuring accessibility compliance with a focus on underrepresented cognitive disabilities is an essential step toward creating an inclusive and accessible online environment. Accessibility benefits everyone and should be at the core of every web development project. By implementing the practical strategies outlined in this paper, web developers can contribute to the development of websites and web applications that cater to the needs of all users. Web accessibility is not merely a legal requirement for federal websites but a moral imperative for all, and it is incumbent upon organizations to embrace these strategies to make the internet a more inclusive space for everyone.

Works Cited

Druss, Benjamin G., et al. "Understanding disability in mental and general medical conditions." *American Journal of Psychiatry* 157.9 (2000): 1485-1491.

Henry, Shawn Lawton. "Just ask: Integrating accessibility throughout design." A List Apart, 4 Mar. 2008, https://alistapart.com/article/justaskintegratingaccessibilitythroughoutdesign/.

"How to use accessibility widgets in order to improve usability." Overlay Facts Org https://overlayfacts.org/blog/how-to-use-accessibility-widgets-to-improve-usability/ 26 Aug. 2022.

"Inclusive teaching: Intellectual disability." *ADCET* www.adcet.edu.au/inclusive-teaching/specific-disabilities/intellectual-disability. 9 Oct. 2023.

"Introduction to web accessibility." *Web Accessibility Initiative (WAI)* - W3C. https://www.w3.org/WAI/fundamentals/accessibility-intro/ 28 Sept. 2023.

Kc, Adasc. "Guide to cognitive disabilities." *ADA Site Compliance*, ADA Site Compliance, 15 Nov. 2022, https://adasitecompliance.com/guide-to-cognitive-disabilities/.

"Making content usable for people with cognitive and learning disabilities." Edited by Lisa Seeman-Horwitz et al., *W3C*, W3C, 29 Apr. 2021, www.w3.org/TR/coga-usable/.

"Mental health disorder statistics." *Mental Health Disorder Statistics | Johns Hopkins Medicine*, 1 Feb. 2023, www.hopkinsmedicine.org/health/wellness-and-prevention/mental-health-disorder-statistics.

Pichiliani, Talita Cristina Pagani Britto, and Ednaldo Brigante Pizzolato. “Underrepresented cognitive disabilities and web accessibility: a survey into the Brazilian web development community.” *Journal on Interactive Systems* 12.1 (2021): 308-327.

Potter, Carsen., et al. “Web accessibility for common websites.” *Festival of Excellence Conference*, 31 March. 2023, Cedar City.

THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

Which AI Systems Create Accurate Alt Text for Picture Books?

Quan (Monica) Zhou, Nathan Samarasena, Kyle Han, Nataly López, Stacy M. Branham
University of California, Irvine
quanz13@uci.edu, nsamaras@uci.edu, kyleh1104@gmail.com,
natalyisalopez@gmail.com, sbranham@uci.edu

Abstract

In the last decade, there has been a surge in development and mainstream adoption of Artificial Intelligence (AI) systems that can generate textual image descriptions from images. However, only a few of these, such as Microsoft's SeeingAI, are specifically tailored to needs of people who are blind screen reader users, and none of these have been brought to bear on the particular challenges faced by parents who desire image descriptions of children's picture books. Such images have distinct qualities, but there exists no research to explore the current state of the art and opportunities to improve image-to-text AI systems for this problem domain. We conducted a content analysis of the image descriptions generated for a sample of 20 images selected from 17 recently published children's picture books, using five AI systems: asticaVision, BLIP, SeeingAI, TapTapSee, and VertexAI. We found that descriptions varied widely in their accuracy and completeness, with only 13% meeting both criteria. Overall, our findings suggest a need for AI image-to-text generation systems that are trained on the types, contents, styles, and layouts characteristic of children's picture book images, towards increased accessibility for blind parents.

Keywords

AI, alt text, image description, blind, screen reader user, picture book illustrations

Introduction

Advances in technology have enabled people with vision disabilities the ability to access information in digital images using AI-generated alternative texts. Alternative text, sometimes called alt text, provides a textual description of digital images that can be read aloud by a screen reader or a refreshable braille display ("Guideline 1.1 - Text Alternatives"). Yet, these AI-generated solutions are not tailored to the scenario where blind parents want to read with their sighted children. First, studies show that when blind parents read with children, they desire access to many details of the images that are typically not included in alternative text, in order to discuss the images together (Park et al.). Second, prior research shows that children's picture book images have distinct features that are not often represented in AI training sets, such that AIs may not be able to accurately detect objects in such images (Hicsonmez et al.).

Co-reading is an important scenario to design for, as it contributes to children's cognitive and literacy skills development (Mason; Fletcher and Reese). When children listen to their parents reading storybooks, they are exposed to new words in a meaningful context, which brings them "a larger, more fully featured oral vocabulary" (Mason). Images in picture books are also important; labeling objects in the images increases children's exposure to novel vocabulary and concepts, as well as helping parents guide children's attention and participation (Fletcher and Reese). Therefore, when blind parents co-read with sighted children, they desire access to not only the print text, but also the images, so they can engage their child in educational dialogues about the visuals (Park et al.).

Unfortunately, initial studies suggest that blind parents do not find images in current reading technologies very accessible (Park et al.; Storer and Branham). There have been several studies aimed at improving co-reading experiences, but none of them are designed to enable parents with vision disabilities to read with their sighted children. Attarwala et al. designed a listening and talking e-book, which allows sighted users to audio record themselves reading a book, and people who are blind can later listen to the recording while viewing the book in large digital text. There are also several studies focusing on sighted parents reading to children with vision disabilities, utilizing tactile books to help children feel and learn about images (Kim and Yeh). Some devices that are designed for the non-disabled population show promise for interactive reading applications. For example**,** Zhang et al.'s research presents an AI-enabled system, StoryBuddy, which can automatically generate dialogic questions about the story for children without the parent being involved. However, research suggests that blind parents do not want to be cut out of the reading process (Park et al.; Storer and Branham). Methods that provide access to rich images in children's picture books—to facilitate synchronous, accessible (for blind parents), yet visually oriented (for sighted children) co-reading experiences—are still needed. Therefore, in this paper, we explore opportunities for AI to generate image descriptions as opposed to replacing the parent in co-reading interactions.

A growing body of research has been exploring automatic alt text generation and optimization using AI image-to-text generation software (Mack et al.; Kreiss et al.; Gleason et al.). The models used by these systems are based on a wide range of image sources, including Wikipedia images (Kreiss et al.) and social media images from Facebook (Wu et al.) and Twitter (Gleason et al.). However, according to a study of AI recognition of different illustrator's styles in children's picture books, the imaginary nature of these formats may lead to "extreme characters and settings," such that illustrations are distinct from common images (Hicsonmez et al.). This begs the question: which AI systems create accurate alt text for picture books? There has yet to be research that explores the current state of the art and opportunities to improve AI image-to-text generation systems for this problem domain.

To address this gap, we conducted a content analysis (Krippendorff) of five current AI systems: asticaVision, BLIP, SeeingAI, TapTapSee, and VertexAI. We then selected a random sample of 17 recently published children's picture books from popular recommendation lists (e.g., The New York Times Best Children's Books of 2022). We extrapolated a total of 669 digital images from the EPUBs of these books offered by Google Play Books. Then, we selected a subset of 20 images that represented a range of image *types* (i.e., decorative, informative, complex), *contents* (i.e., people/animals, abstract/patterns, background/scenery), *styles* (i.e., cartoon, realistic, stylized), and *layouts* (e.g., half page, single page, full spread, panels). We ran these images through each AI system, generating 100 descriptive texts, and then we conducted quantitative and qualitative analysis to ascertain the quality of these descriptions according to W3C's and similar image description guidelines.

Methods

Our content analysis had three phases. First, we identified a sample of digital children's book images (Phase 1). Next, we ran a search to identify viable AI tools (Phase 2). Finally, we conducted the content analysis on the generated text (Phase 3).

Phase 1: Digital Image Selection

This work borrows the digital children's book selection process used by Jinseo Kim in his 2023 master's thesis titled "Are Digital Children's Books Accessible to Blind Parents with Sighted Children?" (Kim). Because blind parents report lack of access to the wide variety of recently published books available in print, Seo identified popular books from a range of lists (e.g., New York Times from 2015 to 2022, NPR's 100 Children's Books list). Only 17 of these

120 children's books were available from a mainstream eBook vendor (i.e., Google Books). He then extracted 669 digital images from the books for accessibility research purposes. From this set of images, our research team selected a subset of 20 images that represented a range of image *types* ("Decorative Images"), *contents* (identified through inductive thematic analysis), *styles* (Guru Staff), and *layouts* (Ferreira; "Panel (Comics)"). The categories and examples of how they were applied to particular images can be found in Tables 3, 4, 5, and 6 respectively.

Phase 2: AI Tool Selection

Next, we conducted a search for off-the-shelf AI image-to-text generation software using both the Google search engine and Apple App Store search. Our search strings were: "image captions ai," "alt text ai," "alt text app," "image description app," and "computer vision." We identified 14 AI tools using this method. We then filtered our list down to five viable AI systems, including those that had distinct output and were therefore using distinct underlying AI systems, those that most reasonably adhered to W3C guidelines for alt text ("Technique G94"), and those that were specifically designed for blind people (TapTapSee and SeeingAI). Figure 1 depicts the details of our filtering process, which led us to ultimately select the following five AI systems: VertexAI by Google, BLIP by Salesforce, SeeingAI by Microsoft, TapTapSee by Cloughtsight, Inc., and asticaVision by Onomal Inc.

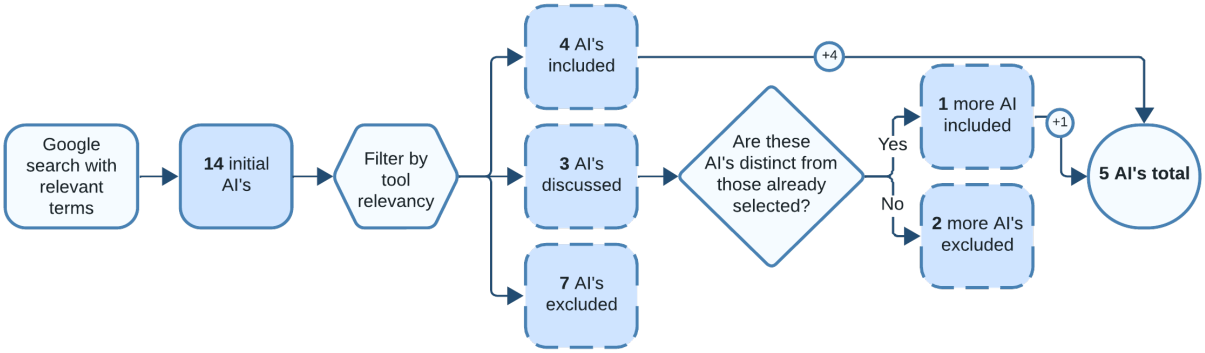


Fig. 1. Flow Chart of AI tool selection process.

Phase 3: Data Collection and Analysis

We used each of the five AI systems to generate image descriptions for each of the 20 selected images, resulting in 100 image descriptions. We then qualitatively assessed the accuracy and completeness of each description. Accurate descriptions are those that are unlikely to lead to misunderstandings about the story content; complete descriptions are those that address each part of the image that contains useful information in the context of reading a children's book. The lead author applied this schema to all images, discussing and iteratively refining its application in conversation with the last author. Table 1 shows five examples of how this schema was applied. Finally, we used descriptive statistics to identify trends in images and image descriptions.

Table 1. Image classifications, AI descriptions & quality assessment data for five sample images.

Metadata	**Image #1**	**Image #7**	**Image #13**	**Image #9**	**Image #14**
Image File					
Type	Complex	Complex	Informative	Decorative	Decorative
Content	Human(s) / Animal(s)	Human(s) / Animal(s)	Human(s) / Animal(s)	Abstract / Pattern	Background / Scenery
Style	Realistic	Cartoon	Cartoon	Stylized	Stylized
Layout	Spread	Panels	Single	Single	Spread
AI Description	"A drawing of a family sitting around a table" (VertexAI)	"Mickey mouse and minnie mouse wall decor" (TapTapSee)	"A cartoon character is hitting a pink speech bubble" (BLIP)	"A close up of orange dots" (SeeingAI)	"This painting depicts a landscape with trees, animals, and a house.…" (asticaVision)
Accurate / Complete?	Accurate, Incomplete	Inaccurate, Incomplete	Inaccurate, Complete	Accurate, Complete	Accurate, Incomplete

Findings

Table 2. Accuracy and Completeness of Image Descriptions for each AI system.

AI	Accurate, Complete	Accurate, Incomplete	Inaccurate, Complete	Inaccurate, Incomplete
asticaVision	4	3	8	5
BLIP	2	4	2	12
SeeingAI	3	4	2	11
TapTapSee	0	0	0	20
VertexAI	4	8	2	6

Accuracy and Completeness Across AI systems

The text generated by VertexAI showed the best performance, with 60% accuracy across all images. SeeingAI and asticaVision tied for second, with 35%, followed by BLIP with 30%. asticaVision had a 60% completeness rate, while others were less than 30%. This is because asticaVision tends to generate a detailed yet redundant description, which raises the possibility of identifying more elements in an image, thus increasing the completeness. We also calculated the "incomplete, accurate" ratio, which is the rate of being "incomplete" for all the "accurate" text. All systems except TapTapSee had "incomplete, accurate" ratios over 40%. We speculate that AI systems tended to focus on the main object and action of an image, rather than details. Surprisingly, 100% of descriptions generated by TapTapSee—an app specifically designed for blind users—were categorized as "inaccurate, incomplete." We suspect that TapTapSee has used few or no images of cartoon and abstract art for their data training, explaining why the AI system performed poorly. Also, TapTapSee's descriptions were concise, making it less likely to contain relevant elements depicted in an image.

Image Type

Table 3. Accuracy and Completeness of Image Descriptions by Image Type.

Image Type	# of Image	Accurate, Complete	Accurate, Incomplete	Inaccurate, Complete	Inaccurate, Incomplete
Decorative	6	9	6	2	13
Informative	5	4	4	4	13
Complex	9	0	9	8	28

Decorative images' generated text shows the highest accuracy, while that of informative and complex images are both below 40%. Decorative images usually contain symbolic objects and tend to precede or succeed the story itself, so they will not substantially affect readers' understanding of the story. Therefore, we did not require "accurate" image descriptions for decorative images to include the symbolic meaning, leading to high rates of accuracy among this image type. For example, Image #9 (Table 1) is composed of countless orange dots. A close reading of the story reveals that the orange dots symbolize connection and hope, as the color orange appears in scenes where the boy spends time with the dog; further, as the image appears on the last page of the book, it suggests hope that the relationship will continue. We labeled descriptions "accurate" if they mention the "large number of orange dots." In contrast, informative and complex images contain more information that are critical to the main story (e.g., Table 1., Images #1 and #7). In order to be considered "accurate," descriptions for these image types needed to be both correct and complete, leading to lower levels of accuracy among this image type.

Image Content

Table 4. Accuracy and Completeness of Image Descriptions by Image Content.

Image Content	# of Image	Accurate, Complete	Accurate, Incomplete	Inaccurate, Complete	Inaccurate, Incomplete	Accuracy
human(s) / animal(s)	13	4	8	13	40	18%
abstract / pattern	3	9	1	1	4	67%
background / scenery	4	0	10	0	10	50%

Our data suggests that image content is correlated with accuracy and completeness of the generated text. As shown in Table 4, descriptions for images that were coded as “abstract / pattern” or “background / scenery” were accurate at rates of over 50%. However, the descriptions for images coded as “human / animal” were only accurate at a rate of 18%. This may be a result of these images portraying more complex stories, including character identities and actions (e.g., Table 1, Images #1 and #7), leading to higher standards of accuracy and completeness. In contrast, images coded as “abstract / pattern” tended to have simple stories (e.g., Table 1, Image #14), which led to a lower standard of accuracy. Examining completeness for images coded as “background / scenery,” the descriptions were all determined to be “incomplete.” These visuals contained many elements and thus had an elevated standard for complete descriptions. As an example, Image #14 (Table 1) depicts a landscape with rich details, including palm trees, houses, a river, and a boat. None of the AI systems generated descriptions with all of the elements above. The description closest to “complete,” generated by asticaVision, still misses the river and boat.

Image Style

Table 5. Accuracy and Completeness of Image Descriptions by Image Style.

Image Style	# of Image	Accurate, Complete	Accurate, Incomplete	Inaccurate, Complete	Inaccurate, Incomplete	Accuracy
Cartoon	10	3	6	9	32	18%
Realistic	4	4	6	1	9	50%
Stylized	6	6	7	4	13	43%

Our data also suggest that differences in image style are correlated with differences in accuracy and completeness of generated descriptions. The text generated for images classified as "realistic" had an accuracy of 50%, the highest figure. For "stylized," accuracy fell to 43%, while that of "cartoon" - the most plentiful in our sample - was only 18%. We consider several reasons for such a result. First, "stylized" images in our sample tended to represent simple aspects of the narrative (e.g., Table 1, Image #9 and #14), making it easier for AI systems to provide accurate descriptions. "Cartoon" and "realistic" images, conversely, tended to depict complex stories that included human and animal characters, which raised the standard for accuracy. Second, some "cartoon" images tended to be abstract, leading to a wide variety of "inaccurate" descriptions generated by AI systems. For example, Image #13 (Table 1) depicts a black cat with a pink speech bubble. Yet, this was interpreted as a "green and purple owl illustration" (TapTapSee), a monster (asticaVision), and a spider (asticaVision, SeeingAI). Only one AI recognized the cat (VertexAI). Third, "cartoon" images in our sample tended to include "panel" layouts; all AI systems performed poorly on such images.

Image Layout

Table 6. Accuracy and Completeness of Image Descriptions by Image Layout.

Image Layout	# of Image	Accurate, Complete	Accurate, Incomplete	Inaccurate, Complete	Inaccurate, Incomplete	Accuracy
Spread	4	0	8	2	10	40%
Single	12	13	10	8	29	38%
Panels	4	0	1	4	15	5%

While none of the image layouts attained more than 50% accuracy on their descriptions, it is worth noting that "panel" images were by far the lowest, with a mere 5% accuracy. Further, though content in these images was accurately detected, none of the AI systems could distinguish between scenes of a paneled image. Consider Image #7 (Table 1), where both asticaVision and SeeingAI recognized "skeletons," "dog," and the corresponding action, "running." Yet, VertexAI did not detect the "running" action. Instead, it generated the *only* accurate translation of the words inside the speech bubbles. None of the AI systems could successfully identify the full story, composed of the dog chasing two skeletons and scaring them by barking.

Discussion

In our data analysis, we found several patterns which led us to believe that AI models are poor at describing children's book images. TapTapSee, an AI system designed for blind users, had 100% of its descriptions categorized as "inaccurate, incomplete." "Panel," "cartoon," and "complex" images most commonly led to poor descriptions. Prior work has shown that AI models are largely trained on utilitarian image types: Wikipedia images (Kreiss et al.) and social media images from Facebook (Wu et al.) and Twitter (Gleason et al.). Our work indicates a gap in the capabilities of AI models to generate appropriate children's book image descriptions. The gap is particularly pronounced for images with properties that are distinct or common among children's picture books (e.g., "panel" and "cartoon" images). We propose that AI models have substantial room for improvement within this domain.

For particular images, AI systems performed relatively well in accurately detecting visual content. VertexAI was the highest performer in terms of accuracy, with 60% of descriptions coded as "accurate." AsticaVision was the highest performer in terms of completeness, with 60% of descriptions coded as "complete." Unfortunately, when considering both measures, VertexAI and AsticaVision tied for top performance, with a mere 20% of descriptions coded as "accurate, complete." Research shows that when parents read with children, they do not simply want to know the presence of a character or object; they additionally want to access details like the emotions on a character's face, the actions taking place, as well as colors and shapes (Storer and Branham; Park et al.). Therefore, we consider that descriptions which are accurate but lack completeness are not viable for co-reading scenarios.

We found images in our sample that are best described as "decorative" which contain details that parents might enjoy talking about with their children. As described in Sections 2 and 3 of the Findings, Images #7 and #14 precede or follow the actual story itself. Thus, the image content is not necessary for readers to access the main story. However, as co-reading is an opportunity for children to learn, parents may want to explore this type of image with their child. The detailed landscape in Image #14, for instance, could allow parents to guide children to label the objects, practice and expand their vocabulary, and have educational dialogues around each element. With Image #7, parents may want to discuss the abstract and symbolic relationship of the orange dots to the main story. However, W3C states that decorative images "don't add information to the content of a page" ("Decorative Images"), and thus, they "require empty alternative text that convey their ornamental purpose" ("Module 4: Images"). Therefore, we argue that the decorative images in children's storybooks should be considered important in the co-reading process, though it may not add information to the story itself. In other words, decorative images *should* have alt text that is both accurate and complete.

Conclusion

While AI systems have seen tremendous growth over past years, no research studies have looked into the prospect of utilizing AI systems to generate image descriptions for children's picture book illustrations. Our study analyzed the descriptions of 20 images from 17 digital children's picture books. Five AI systems (asticaVision, BLIP, SeeingAI, TapTapSee, and VertexAI) each produced 20 descriptions. The accuracy and completeness of descriptions varied greatly by AI system, with variations apparent across image type, content, style, and layout. Our research highlights the necessity for AI systems—especially those like TapTapSee, which are widely adopted within the blind community—to be trained on children's picture books and address the needs of the blind community. We foresee a need to not only improve our AI systems, but to also revise alt text guidelines in ways that are sensitive to co-reading use cases. Progress on these fronts is the first step toward full inclusion of blind parents in the important educational and familial bonding experience of reading print books with children at home.

Work Cited

"Alternative Text." Digital Accessibility at Princeton, The Trustees of Princeton University, accessibility.princeton.edu/how/content/alternative-text. Accessed 12 Oct. 2023.

Attarwala, A., Munteanu, C., and Baecker, R. 2013. An accessible, large-print, listening and talking e-book to support families reading together. In Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services. 440–443.

"Decorative Images." W3C, 27 July 2019, www.w3.org/WAI/tutorials/images/decorative/.

Ferreira, Karen. "Types of Illustrations for Children's Books." Get Your Book Illustrations, 7 Feb. 2020, getyourbookillustrations.com/types-of-illustrations-for-childrens-books/.

Fletcher, K. L. and Reese, E. 2005. Picture book reading with young children: A conceptual framework. Developmental Review, 25(1), 64–103. https://doi.org/10.1016/j.dr.2004.08.009

Gleason, Cole, et al. "Twitter A11y: A browser extension to make Twitter images accessible." Proceedings of the 2020 chi conference on human factors in computing systems. 2020.

"Books on Google Play." Google, Google, play.google.com/store/books?hl=en_US&gl=US.

"Guideline 1.1 – Text Alternatives." W3C. 2019. https://www.w3.org/WAI/WCAG21/Understanding/text-alternatives

Guru Staff. "Types of Illustration Styles for Children's Books." Guru Insights, 4 Nov. 2021, www.guru.com/blog/types-of-illustration-styles-for-childrens-books/.

Hicsonmez, S., Samet, N., Sener, F., and Duygulu, P. 2017. Draw. Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval. https://doi.org/10.1145/3078971.3078982

Onamal Inc. "Cognitive Intelligence API: See, Speak, and Hear with Astica." Astica AI, https://astica.ai/.

Kim, J. (2023). "Are Digital Children's Books Accessible to Blind Parents with Sighted Children?" Master's thesis, University of California, Irvine.

Kim, Jeeeun and Yeh, Tom. 2015. Toward 3d-printed movable tactile pictures for children with visual impairments. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pages 2815–2824, 2015.

Krauss, Jennifer. "The Best Children's Books of 2022." The New York Times, The New York Times, 8 Dec. 2022, www.nytimes.com/2022/12/08/books/review/the-best-childrens-books-of-2022.html.

Kreiss, Elisa, Goodman, N. D., and Potts, C. 2021. "Concadia: Tackling image accessibility with context." https://arxiv.org/pdf/2104.08376v1)%3C%22

Krippendorff, Klaus. (2018). Content Analysis: An Introduction to Its Methodology, 4th ed., SAGE Publications, pages 51–52.

Li, Junnan, et al. "Blip: Bootstrapping Language-Image Pre-Training for Unified Vision-Language Understanding and Generation." Salesforce AI, 5 Nov 2022, https://blog.salesforceairesearch.com/blip-bootstrapping-language-image-pretraining/.

Mack, K., Cutrell, E., Lee, B., and Morris, M. R. 2021. Designing tools for high-quality alt text authoring. The 23rd International ACM SIGACCESS Conference on Computers and Accessibility. https://doi.org/10.1145/3441852.3471207

Mason, Jana M. 1990. Reading Stories to Preliterate Children: A Proposed Connection to Reading. Technical Report No. 510.

“Module 4: Images.” W3C, 29 Sept. 2022, www.w3.org/WAI/curricula/content-author-modules/images/.

“Panel (Comics).” Wikipedia, Wikimedia Foundation, 5 Aug. 2023, https://en.wikipedia.org/wiki/Panel_(comics)#:~:text=A%20panel%20is%20an%20individual,drawing%20depicting%20a%20frozen%20moment.

Park, S., Cassidy, C., and Branham, S. M. 2023. “It’s All About the Pictures: Understanding How Parents/Guardians with Visual Impairments Co-Read with Their Child(ren).” In The 25th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS ’23), October 22–25, 2023, New York, NY, USA. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3597638.3614488

“Seeing AI App from Microsoft.” Microsoft, www.microsoft.com/en-us/ai/seeing-ai

Storer, Kevin and Branham, Stacy M. 2019. “That’s the Way Sighted People Do It” What Blind Parents Can Teach Technology Designers About Co-Reading with Children. In Proceedings of the 2019 on Designing Interactive Systems Conference. 385–398.

“TapTapSee - Blind and Visually Impaired Assistive Technology - powered by CloudSight.ai Image Recognition API.” TapTapSee, https://taptapseeapp.com/.

“Technique G94: Providing Short Text Alternative for Non-Text Content That Serves the Same Purpose and Presents the Same Information as the Non-Text Content.” W3C, 20 June 2023, www.w3.org/WAI/WCAG21/Techniques/general/G94.

“Vertex AI.” Google, https://cloud.google.com/vertex-ai?hl=en.

Vezzoli, Y., Kalantari, S., Kucirkova, N., and Vasalou, A. 2020. Exploring the design space for parent-child reading. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3313831.3376696

Wu, S., Wieland, J., Farivar, O., and Schiller, J. 2017. Automatic alt-text. Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. https://doi.org/10.1145/2998181.2998364

Zhang, Z., Xu, Y., Wang, Y., Yao, B., Ritchie, D., Wu, T., Yu, M., Wang, D., and Li, T. J.-J. 2022. StoryBuddy: A human-ai collaborative chatbot for parent-child interactive storytelling with flexible parental involvement. CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3491102.3517479

2024
The Journal on Technology and Persons with Disabilities

Non-Discrimination, not Copyright Exceptions

Abigail Rekas
University of Galway
abigail.rekas@universityofgalway.ie

Abstract

This paper examines the right to access cultural materials for persons with disabilities and proposes a new approach to implementing that right. It takes a socio-legal approach to critically examine the state of the law and evaluate whether the goals of the Americans with Disabilities Act (ADA) are being met in light of changing understandings of disability.

Keywords

ADA, copyright exception, access to culture, CRPD, Marrakesh, accessibility, A11Y

Introduction

Access to culture is a right. This is not a particularly controversial statement, and yet, access to cultural materials has historically been a challenge faced by persons with disabilities. The full realization of the rights of persons with disabilities has long been one of the goals of the disability rights movement. (*A Brief History of the Disability Rights Movement | ADL*) Unfortunately, despite some exciting technical innovations, particularly those linked to the digitization of cultural materials, and significant legal interventions, such as the Americans with Disabilities Act (*ADA*) promise that persons with disabilities should not be excluded from places of public accommodation (*ADA Title III*) or the adoption of the Marrakesh Treaty to facilitate access to published works for persons who are blind, visually impaired or otherwise print disabled (*Marrakesh Treaty*), this challenge remains insufficiently addressed. Progress has been made, but it is by no means complete or sufficient. In part, this is due to the fact that many persons with disabilities need adapted versions of works, which can implicate copyright and can also involve expensive aftermarket interventions on the part of charitable organizations. The shift in policymaking from the charitable model to the social or human rights models of disability highlights why exception-based access to culture should no longer be the approach adopted by the United States. Instead, a non-discrimination standard should be adopted, and the definition of "public accommodation" should be broadened to include content that has been made available to the public.

Discussion

A note about language: I am committed to being led by, respectful, and affirming of every individual's choice of language about themselves, and while I use "person-first" language, this choice reflects the language of the law rather than a statement of values.

One of the key results of the disability rights movement has been a changing understanding of disability. Policy has shifted from medical or charitable models, which focused on "fixing" the individual as far as possible and provision of services through charitable interventions where fixing was impossible, to a social or human rights model, which focuses on disabling barriers created by society that hinder the full enjoyment of the human experience by persons with disabilities (Lawson and Beckett). Societal barriers, like an architect's choice to only have steps to access a building, can disable a person who would otherwise be able to enter the building. The disability is created by that choice rather than the physical characteristics of the person. These models are frameworks for understanding disability and disabling factors. While they tend to provide a simplistic view of a complex issue, they can be helpful when envisioning broader policy changes (Oliver). The ADA recognizes that societally created barriers should be removed, where it is reasonably feasible, to enable persons with disabilities to participate fully in the life of the community.

Copyright

However, one area which this shifting understanding has not noticeably changed is that of access to cultural content. For the purposes of this paper, the term "cultural content" is taken to include creative or original "works of authorship, fixed in a… medium of expression, from which they can be perceived." (*Copyright - Subject Matter*) In other words, the kinds of works that can be protected by intellectual property, specifically copyright, a category of works that encompasses the vast majority of cultural expression over the past several hundred years. Traditionally, creating accessible versions of copyright-protected cultural content has been a labor-intensive and expensive endeavor, for instance, printing multi-volume braille versions of individual books or pressing multiple vinyl records for audiobooks. At one time, these technologies were the state of the art. However, as technology has moved forward, adaptation of content has become simpler and less expensive to create. Assistive technologies, like refreshable braille output, audio description files, sign language interpretation plug-ins, or alt-text described images, have been so enabling that the present time has become the first time in history “where it is truly possible to include everyone” (David Berman).

While the technology has moved rapidly forward, the enabling framework of law has not. This content is protected by copyright for a very long term, generally the life of the author plus an additional 70 years (*Copyright Duration*). This protection gives the owner of the right the ability to prevent others from copying or creating works based on the original work (*Copyright Exclusive Rights*), which potentially includes making an accessible version of the work. However, there are a set of limitations and exceptions, the most relevant being the Chafee Amendment (*Chafee Amendment*) and fair use (‘Copyright Limitations’), which allow users to create accessible format versions of works without the consent of the author or publisher of the work. This framework allows for private accessible copies to be made but also for third parties, mainly libraries and civil society groups representing disabled persons, to adapt and modify cultural content on behalf of disabled consumers.

This exception-based approach means that the publishers, movie studios, video game studios, and other players in the cultural content industries have no legal obligation to ensure accessibility, and the burden of making works accessible falls on government, libraries, and charitable third parties, perhaps the best known being Benetech Bookshare which has the largest library of accessible ebooks in the world ('Bookshare'). It also means that persons with disabilities are subject to the whims and financial constraints of these third parties in determining what cultural content is available to them. This is why, while wonderful, the Marrakesh treaty will always fall short of full access, leaving aside the fact that it only covers printed material like books. The Marrakesh treaty is premised on works created by third parties, authorized by a copyright exception, moving around the globe. The numbers back this up, as it is estimated that from the 5-7% of annually published books made accessible before the Marrakesh Treaty came into force, that number has not increased significantly and still hovers under 10% (*WBU*). This is even with significant numbers of publishers sharing files with libraries and charities that serve persons with disabilities ('ABC Accessible Publishing').

The other drawback of the present system is not precisely copyright but is strongly related. The Digital Millennium Copyright Act (*DMCA*) imposes strict rules about when the digital locks around a particular work may be broken (*Copyright - Circumvention*) and requires that those who want an exemption from these rules must apply triennially to be able to break copyright management systems legally to enable accessibility (US Copyright Office).

Regulation

While copyright exceptions are the main way works like books are made accessible, there have also been some regulatory interventions that have had an impact on the accessibility of cultural content. For instance, there is a broadcasting regulation that requires some audiovisual media content to be provided with closed captioning and audio description, but generally only if its initial broadcast was on linear television (*CVAA*); purely online content is generally not covered. There are also regulations based on the ADA; for instance, public accommodations that display cultural materials, like movie theaters and museums, are required to be physically accessible (*ADA Title III Definitions*). Interestingly, while there is an obligation on the operators of cinemas to ensure they can display closed captions and audio descriptions and have the devices necessary to enable those functions on hand (‘DOJ Non-discrimination - Movie Theaters’), under law, there is no obligation on the studio that has released the film to include code in the movie file that can actually interact with those devices.

There has also been a patchwork of federal cases over the accessibility of websites, based on the notion that they are a service of a public accommodation, for instance, the Netflix Settlement, which saw the platform streamer agree to significantly increase the availability of accessible content present on its platform (‘Netflix Settlement’; ‘Netflix Is the New Accessible’). These cases, unfortunately, are the subject of a circuit split, with various federal appeals courts disagreeing over whether or not there needs to be a connection with a physical facility, like a storefront, in order to make a website a service of a public accommodation and subject to accessibility requirements. This obligates primarily retail-based websites in some jurisdictions but, again, does not put the obligations on the publishers, producers, or copyright holders to ensure that the content they release to the public is accessible.

The entertainment and media industry were reported to have generated revenue of around $2.3 trillion globally in 2022 (PWC); that revenue was generated by releasing copyright-protected content to the public. Some of it was released with some accessibility functionality; some of it was not. Some was made accessible by third parties; for instance, Disney+ content has been made accessible for children who are d/Deaf or hard of hearing and communicate with sign language, using an accessibility plug-in for Google Chrome called SignUp. SignUp was built by a 17-year-old on her own time and is available for free (Shanfeld).

Accessibility is caught in a paradigmatic loop, where it has become difficult to step outside the self-completing circle of logic. Exceptions and limitations, with a few media regulations, plus producer goodwill and the pressure of public opinion, are what drive accessible content creation. This approach has been exported worldwide, with even the UN Convention on the Rights of Persons with Disabilities Article 30.3 including a provision that intellectual property rights should not create “a barrier to access by persons with disabilities to cultural materials” (*CRPD*). This provision implies that members of the Convention should ensure there is an exception built into their copyright regimes. What has been ignored is that this is simply an extension of the charitable model of accessibility, where persons with disabilities are expected to be grateful for what is made available to them. Even calling it an exception highlights this fact: Exceptions are not the norm by their very definition. If one truly believes there is a right to access culture, one must step outside the paradigmatic loop and look for new solutions.

A New (Accessible) Hope

One such solution could be achieved by expanding the definition of public accommodation under the ADA to include cultural materials that are made available to the public. Rather than focusing on whether websites are “places” or must have a tie to a storefront, the law should be amended to address what makes those accommodations worth visiting. Looking at the value of the listed public accommodations, one may find the semantic meaning of the guarantee of access: one does not go to the cinema to be able to get in the front door or be able to use the bathroom; one goes to the cinema to experience the film! It does not matter if the cinema has the headset to let one access the audio description if the film file does not have the audio description in its code. There must be a technological handshake between the content and the hardware. However, without a mandate to include that functionality, to make the content "born accessible," that handshake will not happen.

One way to make that handshake happen would be to impose a duty on those who make cultural content available to the public and thereby ensure that when that content is opened to the public, it is opened to the broadest segment of the public possible by applying the principles of Universal Design to cultural content (Connell et al.). Including cultural works as public accommodations brings these works under the purview of the ADA Title III, which traditionally governs places open to the public and allows courts and companies to apply a single accessibility standard to all works, subject to an undue burden analysis. This single standard is one with which policymakers, courts, and the public are familiar. Rather than drawing up a new law, it applies an old law governing public accommodation to a twenty-first-century understanding of the term.

Challenges

While the undue burden analysis is subject to criticism that it provides an escape hatch or an exception to the obligation to ensure accessibility for a public accommodation, when one reviews the cases, the evidence for this is not strong, and much of the criticism arose in the early '90s, the earliest days of the ADA (see, e.g. (Jones; Porter). There are two reasons for continuing to use the undue burden analysis. Firstly, it is the present approach under the ADA Title III, which means that courts and lawyers are used to applying it, and businesses are aware of what it means, making it legally predictable. The second is to address concerns of resources in the age of user-generated content. Many people are small content producers who regularly make TikTok, Instagram, or YouTube videos, making little or no money from their content generation. The obligations placed on those creators need to be correspondingly small, perhaps only opt-out prompts on the platform – those modifications that are “readily available” (*ADA Title III*), in order not to chill speech. That accessibility might interfere with free speech is a criticism raised with accessibility mandates in the broadcasting sphere; for instance, CNN successfully raised the First Amendment as a defense for failing to caption their online web videos in California (*GLAAD, Inc. v. CNN*; Gardner).

While well-resourced entities may be obliged to include a wide variety of accessibility functions, less well-resourced entities or individuals will have less obligation. Anecdotally, when discussing this proposal, the most common response is, “Well, what about someone like me?” the defense provided by the undue burden analysis answers that question. The Department of Justice issues a Technical Manual, which provides guidance for businesses governed by Title III and addresses the factors considered in an undue burden analysis. This manual explicitly refers to the size and resources of the regulated entity (US DOJ) as a factor to be considered.

Another question raised has to do with different public or quasi-public institutions. In the first category, we find educational institutions like universities. This obligation is primarily targeted at private industry, as much educational content is already required to be made accessible by Title II, which regulates public bodies, and course content is generally covered by accessibility mandates, which require schools to provide equal access to all education benefits regardless of disability (*ADA Title II*; *Technology Accessibility*). Student work would not be considered made available to the public unless it was opened beyond the context of the classroom.

In the second category, we find institutions like museums, which present two potential challenges. The first is the nature of their collections; museums primarily display works of visual art, which do not have the same ability to incorporate digital tools into their underlying code –

unless and until they are made digital. Many tools have been used over the years to make museum collections accessible, and the considerations of Museum accessibility are nothing new. The first recorded intervention to benefit persons with disabilities in the US occurred at the American Museum of Natural History in New York City in 1909 (Claudia Haines). These days, digitizing collections is one element of museum curation, and that digitization allows for metadata such as alt-text to be added to the digitized work. 3D printers allow for tactile interaction. In museum practice, Braille displays and headsets describing the art are commonplace. These are all ways for the museum's content, a museum being one of the public accommodations explicitly listed by the ADA, to be made accessible. This content can be made accessible, even when not primarily expressed in a digital medium. Access to the content speaks to the value of the museum as a public accommodation.

The second challenge comes from the language of the proposal, "those who open works to the public," which implies initial publishing. However, the museum acts as a gatekeeper. It controls public access to the work, including charging for access, charting for prints, and limiting the ability of the public to make copies or take photographs. As this is the case, they comfortably stand in the shoes of the publisher of the work and assume some of the obligations of that actor. Further, this is another reason for keeping the undue burden language of Title III, as it can accommodate the fact that many museums are underfunded and poorly resourced.

The last challenge worth mentioning in this short policy recommendation is the burden placed on persons with disabilities. This approach is premised on the ability of a disabled person to make a complaint that the cultural material is not accessible. It means that the burden of policing is placed on the individual rather than on the state. The impact of anti-discrimination litigation on persons with disabilities is a seriously under-researched area. It is possible to draw parallels between reasonable accommodation lawsuits in the employment context, which indicate that there are emotional costs to asserting one's rights, even when doing so is justified (*Burden vs. Entitlement*; Levy). Setting up a complaints procedure and issuing clear and concrete regulatory guidance is necessary when bringing cultural content under the umbrella of public accommodation, and lessons should be taken from the challenges of workplace accommodation regulation.

Conclusion

This paper has given a brief introduction to the idea that in order to fully realize access to culture as a right and to give semantic meaning to the ADA promise of accessible museums, concert halls, libraries, galleries, cinemas, and so on, policymakers should consider a shift from copyright exceptions as the norm to a non-discrimination framework incorporating content into the Americans with Disabilities Act Title III definition of cultural materials. Due to space constraints, much of this argument has been presented in brief, and the challenges mentioned are just some that this policy recommendation will likely face. It is to be anticipated that the large content-producing industries will resent any policy intervention that will take away from the bottom line, and the lobby will inevitably be strong. However, increasing access to cultural materials seems to have plateaued, although the assistive technologies continue to race forward. Moving out of the self-completing loop of regulation and exception may kickstart a greater focus on access for everyone to cultural materials and shift the burden of making that content accessible to an industry that is best poised to shoulder it.

Works Cited

Case law, Statutes, and Treaties

GLAAD, Inc. v. CNN. 12-15807, https://law.justia.com/cases/federal/appellate-courts/ca9/12-15807/12-15807-2014-02-05.html.

Americans with Disabilities Act of 1990, 42 U.S.C. §§12101-12213. https://www.law.cornell.edu/uscode/text/42/chapter-126.

ADA Title II 42 U.S.C. §§12131-12165.

ADA Title III 42 U.S.C. §§12181-12189.

Copyrights 17 U.S.C. §§121-1511 https://www.law.cornell.edu/uscode/text/17

Subject Matter of Copyright: In General 17 U.S.C. §102

Exclusive Rights in Copyrighted Works 17 U.S.C. §106

Limitations on Exclusive rights: Fair Use 17 U.S.C. §107

Chafee Amendment - Limitations on Exclusive rights: Reproduction for Blind or Other People with Disabilities 17 U.S.C. §121

Duration of Copyright: Works Created on or After January 1 1978. 17 U.S.C. §302

Digital Millenium Copyright Act (DMCA) - Circumvention of Copyright Protected Systems 17 U.S.C. § 1201

Twenty-First Century Communications and Video Accessibility Act of 2010 (CVAA), Pub. L. No. 111-260, 124 Stat. 2751 (2010) (CVAA), also at https://www.fcc.gov/accessibility; codified in various locations in 47 U.S.C.

United Nations Convention on the Rights of Persons with Disabilities G.A. Res 61/106, U.N. Doc. A/RES/61/106. (Jan. 24, 2007).

World Intellectual Property Organization (WIPO) Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired, or Otherwise Print Disabled, VIP/DC/8, 27 June 2013, [Marrakesh Treaty], available at: http://www.wipo.int/dc2013/en/.

Manuals and Technical Guidance

Department of Justice (DOJ). *The Americans with Disabilities Act Title III Technical Assistance Manual § III-4.2000 Reasonable Modifications*. 1994, https://www.ada.gov/reachingout/title3l1.html. Accessed 11 Aug. 2021.

DOJ 'Nondiscrimination - Movie Theaters'. *Federal Register*, Web, 2 Dec. 2016, https://www.federalregister.gov/documents/2016/12/02/2016-28644/nondiscrimination-on-the-basis-of-disability-by-public-accommodations-movie-theaters-movie. Accessed 24 Feb. 2020.

Department of Education (ED) *Technology Accessibility*. 19 May 2023, https://www2.ed.gov/about/offices/list/ocr/frontpage/pro-students/issues/dis-issue06.html. Accessed 9 Oct. 2023.

US Copyright Office. *Rulemaking Proceedings Under Section 1201 of Title 17 | U.S. Copyright Office*. Web, https://www.copyright.gov/1201/. Accessed 9 Oct. 2023.

Scholarly Work

Connell, Betty Rose, et al. *The Principles of Universal Design*. The Center for Universal Design - North Carolina State University, 1997, https://projects.ncsu.edu/design/cud/pubs_p/docs/poster.pdf. Accessed 26 Apr. 2019.

Haines, Claudia. 'Please Touch: The History of Museum Accessibility for Blind Visitors'. *Museum Studies at Tufts University*, https://sites.tufts.edu/museumstudents/2021/07/27/please-touch-the-history-of-museum-accessibility-for-blind-visitors/. Accessed 14 Sept. 2021.

Jones, Augustus J., Jr. 'Federal Court Responses to State and Local Claims of "Undue Burden" in Complying with The Americans with Disabilities Act'. *Publius: The Journal of Federalism*, vol. 25, no. 3, Jan. 1995, pp. 41–54.

Lawson, Anna, and Angharad E. Beckett. 'The Social and Human Rights Models of Disability: Towards a Complementarity Thesis'. *The International Journal of Human Rights*, vol. 25, no. 2, Feb. 2021, pp. 348–379.

Levy, Nathaniel. 'Workplace Ableism Is a Problem for ADA Rights'. *OnLabor*, 19 June 2019, https://onlabor.org/workplace-ableism-is-a-problem-for-ada-rights/. Accessed 11 Aug. 2021.

Oliver, Mike. 'The Social Model of Disability: Thirty Years On'. *Disability & Society*, vol. 28, no. 7, Oct. 2013, pp. 1024–1026, doi:10.1080/09687599.2013.818773.

Porter, N. 'A New Look at the ADA's Undue Hardship Defense'. *Missouri Law Review*, Jan. 2019. *Semantic Scholar*, https://www.semanticscholar.org/paper/A-New-Look-at-the-ADA%27s-Undue-Hardship-Defense-Porter/f1360b5477ba7ad93c0bf7e77a97f1a6de3d8e4d. Accessed 9 Oct. 2023.

Internet Resources

A Brief History of the Disability Rights Movement | ADL. Web, https://www.adl.org/resources/backgrounder/brief-history-disability-rights-movement. Accessed 9 Oct. 2023.

‘ABC Accessible Publishing’. *Abc*, Web, https://www.accessiblebooksconsortium.org/publishing. Accessed 9 Oct. 2023.

‘Bookshare’. *Benetech | Software for Social Good*, Web, https://benetech.org/work-area/bookshare/ . Accessed 9 Oct. 2023.

Burden vs. Entitlement. https://dcp.ucla.edu/node/88. Accessed 11 Aug. 2021.

Berman, David. *12 Things You Need to Know About WCAG 2.1: How It Impacts Your Work & Laws Around the Globe*. https://www.3playmedia.com/resources/recorded-webinars/wbnr-10-01-2020-wcag/. 3 Play Media Webinar.

Gardner, Eriq. ‘CNN Gets First Amendment Victory in Video Captioning Dispute’. *The Hollywood Reporter*, 5 Feb. 2014, https://www.hollywoodreporter.com/business/business-news/cnn-gets-first-amendment-victory-677446/. Accessed 17 Aug. 2021.

Markey. *Sen. Markey, Rep. Eshoo Reintroduce CVTA to Make Communication Technologies Accessible for Americans with Disabilities | U.S. Senator Ed Markey of Massachusetts*. Web, https://www.markey.senate.gov/news/press-releases/sen-markey-rep-eshoo-reintroduce-cvta-to-make-communication-technologies-accessible-for-americans-with-disabilities. Accessed 9 Oct. 2023.

‘Netflix Is the New Accessible’. *Perkins School for the Blind*, Web, https://www.perkins.org/stories/netflix-is-the-new-accessible. Accessed 23 Apr. 2019.

‘Netflix Settlement’. *Disability Rights Advocates*, Web, https://dralegal.org/case/netflix-settlement/. Accessed 28 June 2021.

PWC. 'Global Entertainment & Media Revenues Surge to $2.3 Trillion; Virtual Reality Sees 36% Growth as Gaming and Esports Are on Pace to Become a $324 Billion Business'. *PwC*, Web, https://www.pwc.com/gx/en/news-room/press-releases/2022/global-entertainment-and-media-outlook-2022-2026.html. Accessed 5 Sept. 2023.

Shanfeld, Ethan. 'New Google Chrome Extension SignUp Offers ASL Captions for Three Films on Disney Plus'. *Variety*, 29 Aug. 2021, https://variety.com/2021/digital/news/signup-asl-captions-disney-plus-1235051323/. Accessed 12 Mar. 2023.

WBU. Web, https://worldblindunion.org/programs/accessibility/accessibility-go-a-guide-to-action/. Accessed 9 Oct. 2023.

THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

D2L's Process for Developing an Accessibility Maturity Model

Aggi Jeemon, PhD, Lindsay McCardle, PhD, Laura Windhorst, Kate Wilhelm, Sambhavi Chandrashekar, PhD
D2L Corporation
Aggi.Jeemon@d2l.com, Lindsay.McCardle@d2l.com, Laura.Windhorst@d2l.com, Kate.Wilhelm@d2l.com, Sambhavi.Chandrashekar@d2l.com

Abstract

Accessibility Maturity Model (AMM) is a framework for establishing a robust accessibility program across an organization. As a global education technology company, D2L is developing a research-based AMM internally to facilitate a sustainable and extensible accessibility program to lead our organization toward maturity and excellence. Through primary research with internal product teams, we identified themes that drive our product process. Using secondary research on prevailing AMMs, we shortlisted 10 models based on their relevance to our operational processes and business functions for comparative analysis. We then selected a base model that aligns with the themes identified through research and developed the model into the first iteration of our organization's AMM, with plans for further iterations to cover other departments in the organization.

Keywords

Accessibility maturity model, digital accessibility metrics, leadership, strategy, process maturity

Introduction

Accessibility is important to businesses specializing in the design and development of digital experiences as it holds moral, legal, and economic considerations. The moral perspective aligns with Benjamin Franklin's philosophy of 'doing well by doing good,' seeking to strike a balance between business and societal aspirations (Forbes, 2020). The legal viewpoint focuses on the risks associated with non-compliance, pushing organizations into a reactive mode due to the divergence between the literal interpretation of the law and its intended spirit. When viewed through an economic lens, organizations recognize the emergence of a market comprising 1.85 billion people with disabilities, surpassing even the population of China, holding an estimated after-tax disposable income exceeding $400 billion ("Return on Disability Report"). Consequently, organizations comprehend the significance of accommodating this potential market segment and the repercussions of failing to do so. Given the growing influence of people with disabilities and the potential for accessibility to differentiate businesses in the market, it becomes imperative to explore and discuss how accessibility can be integrated into organizations. Accessibility has evolved beyond being solely a corporate social responsibility or risk mitigation activity confined to specific departments into an organizational transformation and change management initiative that engages the entire organization. Accessibility is no longer a one-time project but a program that entails continuous efforts to enhance business and operational functions, fostering consistency in practices and maturity in processes.

The Accessibility Maturity Model (AMM) is a framework for establishing a robust accessibility program across an organization. AMMs could be related to diversity inclusion or process capability in product building. Our focus is on the latter. It is normative and does not set mandatory requirements for an organization to follow. It is intended to be independent of the requirements set forth in relevant technical accessibility standards, such as the Web Content Accessibility Guidelines known as WCAG (W3C Accessibility Standards Overview) and is complementary to Accessibility Conformance testing of products (W3C Group Draft Note, 2023).

D2L is a global education technology company with a vision to make learning accessible, engaging and inspiring. We have expertise and best practices in incorporating accessibility in the product process. To achieve a robust product process with seamless accessibility practice in every stage, we began building an AMM. As a first step, we conducted research in the Technology, Engineering and Design department among teams involved in crafting product experiences to gain insights about ideal accessibility practices and explored prevailing maturity models in accessibility in the literature. This paper describes the outcomes our research. Following this, we plan to work with other relevant departments to finetune the AMM and implement a sustainable and extensible organization-wide accessibility program.

Research Methodology

We adopted a case study methodology to explore our accessibility processes in-depth using feasible and ethical qualitative methods for primary research, supplemented by secondary research in the form of a narrative literature review (Priya, 2021). Our research objectives were to:

1. Explore the available maturity models in accessibility practice,
2. Select a base model for reference that aligns with the relative operational and business functioning of our organization, and
3. develop the base model into D2L's AMM by identifying relevant dimensions.

Aligned with the objectives, we sought to answer the following research questions:

1. Which are the prevailing AMMs relevant to digital technology organizations?
2. Which AMM suits best as a base model for D2L?
3. How can that model be customized and developed into D2L's AMM?

We attempted to answer the first research question through a narrative literature review of AMMs in the digital technology, education, and education technology industries. To get a holistic view of accessibility benchmarking in the market, we reviewed 16 AMMs relating to process capability, disability inclusion and content accessibility. Of these, we selected 10 models relating to process capability for a comparative analysis.

To answer the second and third research questions, we collected data from internal stakeholders through facilitated group discussion sessions labelled as 'listening tours.' We gathered data through 14 sessions with UX designers, UX Researchers, Engineers, and Product Managers (*n* = 151). Prompts for discussion centered on current and ideal accessibility practices in product building. We analyzed the data using affinity mapping and arrived at the following themes: (a) Consistency in product development process; (b) Clarity in accessibility requirements; (c) Expertise in knowledge and insights; (d) Prioritization of accessibility; (e) Role clarity and ownership and (f) Organizational culture. We also conducted open-ended interviews with members from several other departments to gather further inputs based on these themes.

Discussion

In the following three subsections, we summarize the research outcomes that answer the three research questions.

Comparative Analysis of Various Accessibility Maturity Models

Out of all available accessibility maturity measuring tools reviewed in the literature, 16 models focused on software industry were reviewed in depth. Out of these 10 models were selected for comparative analysis because they had at least one dimension focused on product life cycle and had dimensions for other business processes where accessibility plays a major role.

The analysis indicated that the models constitute a range of 2 to 12 dimensions and 3 to 6 levels of maturity, with some models having sub dimensions. The structure of the models included checklists, matrix of dimensions and maturity stages, or questionnaires. These models used qualitative, quantitative and mixed indicators for analysis. Some models were developed through individual case methods for use by individual organizations while others were developed by accessibility consulting organizations and government bodies to apply broadly across the industry. Table 1 shows a comparative analysis of the 10 models.

Table 1. Comparative analysis of Accessibility Maturity Models.

Model	**Source**	**Dimensions**	**Maturity levels**	**Structure**	**Indicators for Analysis**
Accessibility Maturity Model	Business Disability Forum, 2016	10	5	Matrix	Qualitative Analysis
Accessibility Evolution Model	Microsoft, 2016	8	5	Matrix	No indicators published
Accessibility Maturity Model	W3C, 2021	7	4	Matrix	Qualitative Analysis
Policy-Driven Adoption for Accessibility	National Association of State Chief Information Officers, 2015	6	3	Questionnaire	Qualitative plus Quantitative analysis

Model	Source	Dimensions	Maturity levels	Structure	Indicators for Analysis
Strategic Accessibility Companionship	nomensa.com/aa	2 main dimensions with 3 subdimensions	-	Questionnaire	Quantitative analysis
Digital Accessibility Maturity Model	Level Access, 2015	10	5	-	Qualitative analysis
Digital Accessibility Maturity Score Card	Hassell Inclusion Limited, 2020	9	4	Questionnaire	Quantitative analysis
Usability and Accessibility Focused Process Improvement Capability Model (MODECUA)	Quintal and Macias, 2018	6 phases with 5 levels	-	Checklist	-
ISO/IEC 30071-1:2019	Jonathan Hassell, 2019	9	4	Questionnaire	Quantitative analysis
Rough Accessible Design Maturity Continuum	TPGi, 2014	-	5	-	-

Selecting W3C Accessibility Maturity Model as our base model

Of the 10 models compared above, we selected the W3C Accessibility Maturity Model (WCAG Group Draft Note, 2022) as the base to build D2L's AMM because its dimensions aligned most closely with our organizational processes and internal stakeholder responses. For example, the W3C model's dimensions *culture*, *knowledge and skills*, and *ICT development life cycle* matched closely with the themes from our research mentioned in the Research Methodology section of this paper. Further, all items in the *ICT development life cycle* dimension are helpful for D2L product teams as critical enablers of consistency and repeatability in the process, an area as identified during data analysis as needing attention. The items in Table 2 listing responses from our product teams on ideal accessibility practices in the product process resonated well with the ICT Development Life Cycle dimension of the W3C AMM.

Table 2. Ideal accessibility practices in the product process.

No.	Ideal accessibility practice
1	Consistent approach to designing, implementing, testing, and releasing accessible products
2	Design artifacts with accessibility information
3	Accessibility review with 3-legged stool at D2L (Product Management, Design, Engineering)
4	Documentation - standards, processes, guidelines, checklists
5	User testing and dev testing with people with disabilities
6	Tools and devices for manual and automated accessibility testing
7	Accessibility requirements specified in job descriptions
8	Clarity around roles and responsibilities
9	Accessibility as a product release gate
10	Accessibility learning gap assessment
11	Accessibility resources creation and knowledge sharing
12	Role-specific training plans and curriculums

Building D2L's Accessibility Maturity Model

This subsection details how we enhanced the AMM using the primary and secondary research data as well as D2L's key business processes relevant to accessibility. As described below, we included the following six dimensions in the first iteration: Product, Strategy, Culture, Competency, Customer, and Community.

Product

At D2L, the product teams follow a 3-legged stool (3LS) model, where the responsibility for product accessibility is shared across three roles: Product Management, UX Research & Design, and Engineering. This 3LS approach is not reflected in any of the existing models including the W3C model, which was a key misalignment of the model with our processes. So, we added a new dimension named Product in D2L's AMM to include the 3LS structure and reflect accessibility practices in the end-to-end process of product building.

We added three subdimensions under Product as *Product Management* (understanding accessibility market trends and competitive intelligence); *UX Research and Design* (understanding user accessibility needs and pain points to define design specifications) and *Engineering* (developing and testing for edge cases considering various disabilities and adaptive technologies.)

Strategy

Enabling products for accessibility requires the integration of accessibility criteria in all phases of a product life cycle, and other business processes where accessibility plays a role (Minnesotta IT Services). Our goal was to take a holistic strategic approach to accessibility covering other relevant areas of the business process and organization system. Having a solid strategy specific to accessibility in relevant business processes is important. An accessibility strategy helps advance organizational accessibility maturity, reducing risk and building capacity (TPGi). With a goal to ensure that accessibility is integrated into all business processes that impacts our business outcomes, we developed the dimension Strategy.

The Strategy dimension constitutes the goals, plans, processes set by organization to achieve accessibility-specific business outcomes. It has 5 subdimensions: Vision and mission, Policy and standards, Key Performance Indicators, Investment, Partnerships and Procurement. Every organization in the journey towards accessibility maturity should have accessibility vision and mission. For D2L, the vision of the company is 'to create accessible learning experiences.' The mission explains how that goal is achieved. Policy and standards refer to the guidelines and regulations that the organization adheres to create products and meet stakeholder needs. Key Performance Indicators constitutes of actionable and quantifiable metrics of progress in accessibility related to product, marketing, sales, legal and risk management. Investment is very crucial to mature in practices such as budget allocation for accessibility. Partnerships and procurement subdimension is marked by proof points to ensure that the vendors and partner products have integrated accessibility into their policies, practices, and processes.

Culture

Organizational culture has an enormous impact on how accessibility strategy is developed, received, adopted, implemented, and managed (Implementing Accessibility: Understanding Organizational Culture, 2021). Committing to accessibility by an organization means having a culture of commitment to accessibility in an organization. We identified supporting organizational culture as a best practice during the listening tours and interviews with members of various departments. So, we decided to include Culture as a dimension in D2L's Accessibility Maturity Model.

Culture has three subdimensions: Leadership, People and Culture, and Communication. Culture indicates the values and behaviours of organizational members at all levels towards accessibility. The subdimensions deals with commitment by leadership teams, human resource teams and the integration of accessibility in how the organization communicates with its stakeholders.

Competency

Another major theme emerged from the data analysis was accessibility skills and knowledge required in various roles. We did further research and interviewed learning experts within the to develop a dimension focused on continuous learning. The dimension *Competency* focuses on learning aspects of all roles within the organization where accessibility comes into play.

The dimension Competency has three sub dimensions: Learning gap analysis, competency mapping, and learning design. Competency involves the knowledge, skills, attitudes and behaviours essential to do accessibility-related jobs in each role. For this dimension, the subdimension learning gap analysis includes the process of regular check ins on the expected and current state of learning within the organization. Competency mapping includes understanding the availability of accessibility skills within each team. Learning design is the process of methodically and systematically crafting learning experiences for all employees to support them in their jobs related to accessibility.

Customer

The customer is the central factor in each company. Sustainable success is dependent on how capable a company is in meeting the needs of the users of its products and services. Customer is the dimension that takes care of the factors essential to ensure we cater to the customer needs of an accessible learning platform.

Customer dimension constitutes four subdimensions: Implementation, Support, Feedback, and Continuous engagement. All these subdimensions ensure that customer needs are taken care of continually through regular follow-up.

Community

Fostering a strong community helps to co-create and co-design accessible experiences. Community breeds collaboration and growth (Forbes, 2022). We decided to include Community as a dimension in the model inculcating some of the best practices the company is already following. We already have a community team with strong accessibility allyship and monthly meetings knowns as Accessibility Interest group where like-minded stakeholders converge and discuss accessibility-related topics.

Community has three sub dimensions; Events, Conversations, and Knowledge sharing. Each of these subdimensions focuses on D2L's core strength in community building through various channels.

AMM Maturity Stages

Maturity stages can be derived through cumulative measurement of accessibility performance to infer the current and potential next stage of accessibility maturity. W3C model suggests the following maturity levels, which we plan to finetune in further AMM iterations:

- 0 = no action, no plan;
- 1 = planning, no action;
- 2 = some action, needs improvement;
- 3 = repeated, consistent action.

Conclusion

A digital technology organization's AMM must have a holistic focus beyond its digital products. Our research provided insights about the benefits of adopting an organization-wide approach. Starting with primary research with product teams, we are iterating towards and organizational AMM as a tool to integrate accessibility into the product development, competency building, business strategies, organizational culture, and overall organizational functioning in a consistent and measurable manner. The rolling out and iterative improvement of this model as it is implemented in the actual work environment is a change management initiative. Organizational transformation in accessibility practice is like any other business transformation involving adoption of new behaviors and skills (Barrel "Integrating a Sustainable

Accessibility Program”). Our goal is to build an accessibility governance system that integrates accessibility into policies and key business processes in a consistent, repeatable and measurable fashion.

Works Cited

"Accessibility evolution model" Microsoft. https://blogs.microsoft.com/accessibility/accessibility-evolution-model/. Accessed August 29, 2023.

"Accessibility maturity model" Business Disability Forum. https://businessdisabilityforum.org.uk/knowledge-hub/resources/tech-taskforce-accessibility-maturity-model/. Accessed April 24, 2023

"Accessibility program fundamentals: Choosing the right accessibility metrics". Deque. https://www.deque.com/blog/accessibility-program-fundamentals-choosing-the-right-accessibility-metrics/ Accessed October 1, 2023

Auer, Nadine, et al. "How to measure the accessibility maturity of organizations – A survey on accessibility maturity models for higher education." *Frontiers in Computer Science* 5: 1134320. 2023.

"Accessibility maturity model". Business Disability Forum. https://businessdisabilityforum.org.uk/knowledge-hub/resources/tech-taskforce-accessibility-maturity-model/. Accessed October 5, 2023

"Accessibility strategy". TPGi. https://www.tpgi.com/accessibility-solutions/accessibility-strategy/ Accessed September 30, 2023

Barrell, Dylan. "Integrating a sustainable accessibility program: A digital transformation imperative." *Forbes Technology Council Post*, https://www.forbes.com/sites/forbestechcouncil/2020/08/24/integrating-a-sustainable-accessibility-program-a-digital-transformation-imperative/?sh=2f03321616d0. Accessed August 29, 2023.

McClimon, Timothy, J. "Doing well and doing good." https://www.forbes.com/sites/timothyjmcclimon/2020/04/20/doing-well-and-doing-good/?sh=6f0037023da0. Accessed December 5, 2023

"Digital accessibility maturity model". Level Access. https://www.levelaccess.com/blog/the-digital-accessibility-maturity-model-introduction-to-damm/ Accessed August 29, 2023.

"Digital accessibility maturity score card". Hassell Inclusion. https://www.hassellinclusion.com/accessibility-maturity-scorecard/ Accessed February 10, 2023

Donovan, Rich. "Annual report: The global economics of disability". *Return on Disability* (2020): 9.

"ISO/IEC 30071-1:2019 Information technology — Development of user interface accessibility — Part 1: Code of practice for creating accessible ICT products and services". https://www.iso.org/standard/70913.html#:~:text=This%20document%20gives%20guidelines%20for,to%20all%20types%20of%20organizations. Accessed April 11, 2023

"Policy driven adoption for accessibility (PDAA) FAQ for companies and vendors". National Association of State Chief Information Officers. https://mn.gov/mnit/about-mnit/accessibility/pdaa-faq-companies-and-vendors.jsp Accessed April 12, 2023

Priya, Arya. "Case study methodology of qualitative research: Key attributes and navigating the conundrums in its application." *Sociological Bulletin* 70.1 (2021): 94-110.

Quintal, Clara and Macías, José A. "Measuring and improving the quality of development processes based on usability and accessibility." https://dblp.org/db/journals/uais/uais20.html#QuintalM21 Accessed February 20, 2023

"Strategic accessibility companionship". https://www.nomensa.com/aa/. Accessed January 10, 2023.

"The importance of community as a business owner". Forbes Business Council Post. https://www.forbes.com/sites/forbesbusinesscouncil/2022/05/03/the-importance-of-community-as-a-business-owner/?sh=10359b2a6481 Accessed October 3, 2023

"W3C accessibility maturity model". WCAG Group Draft Note. https://www.w3.org/TR/maturity-model/. Accessed October 30, 2022

"W3C accessibility maturity model". World Wide Web Consortium. https://www.w3.org/TR/maturity-model/#About-maturity-model. Accessed August 29, 2023.

"W3C accessibility standards overview". World Wide Web Consortium https://www.w3.org/WAI/standards-guidelines/. Accessed August 28, 2023

THE JOURNAL ON
TECHNOLOGY AND
PERSONS WITH
DISABILITIES

Transforming Language Access for Deaf Patients in Healthcare

Roshan Mathew, Wendy A. Dannels
Rochester Institute of Technology, Rochester, New York, USA
rm1299@rit.edu, w.dannels@rit.edu

Abstract

Deaf patients often depend on sign language interpreters and real-time captioners for communication during healthcare appointments. While they prefer in-person services, the scarcity of local qualified professionals and the faster turnaround times of remote options like Video Remote Interpreting (VRI) and remote captioning often lead hospitals and clinics to choose these alternatives. These remote methods often cause visibility and cognitive challenges for deaf patients, as they are forced to split their focus between the VRI or captioning screen and the healthcare provider, hindering their comprehension of medical information. This study proposes the use of augmented reality (AR) smart glasses as an alternative to traditional VRI and remote captioning methods. Certified healthcare interpreters and captioners are key stakeholders in improving communication for deaf patients in medical settings. Therefore, this study explores the perspectives of 25 interpreters and 22 captioners regarding adopting an AR smart glasses application to facilitate language accessibility in healthcare contexts. The findings indicate that interpreters prefer in-person services but recognize the advantages of using a smart glasses application when in-person services are not feasible. In contrast, captioners show a strong inclination toward the smart glasses application over traditional captioning techniques.

Keywords

Accessibility design and evaluation methods, Deaf or hard of hearing, healthcare interpreting and captioning, augmented reality (AR)

Introduction

According to the latest national health interview survey, approximately 15.5% (40 million) of individuals 18 years or older in the United States indicated experiencing various degrees of hearing loss (State of Hearing Loss in America). People with hearing loss (in this study, we will refer to them as deaf) utilize various communication methods and supplementary aids, such as sign language interpreting, real-time captioning, and assistive listening devices. However, research shows that despite increased healthcare utilization among the deaf community, the current healthcare system falls short of addressing their communication needs (McKee et al.). Sign language interpreters and captioners play a crucial role in bridging this gap during clinical consultations. While deaf patients often prefer to have interpreters and captioners in person, hospitals and clinics often resort to providing remote services such as Video Remote Interpreting (VRI) or remote captioning services (James et al.). Previous research has investigated the use of VRI in healthcare for deaf patients and the preferences of healthcare providers and deaf patients for VRI versus in-person interpreting (Kushalnagar et al.; Yabe). In critical care settings, healthcare personnel favored in-person interpreters, with no distinct preference in non-critical settings. On the other hand, deaf patients preferred in-person interpreting in both critical and non-critical settings (Yabe). To the best of our efforts, no existing studies were found addressing the preferences of healthcare providers and deaf patients regarding captioning services in healthcare. However, deaf patients continue to experience several visibility and cognitive challenges with these remote services (see fig. 1), such as maintaining eye contact with medical professionals, understanding their body language and facial expressions, and scheduling access services during emergencies (Sheppard; Kushalnagar et al.; Olwal et al.). These challenges stem from the need for deaf patients to split their attention between interpreters, captions, caregivers, and any demonstrations, potentially impacting their ability to fully comprehend health-related information during medical consultations.

Visual dispersion for a deaf patient with an onsite interpreter

Visual dispersion for a deaf patient with an onsite/remote captioner

Fig. 1. Scenarios demonstrating deaf patients' challenges with full visual access to information.

One potential solution to these challenges involves using augmented reality (AR) smart glasses (henceforth, smart glasses) as assistive technology (see fig. 2). Our research team developed a prototype called Access on Demand (AoD) (see fig. 3) that projects sign language interpretation and real-time captions onto these smart glasses (Access on Demand). We proposed that smart glasses can be a viable alternative to conventional VRI and remote captioning services within medical centers and clinics.

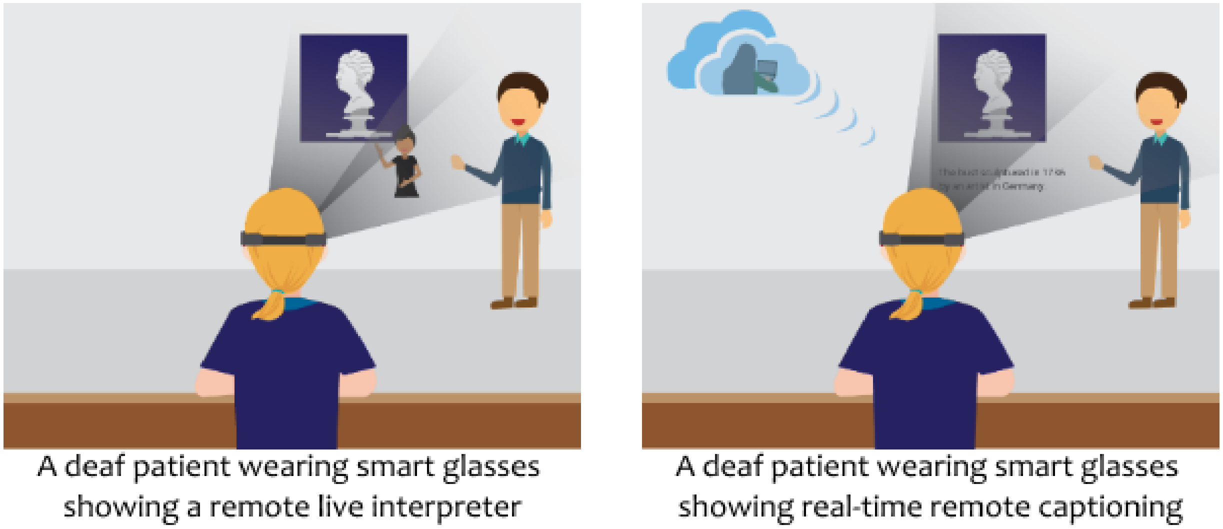


Fig. 2. Scenario demonstrating potential improvements in visual access for deaf patients using augmented reality (AR) smart glasses.

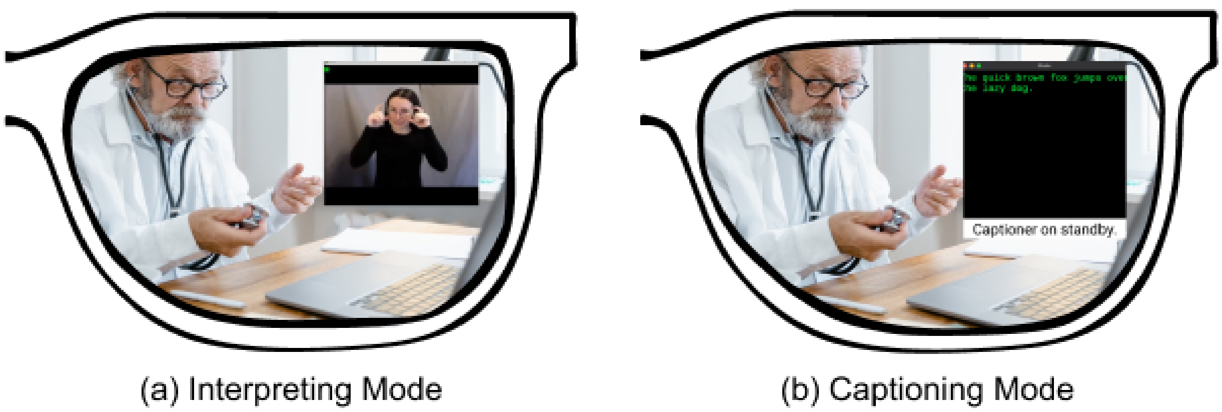


Fig. 3. Interpreting (a) and captioning (b) modes on the smart glasses.

To explore this, we conducted a study comparing the perspectives of deaf patients, healthcare interpreters, and captioners on using smart glasses to improve communication in medical settings (Mathew et al.). The study found that both interpreters and deaf patients who use sign language preferred on-site services over remote ones. Additionally, deaf patients who used captions and captioners preferred using smart glasses over traditional captioning methods. In this paper, we delve deeper into the views of interpreters and captioners regarding the advantages and challenges of incorporating smart glasses. The findings will contribute to the broader understanding of integrating AR technology as an assistive tool in clinical settings.

Discussion

Methodology

We conducted an initial literature review to investigate the experiences of access service providers working with deaf patients in healthcare settings. We then surveyed sign language interpreters and captioners to gather insights into their experiences and perspectives regarding using smart glasses to enhance these services. The study was approved by the Institutional Review Board at the authors' institution. To participate, interpreters and captioners needed to be 18 years or older, certified, and have healthcare interpreting or captioning experience. The survey was available from May to December 2022, and we recruited participants from known access service providers.

We received 25 valid responses from interpreters and 22 from captioners. To analyze participant feedback, we employed a qualitative approach, specifically, an inductive and semantic method. We condensed open-text responses by creating codes, summarizing the advantages and disadvantages of using smart glasses for access services. Then, we identified recurring themes and compared them across participant groups. These themes were further organized into broader categories, yielding insights into the challenges, factors facilitating adoption, and participant preferences concerning the use of smart glasses.

Interpreters' Perspectives on Using Smart Glasses

Although interpreters generally preferred to be physically present, they acknowledged the advantages of using smart glasses for deaf patients during appointments. One of the most significant benefits, as highlighted by interpreters, is that smart glasses enable deaf patients to maintain eye contact with clinical providers while having interpreters in their direct line of sight. Interpreter IP24 stated, "*The client's view of the interpreter would move with the client's movements. The client would be in control of the access, not the medical provider.*" Additionally, smart glasses facilitate a more personalized interaction, as explained by participant IP25, "*The providers no longer try to speak to the interpreter but have a more personalized approach with the patient*." Moreover, interpreter IP16 added, "*[Using smart glasses will allow for a] better understanding of demonstrations while looking at [the] source material [as they do] not have to look back and forth between interpreter and demonstration*."

Interpreters also saw the potential for smart glasses to significantly reduce situations where deaf patients struggle to arrange for an in-person interpreter during their medical appointments. Interpreter IP1 pointed out, "*Remote interpreting allows a greater pool of interpreters to be reached to fill the assignment, increasing the likelihood that interpreting services will be provided, even for last-minute requests*." This is especially crucial for deaf patients residing in rural areas, where qualified interpreters may be scarce and obtaining their services on short notice could be challenging. Another challenge with offering interpreting services through smart glasses stems from the lack of live context. Interpreter IP06 pointed out that in a remote setting, "*the interpreter has a limited view - interpreters can't see what's happening in the room or what's going on, which impacts communication.*" Drawing from their experience in providing VRI services, interpreters found the lack of access to the surrounding environment to be a substantial challenge, as emphasized by interpreter IP24, "*One of the biggest challenges with VRI is the interpreter's access to visual and audio cues.*" Therefore, interpreters were concerned about whether they would have access to the necessary spatial information for precise interpretation, "*Does the interpreter that is in the smart glasses have the ability to see the* *doctor, images on a screen, and the deaf patient? If not, this would be a barrier* [IP14]." Interpreters were also concerned whether the smart glasses would provide a clear and customizable view for deaf patients to see the interpreters and understand complex information comfortably. An interpreter, IP06, asked, "*Is the screen adjustable? Can you adjust the screen size or [video] quality and change the interpreter if needed? Can this be used lying down?*"

Interpreters were also concerned about potential technical glitches that could impact the use of smart glasses for interpreting services, which corroborates findings from prior studies (Kushalnagar et al.). They noted that healthcare facilities often have unreliable Wi-Fi, which could cause video streaming problems and communication delays. Furthermore, there were concerns about the reliability of smart glasses and whether the hospital personnel would have the necessary expertise to address technical problems. Interpreter IP13 emphasized these concerns, asking, "*Providers don't always know how the VRI systems work, so who could support troubleshooting with glasses?*"

Captioners' Perspectives on Using Smart Glasses

Similar to the feedback received from interpreters, captioners also believed that the use of smart glasses would enable deaf patients to "*focus more on visual aids being presented and face the provider at all times* [CP04]," which offered them the advantage of "*view[ing] chart/figures/models and still keep real-time captioning in [their] line of sight* [CP21]." Besides, unlike traditional captioning methods, where deaf patients must carry a handheld tablet or a similar device to view the captions, using smart glasses would offer a "*hands-free [experience], [as] deaf patients do not need to carry anything, nor have [to] separate their attention from the provider."* According to a captioner, CP11, all these benefits "*will greatly reduce the communication time between deaf people and health care providers.*" For the captioners, the use of smart glasses would afford them the benefit of directional audio, which could be crucial in many patient-doctor interactions. A captioner, CP03, stated, "*I did a small amount of captioning with Google Glass when it first came out, and while the technology wasn't yet sufficient at that time, I could see the tremendous potential of head-mounted captioning displays and eagerly await the maturity of the technology. I think both directional audio for the captioner, plus increased mobility and less obtrusive captions for the caption user, are huge advantages that I would love to see realized.*"

One of the primary concerns expressed by captioners regarding the use of smart glasses is the potential limitation in providing deaf patients access to the full transcript of the interaction. This limitation arises from the relatively smaller caption container within the AR projection, which cannot display as much content as larger screens on smartphones and tablets. Captioner CP06 noted, "*I would worry about how much text could be displayed at any one time on the glasses and if it would scroll past too fast. I also wonder about read-back capabilities, say, if the consumer missed something or wanted to go back and read something again and only had the glasses on*." Therefore, captioners expressed a desire for deaf patients to have "*the ability to change the font size, color, and type, as well as the caption position within the viewing area* [CP02]." This customization would enable individuals to adjust the captions to suit their personal preferences.

Captioners were also apprehensive about potential technical challenges affecting the delivery of captioning services through smart glasses, particularly issues related to Wi-Fi and hardware. They were worried about Wi-Fi problems causing connectivity issues and delays in displaying captions. Additionally, like interpreters, many captioners noted a lack of expertise among healthcare facility staff in troubleshooting hardware-related technical problems.

Captioner CP8 shared, "*I would also be concerned about medical staff knowing how to use the glasses and keeping them charged and ready to be used. I'm a captioner but also an ASL interpreter, and this is a frequent issue that I've seen with VRI equipment and hospital staff*."

Overall Preferences

Most interpreters preferred in-person services over remote methods and did not express a strong preference between VRI and smart glasses if they had to deliver their services remotely. According to interpreter IP11, this preference for onsite services is founded on the fact that "*ASL is a 3D language that is best provided by someone qualified in person*." Interpreter IP15 expanded on this point, stating, "*I think it would be difficult for an interpreter not on site to provide the in-depth interpretation required of many health care interactions. An offsite interpreter will have difficulty in managing pace; clarification; viewing necessary imaging, reports, evaluations, and interactions with the provider requiring visual cues.*" Interpreter IP02 shared their dilemma, "*I'm going back and forth between smart glasses and VRI. No current infrastructure. In both cases, the interpreter is not "present" to be looking at the X-rays, MRI scans, or other personal information, so it seems it would feel less invasive. The biggest plus to having an in-person interpreter is [that] you can see each other, you can hear the hearing person well - even when they are not visible, and the Deaf person and the interpreter can bond, so the Deaf person doesn't feel alone when trying to talk about scary medical things - it's hard to provide that bit of culture in VRI - but it can be done."* Nevertheless, some interpreters saw smart glasses as a more suitable option in specific situations, especially when deaf patients sought greater privacy or when VRI equipment wasn't feasible. Interpreter IP16 acknowledged the appeal of smart glasses for maintaining privacy, "*I can see the appeal for privacy purposes if the deaf adult would prefer not to use a local interpreter*." Another interpreter, IP23, thought "*smart* *glasses offer more parity with hearing patients in terms of privacy*" because deaf patients may feel more comfortable not having a third person in the room while discussing sensitive health information. Interpreter IP25 shared that most deaf patients also almost always prefer to have an onsite interpreter and summed it up by stating, "*[The] feedback from the Deaf community is mostly for onsite. However, a growing number are accepting VRI services. Smart glasses might be a game changer for those that do not prefer an onsite interpreter, but others may resent the headset setup*."

On the other hand, captioners have shown a preference for utilizing smart glasses to deliver their services to deaf patients, as opposed to relying on traditional captioning methods. One captioner, CP01, said, "*If they work as described, without freezing and without frustration experienced by the wearer, smart glasses would edge out remote captioning being read on a device.*" Another captioner, CP03, went into more detail, explaining, "*Smart glasses seem an ideal modality to caption to. Remote captioning [traditional] can work well, but audio quality can be challenging, and the caption user needs to look away from the person they're talking to every few seconds to read the screen. Onsite captioning can be used in case of audio quality issues or poor internet connectivity, but it requires travel time, has potential health consequences in the case of infectious disease consultations, can make the caption user feel awkward, and is less flexible logistically.*"

Conclusion

Healthcare interpreters and captioners play a vital role in improving communication for deaf patients in medical settings. Drawing from their extensive expertise in the varied range of access-related situations that deaf patients encounter during consultations, their input becomes invaluable. Consequently, their feedback offers crucial insights for successfully integrating smart glasses to enhance language access services for deaf patients within healthcare contexts. Interpreters emphasized their preference for in-person services over remote options. However, they also recognized the benefits of using smart glasses applications like AoD when in-person services are unavailable or impractical. On the other hand, captioners showed a clear preference for providing captions through smart glasses over conventional captioning methods.

Limitations and Future Work

This study is preliminary, and it is essential to note that we have not yet tested our prototype in a live clinical setting to evaluate the AR view of interpreters and captions during live healthcare consultations involving deaf patients. Furthermore, we did not collect feedback from healthcare professionals, such as doctors, nurse practitioners, and administrators, who are critical stakeholders in ensuring effective communication access within medical settings. In the future, we intend to conduct technical feasibility studies in live medical environments involving deaf patients and caregivers to evaluate the usability of AoD for all relevant stakeholders.

Acknowledgments

1. This material is based upon work supported by the National Science Foundation under Award No. 1811509. Any opinions, findings and conclusions, or recommendations expressed in this material are of the author(s) and do not necessarily reflect the views of the National Science Foundation.

2. This project was also partially supported by the American Recuse Plan (ARP) and Public Law 117-2.

3. The authors acknowledge Research Computing at the Rochester Institute of Technology for providing computational resources and support that have contributed to the research results reported in this publication.

Works Cited

Access on Demand. https://www.accessondemand.io/. Accessed 10 Oct. 2023.

James, Tyler G., et al. "'They're not willing to accommodate Deaf patients': Communication experiences of Deaf American Sign Language users in the emergency department." *Qualitative Health Research*, vol. 32, no. 1, Jan. 2022, pp. 48–63. *SAGE Journals*, https://doi.org/10.1177/10497323211046238.

Kushalnagar, Poorna, et al. "Video remote interpreting technology in health care: Cross-sectional study of Deaf patients' experiences." *JMIR Rehabilitation and Assistive Technologies*, vol. 6, no. 1, Mar. 2019, p. e13233. *rehab.jmir.org*, https://doi.org/10.2196/13233.

Mathew, Roshan, et al. "An augmented reality-based approach for optimization of language access services in healthcare for Deaf patients." *Universal Access in Human-Computer Interaction*, edited by Margherita Antona and Constantine Stephanidis, Springer Nature Switzerland, 2023, pp. 29–52. *Springer Link*, https://doi.org/10.1007/978-3-031-35897-5_3.

McKee, Michael, et al. "Reframing our health care system for patients with hearing loss." *Journal of Speech, Language, and Hearing Research*, vol. 65, no. 10, Oct. 2022, pp. 3633–45. *pubs.asha.org (Atypon)*, https://doi.org/10.1044/2022_JSLHR-22-00052.

Olwal, Alex, et al. "Wearable subtitles: Augmenting spoken communication with lightweight eyewear for all-day captioning." *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, Association for Computing Machinery, 2020, pp. 1108–20. *ACM Digital Library*, https://doi.org/10.1145/3379337.3415817.

Sheppard, Kate. “Deaf adults and health care: Giving voice to their stories.” *Journal of the American Association of Nurse Practitioners*, vol. 26, no. 9, Sept. 2014, p. 504. journals.lww.com, https://doi.org/10.1002/2327-6924.12087.

“State of hearing loss in America | NCOA.Org.” *National Council on Aging*, https://www.ncoa.org/adviser/hearing-aids/hearing-loss-america/. Accessed 1 Sept. 2023.

Yabe, Manako. “Healthcare providers’ and Deaf patients’ interpreting preferences for critical care and non-critical care: Video remote interpreting.” *Disability and Health Journal*, vol. 13, no. 2, Apr. 2020, p. 100870. *ScienceDirect*, https://doi.org/10.1016/j.dhjo.2019.100870.

20
24
THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

Predictors of Postsecondary Web Accessibility, 2012 to 2022

Terrill Thompson[1] Dan Comden[1], Scott Ferguson[1], Dustine Bowker[1], Sheryl Burgstahler[1], Jared Smith[2], Elizabeth Moore[3]
University of Washington[1], Utah State University[2], Applied Inference[3]
tft@uw.edu, danc@uw.edu, smf23@uw.edu, drbowker@uw.edu, sherylb@uw.edu, j.smith@usu.edu, liz@appliedinference.com

Abstract

The current study, conducted in 2022, is a follow-up to research conducted in 2012. Researchers compared data collected ten years apart in order to determine whether and how web accessibility, related policies, and the presence of “accessibility” links on home pages had changed over time. The top ten web pages for all U.S. higher institutions were collected using the Google Search API, and were evaluated for accessibility using WAVE, by WebAIM. Automated Accessibility Scores (AAS) were calculated for each institution. Additional data were collected from various sources, including data related to demographics, federal funding, endowment, specific characteristics of accessibility policies, whether institutions have “accessibility” links on their home pages, and whether institutions had resolved complaints with the U.S. Department of Education Office for Civil Rights (OCR) or been the recipient of one or more lawsuits related to digital accessibility. Some of these measures were collected in the 2012 study as well, allowing for comparison over time. The longitudinal comparison found significant increases in web accessibility scores, number of accessibility policies, and number of accessibility links on home pages across all types of institutions. Analysis of variance found that institutions with the highest average web accessibility scores in 2022 have had OCR resolutions; are categorized as Coalition of Urban Serving Universities; or have a digital accessibility policy, especially if it identifies a responsible party. Of the institutions that serve traditionally underserved populations, Women’s Colleges and Hispanic-Serving Institutions have higher web accessibility scores, while Historically Black Colleges and

The Journal on Technology and Persons with Disabilities

Predictors of Postsecondary Web Accessibility, 2012 to 2022

Terrill Thompson[1] Dan Comden[1], Scott Ferguson[1], Dustine Bowker[1], ,
Sheryl Burgstahler[1], Jared Smith[2], Elizabeth Moore[3]
University of Washington[1], Utah State University[2], Applied Inference[3]
tft@uw.edu, danc@uw.edu, smf23@uw.edu, drbowker@uw.edu,
sherylb@uw.edu, j.smith@usu.edu, liz@appliedinference.com

Abstract (continued)

Universities (HBCU) and Tribal Colleges have lower scores. Overall, institutions with the lowest average accessibility scores are private for-profit institutions, those categorized as "Exclusively graduate or professional," very small institutions, and institutions with no digital accessibility policy. Multiple regression analysis found that taken together, the strongest, most stable combination of predictors of web accessibility are size of the institution, having resolved at least one OCR complaint, having some type of web accessibility policy, being a public institution, being a private for-profit institution (negative association), having received a National Science Foundation (NSF) award, having received a National Institutes of Health (NIH) award (negative association), having at least one lawsuit related to digital accessibility, having a responsible party identified in an accessibility policy, being an HBCU or a Tribal College (negative association), and total degrees conferred (negative association).

Keywords

Web accessibility, higher education, policy, OCR, lawsuits, multiple regression analysis

Introduction

The current study, conducted in 2022, is a follow-up to research conducted in 2012. The original study reported on the current state of web accessibility at postsecondary institutions in the United States, the number of institutions that have policies related to web and/or IT accessibility, the number with “accessibility” links on their home pages, and the best predictors of web accessibility among the measures collected (Thompson et al.).

The current study repeated aspects of the original study to determine whether and how these measures had changed over ten years. It also introduced new independent variables and a more extensive assessment of web accessibility. The following research questions were explored:

1. Has the state of accessibility on postsecondary websites improved in 10 years?
2. How have institutions’ web/technology accessibility policies changed in 10 years?
3. Has the number of institutions with “accessibility” links on their home pages changed over the last 10 years?
4. What variables are most strongly associated with institutions’ automated web accessibility scores?

Methods

Subjects

Internet domain names for all U.S. postsecondary institutions were collected using the Google Custom Search JSON API, then manually checked for accuracy. Where multiple institutions shared a web domain, researchers chose one institution to represent the set (generally, the “main campus”). A few additional institutions were excluded from the study because no website could be found, their website consisted of fewer than 10 web pages, or their website was in a language other than English. After filtering, 3,386 institutions were included in the study. Of these, 2,717 institutions had a matching record in the 2012 study.

Sources and Procedures

The study collected the following data for each U.S. postsecondary institution:

- **Demographics:** The American Council for Education provided about 100 data elements.
- **Federal Funding (new in 2022):** The amount received from the National Science Foundation (NSF) in 2020 and National Institutes of Health (NIH) in 2022.
- **Endowment (new in 2022):** Endowment assets (year-end) per 12-month FTE enrollment as of 2021, retrieved from the National Center for Education Statistics.
- **Accessibility Policy:** Google's Search API was used to search each institution's website for digital accessibility policies. Policies were manually verified and categorized by type: Formal standalone, Formal incorporated, Standards or Guidelines, and General statement. **New in 2022:** Policies were rated on inclusion of several variables: effective date or milestones, issuing authority, accessibility standard, responsible parties, scope of covered technologies, requirements for third party content, enforcement, and user feedback.
- **Accessibility Links:** A custom script was developed to search all links on each institution's home page to determine if one or more links contain the word "accessibility." These links tend to be in the footer as part of a theme that is used throughout institutions' websites.
- **U.S. Department of Education Office for Civil Rights (OCR) Resolutions (new in 2022):** A Freedom of Information Act request and the OCR Recent Resolution Search website produced 482 OCR resolutions related to digital access involving 464

postsecondary education institutions since October 2013.

- **Lawsuits (new in 2022):** A search of case dockets since 2009 in three online legal databases produced 204 lawsuits about website accessibility involving 194 institutions.

Sample of Web Pages

For each institution, two sets of sample web pages were collected using the Google Search API, differing by their search terms:

1. Method 1 searched for “site:” + domain (e.g., “site:washington.edu”). This query frequently resulted in popular resources on departmental websites, not necessarily pages that represent the institution. This was the sole method used in the original study in 2012; therefore, this sample of pages was used for longitudinal comparison.
2. Method 2 prepended the name of the institution to the search terms (e.g., “University of Washington site:washington.edu”) and returned more pages that represent the institution (e.g., home page, admissions, libraries). Since this sample is more likely to reflect the institution’s central commitment to accessibility, it was used for all analyses that did not involve longitudinal comparison.

Measuring Web Accessibility

The WAVE API was used to scan all web pages and collect the following information:

- Total errors (number of accessibility errors detected)
- Total elements (number of HTML elements on the web page)
- Error density (number of errors per element)
- Total alerts (potential accessibility issues, verification would require manual inspection)

These data elements were used to calculate each institution's **Automated Accessibility Scores** using a formula that weights and combines these factors, generating a normalized score based on average web accessibility as estimated by WebAIM Million results.

Additionally, five outcome measures were collected for comparison with the 2012 web accessibility findings:

- **Document language:** % pages with lang attribute on outer element
- **Headings:** % pages with HTML headings
- **Page regions:** % pages with ARIA landmark roles
- **Alt text:** % images with alt text, across all pages
- **Form labels:** % form fields with labels, across all pages.

Analysis

Changes since 2012 were evaluated with a simple comparison of values since the dataset includes the universe of available subjects. Analysis of variance and Pearson's correlation coefficient were used to test the relationship between characteristics of the institutions and web accessibility while exploratory multiple regressions were used to develop a multivariate model of web accessibility.

Results

Research Question #1. Has the accessibility of postsecondary websites improved in 10 years?

Web accessibility scores have improved significantly since 2012 (see Table 1). Document language, headings, and page regions have nearly universal adoption in 2022. Page regions changed the most, from 2.9% to 93.4%. The lowest 2022 measures were alt text and form labels, though both increased substantially since 2012.

Table 1. Web accessibility scores on five measures in 2012 and 2022.

Web accessibility measure	2012	2022	Percent increase
Document language	37.1%	98.9%	166%
Headings	77.9%	98.3%	26%
Page regions	2.9%	93.4%	3074%
Alt Text	61.5%	87.1%	42%
Form labels	40.7%	67.8%	66%

Research Question #2. How have institutions' web/technology accessibility policies changed in 10 years?

In 2012, 274 institutions (8.4%) had web accessibility policies, most commonly "Formal, standalone" (35% of policies). These institutions tended to be public (83.2%), medium to very large (77.8%), with 4-year degrees (71.5%).

Table 2. Accessibility policies in 2012 by type of institution and type of policy.

Type of Institution	Formal, incorporated	Formal, standalone	General statement	Standards or guidelines	Total, All Policy Types
Associate's Colleges (n=1,204)	17	21	26	12	76
Doctorate-granting Universities (n=283)	18	35	11	10	74
Master's Colleges and Universities (n=623)	36	36	14	6	92
Baccalaureate Colleges (n=594)	7	1	8	5	21
Special Focus Institutions (n=518)	1	3	7	0	11

Type of Institution	Formal, incorporated	Formal, standalone	General statement	Standards or guidelines	Total, All Policy Types
Tribal Colleges (n=29)	0	0	0	0	0

In 2022, 1197 institutions (35.4%) had web accessibility policies, most commonly "General statements" (53.2% of policies). Policies were still most prevalent in public, four-year, and medium to very large schools, but substantial increases in numbers of policies were observed at private not-for-profit institutions, two-year schools, and small/very small schools.

Research Question #3. Has the number of institutions with "accessibility" links on their home pages changed over the last 10 years?

In 2012, 240 institutions (7.4%) had accessibility links on their home page. In 2022, this number had grown to 1438 (42.5%).

Table 3. Accessibility policies in 2022 by type of institution and type of policy.

Type of Institution	Formal, incorporated	Formal, standalone	General statement	Standards or guidelines	Total, All Policy Types
Associate's Colleges (n=1,140)	22	74	304	52	452
Doctorate-granting Universities (n=460)	24	152	81	27	284
Master's Colleges and Universities (n=633)	24	79	112	23	238
Baccalaureate Colleges (n=559)	6	44	87	15	152
Special Focus Institutions (n=561)	2	9	51	6	68

Type of Institution	Formal, incorporated	Formal, standalone	General statement	Standards or guidelines	Total, All Policy Types
Tribal Colleges (n=33)	0	1	2	0	3
TOTALS (n=3,386)	78	359	637	123	1197

Research Question #4. What variables are most strongly associated with institutions' automated web accessibility scores?

Figure 1 and Table 4 provide different presentations of the average accessibility score for several categories of institutions, compared against the overall average accessibility score of the entire population (6.95). All tests shown reached statistical significance using analysis of variance. Table 5 provides additional results of univariate statistical tests.

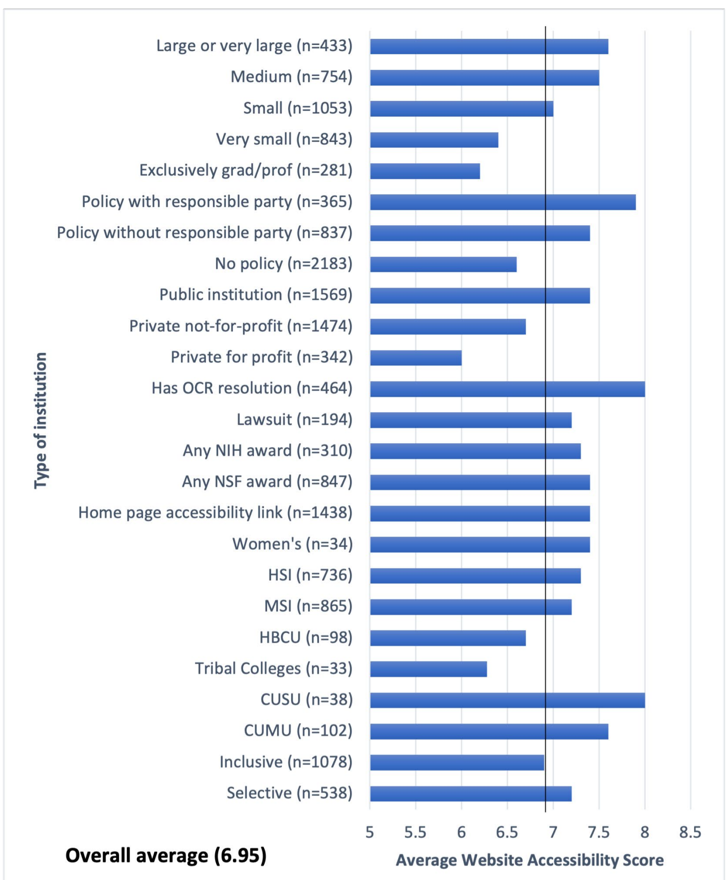


Fig. 1. Average website accessibility score by type of institution.

Table 4. Average website accessibility score (mean and standard deviation) by type of institution.

Independent Variable	N	Mean	SD
Large or very large	433	7.6	1.5
Medium	754	7.5	1.7
Small	1053	7.0	1.9
Very small	843	6.4	1.9
Exclusively grad/prof	281	6.2	1.9
Any Web Accessibility Policy	1197	7.5	1.7
No policy	2188	6.6	1.9
Standards or guidelines	123	7.4	1.6
Formal, incorporated	78	7.4	1.7
General statement	637	7.4	1.8
Formal, standalone	123	7.8	1.4
Policy has responsible party	365	7.9	1.4
Policy does not have responsible party	842	7.4	1.7
Public institution	1569	7.4	1.8
Private not-for-profit	1474	6.7	1.8
Private for profit	342	6.0	1.9
Office of Civil Rights Dispute	464	8.0	1.6
No OCR	2921	6.8	1.8
Lawsuit	194	7.2	1.6
Serial lawsuit	155	7.2	1.6
Non serial lawsuit	44	7.3	1.6
No lawsuit	3191	6.9	1.9
Any NIH award	310	7.3	1.5
No NIH award	3075	6.9	1.9
Any NSF award	847	7.4	1.7
No NSF award	2538	6.8	1.9
Home page accessibility link	1438	7.4	1.8
No home page link	1947	6.6	1.8
HBCU	98	6.7	1.7
Not HBCU	3287	7.0	1.9
Tribal Colleges	33	6.3	2.0

Independent Variable	N	Mean	SD
Not Tribal Colleges	3352	7.0	1.9
Hispanic-Serving Institution (HSI)	736	7.3	1.7
Not HSI	2647	6.9	1.9
Minority-Serving Institution (MSI)	865	7.2	1.8
Non MSI	2520	6.9	1.9
Women’s College	34	7.4	1.9
Not Women’s College	3351	7.0	1.9
Coalition of Urban-Serving Universities (CUSU)	38	8.0	1.4
Not CUSU	3347	6.9	1.9
Coalition of Urban and Metropolitan Universities (CUMU)	102	7.6	1.5
Not CUMU	3283	6.9	1.9
Inclusive	1078	6.9	1.8
Selective	538	7.2	1.8
More selective	268	7.2	1.5

Table 5. Results of univariate statistical tests by type of institution

Independent Variable	Result of univariate statistical test
Size of institution 2021	F(4,3359)=59; p<.001
Web Accessibility Policy	F(1,3383)=199; p<.001
Type of Web Accessibility Policy	F(3,1193)=5.3; p<.01
Web Policy Identifies Responsible Party	F(1,1200)=25; p<.001
Control of Institution	F(2,3382)=108; p<.001
OCR Resolution Related to Web Accessibility	F(1,3383)=162; p<.001
Lawsuit Related to Web Accessibility	F(1,3383)=4.4; p<.05
Receipt of NIH Award	F(1,3383)=13; p<.001
Receipt of NSF Award	F(1,3383)=64; p<.001
Accessibility Link on Home Page	F(1,3383)=133; p<.001
Historically Black Colleges and Universities	F(1,3383)=1; ns
Tribal Colleges	F(1,3383)=4.5; p<.05
Hispanic-Serving Institutions (HSI)	F(1,3383)=28; p<.001
Minority-serving institutions (MSI)	F(1,3383)=16; p<.001
Women’s College	F(1,3383)=2.1; ns
Coalition of Urban-Serving Universities (CUSU)	F(1,3383)=11; p<.001

Independent Variable	Result of univariate statistical test
Coalition of Urban and Metropolitan Universities (CUMU)	$F(1,3383)=13$; $p<.001$
Selectivity (ACT and % admissions)	$F(2,1881)=4$; $p<.05$
Number of positive policy characteristics	$r(3383)=.23$; $p<.001$
Total degrees conferred	$r(3383)=.12$; $p<.001$
Total Fall 2020 enrollment	$r(2330)=.14$; $p<.001$
Endowment 2021	$r(2575)=-.01$; ns
Number of first-time students	$r(2330)=.16$; $p<.001$

Multiple regression analysis (MRA) was used in a three-step process to develop a predictive model of website accessibility (the dependent variable, DV) based on available independent variables (IVs). In the first step, a wide range of candidate IVs were used in a stepwise MRA to produce an initial model. The candidate IVs were chosen based on the initial univariate analyses, exploratory hypotheses, and the quality of available data, then the statistical program evaluated the IVs based solely on the size of their correlation or partial correlation with the DV at each step of the process, until none of the remaining variables have a significant partial correlation with the DV. To test the model's stability, we used the same IVs with a backward stepping approach. In this method, all candidate IVs are entered into the model initially, and then those with non-significant partial correlations with the DV are removed until all remaining variables contribute significantly to the model. In the third step, the models are compared, common IVs are retained, and other surviving IVs may be added to the model depending on their usefulness and theoretical importance.

Forward stepping produced a model with an adjusted R^2 of .121 with 12 IVs. The backward stepping method using the same candidate IVs produced an almost identical model with an adjusted R^2 of .122. The final model, with an adjusted R^2 of .12, included all the IVs that appeared in both models, listed here in order of independent impact on the DV: size of the appeared in both models, listed here in order of independent impact on the DV: size of the institution, having resolved an OCR complaint, having a web accessibility policy, control of the institution (public and private-for-profit), having received an NIH or NSF award, indication of a responsible party in the web accessibility policy, total degrees conferred, being the recipient of at least one web accessibility-related lawsuit, being a Tribal College, and being an HBCU. The zero-order correlations between the institutions' web accessibility scores and three of the IVs were negative: Private for-profit; Tribal College, and HBCU. The partial correlations of two other IVs became negative after other IVs entered the model: being a private for-profit institution and having received an NIH award. According to the model, when all other variables are controlled for, the presence of these five characteristics tends to depress the predicted web accessibility score.

Discussion

Web accessibility scores at postsecondary institutions improved significantly from 2012 to 2022, as did the percentage of institutions with digital accessibility policies and the percentage with "accessibility" links on their home pages. These increases suggest a greater overall awareness of accessibility.

The predictive power of institution size, as well as NSF funding, may speak to the availability of resources. Well-funded institutions are more likely to have personnel with the requisite skills needed to improve website accessibility. The reversal of the positive relationship between NIH funding and web accessibility score that emerged in the MRA requires further study.

OCR resolutions are better predictors of web accessibility than lawsuits, possibly due to the approach OCR takes to resolving complaints. OCR seeks to partner with institutions and help them improve, whereas legal settlements tend to be adversarial and punitive, without similar support.

Any policy is better than no policy, and the type or formality of the policy is not a stable predictor of web accessibility. The one policy characteristic that is positively correlated with web accessibility is identifying responsible parties for policy implementation.

Further research is needed to understand the challenges faced by Tribal Colleges and Historically Black Colleges and Universities and to support these institutions in making their websites welcoming and accessible to students of all abilities.

Conclusion

The substantial improvements observed in this study in web accessibility, the number of policies, and the number of “accessibility” links on home pages suggest that postsecondary institutions became much more aware of digital accessibility between 2012 to 2022. Many variables are associated with web accessibility. Some are obvious, such as having a related policy, having resolved an OCR complaint, or having an accessibility link on the home page. Others are less obvious but still unsurprising, such as being larger institutions with larger enrollment, being a public institution, and being a recipient of NSF or NIH awards. We were surprised to find that the schools’ wealth, based on endowment data, was not predictive of web accessibility; that identifying responsible parties was the only feature of accessibility policies that is positively associated with web accessibility; and that OCR resolutions are a much stronger predictor of web accessibility then lawsuits.

Limitations

These findings, based on the institutions’ most frequently accessed web pages, may be more indicative of the institutions’ top-level, centralized commitment to accessibility than of the accessibility of the institutions' websites as a whole. Additionally, this study is limited by the use of an automated tool to measure accessibility. Future research should also include manual testing of accessibility.

Works Cited

Google, "Custom Search JSON API." *Programmable Search Engine*, 4 Apr. 2023, developers.google.com/custom-search/v1/overview.

Indiana University Center for Postsecondary Research, *Carnegie Classifications 2021 Public Data File*, ver. 10, 30 Mar. 2023, carnegieclassifications.acenet.edu/wp-content/uploads/2023/03/CCIHE2021-PublicData.xlsx.

National Center for Education Statistics, *IPEDS: Integrated Postsecondary Education Data System*, nces.ed.gov/ipeds/datacenter/selectvariables.aspx. Accessed 30 May 2023.

National Institutes of Health, *NIH Awards by Location & Organization*, U.S. Department of Health and Human Services, 22 Dec. 2022. report.nih.gov/award/index.cfm?fy=2022.

National Science Foundation, *Rankings by Total R&D Expenditure*, ncsesdata.nsf.gov/profiles/site?method=rankingBySource&ds=herd. Accessed 18 July 2022.

Thompson, Terrill, et al. "Seeking Predictors of Web Accessibility in U.S. Higher Education Institutions." *Information Technology and Disabilities E-Journal*, vol. 13, no. 1, Apr. 2013. itd.athenpro.org/volume13/number1/thompson.html.

U.S. Department of Education, "Office for Civil Rights Recent Resolution Search," *ED.gov*, ocrcas.ed.gov/ocr-search. Accessed 1 June. 2022.

WebAIM, *WAVE Web Accessibility Evaluation Tool*, wave.webaim.org. Accessed 17 July 2022.

WebAIM, *WAVE AIM (Accessibility IMpact) Testing/Scoring Methodology*, docs.google.com/document/d/10KkQzeQ1JCUEFi_vVjzMrqPMYPDhrnTaELh-oHqwER. Accessed 17 July 2022.

WebAIM, *The WebAIM Million*, 31 Mar. 2022, webaim.org/projects/million/2022.

THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

AAC Social Communication Group Photovoice

Samuel Sennott[1], Linda Akagi[1], Sam Vranizan[1], Laura Moeller[1], Lateef McLeod[2]
Portland State University[1], California Institute of Integral Studies[2]
sennott@pdx.edu, akagilinda@gmail.com, samvranizan@gmail.com, lrottiers23@gmail.com, lmcleod03@gmail.com

Abstract

The AAC Social project used a community-based participatory action research design. The study followed Wang and Burris' Photovoice methodology, as adapted by Sutton-Brown. The Universal Design Lab's AAC Design Council, led by people with the lived experience of communication disability, collaborated with AAC users, graduate students, families, and friends to co-design and implement a social communication group around music during the Covid-19 pandemic. The AAC Social included: (a) making musical playlists using Spotify music streaming service, (b) creating music using Google's Chrome Music Lab, (c) playing djembe drums, and (d) dancing. Results included validation of the importance of AAC mentorship, the co-creation of social experiences by people with the lived experience of using AAC, and the multi-modal nature of the Photovoice approach. The results point to future research expanding the use of Photovoice for people with communication disabilities and those in community with them.

Keywords

Augmentative and alternative communication, AAC mentoring, participatory action research, photovoice, social communication groups

Introduction

In its "Communication Bill of Rights," the National Joint Committee for the Communication Needs of Persons with Severe Disabilities declared, "The right to interact socially, maintain social closeness, and build relationships," as the first in a list of fifteen fundamental communication rights (Brady et al., 2016). Access to community for social interactions is something meaningful and valuable as a component of quality of life. Yet, youth with communication disabilities too often encounter barriers in accessing social opportunities. Youth with communication disabilities who use *Augmentative and Alternative Communication* (AAC) and their families report recurring barriers to relationship building with peers. Barriers include: speed of communication, lack of inclusive opportunities to socialize and low expectations from adults or parents (Ostvik et al., 2018). The purpose of our community-based participatory action research work is to explore access to community and social interaction.

All students need the space to interact socially. Locally, we know that many students and young people in our community desire to develop the social communication skills to participate in social opportunities afforded through recreation activities such as sports and music. The Covid-19 pandemic has only further limited the opportunities for inclusion. While the people who use AAC – and their needs, levels of support, and access to technology vary tremendously, McNaughton et al. (2019) observe that the success of AAC should be measured based on its ability to improve access and participation in everyday life. Access to social interaction is a key area in communication-development, particularly for school-aged children. Social interaction is important across the life course, justifying both the overall approach of this research and the need to create a community upon which we can collaboratively conduct this action research work.

AAC Mentorship

There is emerging research in the AAC community on the use of mentor programs. Adults or peer mentors may be trained in teaching others in a range of areas, and not just how to use their AAC devices. This type of relationship results in a partnership or mentorship that provides more accessibility and understanding of AAC to individuals and their families. Research has shown that this type of mentorship may best support individuals in navigating their AAC, especially if they are new to learning about their system or device (Light et al., 2007). According to Light et al. (2007), in describing “The AAC Mentor Project”, socio-relational skills as well as collaborative problem-solving skills may also be supported when creating these types of mentorship-based partnerships. Some examples of AAC mentorship include The StoryTime, Dress Up, Drama groups discussed by Batorowicz et al. (2006), which used paid and non-paid mentors to teach the children how to use their AAC devices within these social constructs and programs. Mentors were typically teenagers who used AAC devices as well. This type of set-up appeared to allow for more personal experiences and connections between the children attending the program and the (peer) mentors. Another example study by Ballin et al. (2013), used mentors to teach and train individuals with their Speech Generating Devices (SGD). When paired with mentoring, this type of modeling served as an important learning tool and resource to individuals using new SGDs. Both mentor and mentee individuals in this study were able to learn from one another and grow in their own knowledge and experience of using SGDs in their daily lives. Additionally, Grace et al. (2019) utilized an e-mentoring approach in their research, similar to one that this project took. Participants from this research study found e-mentoring to be successful and had positive experiences in learning how to use their AAC devices from one another. Other large project examples include the Bridge School’s extensive AAC Mentorship training and implementation projects and The Penn State AAC Mentor Project (Light et al., 2007) and serve as community resources. In summary, we believe AAC mentors can play a pivotal role in assisting the AAC community to help youth develop their communication competence.

Discussion

Project Origins: The AAC Design Council

The origins of this project came out of what we call in our research lab, the *AAC Design Council*. The AAC Design Council is a leadership group that exemplifies the principle of nothing about us, without us. For over two years, people with communication disabilities gathered weekly on Wednesday afternoons along with AAC researchers, students, and other interested community members. This group was forged out of a desire to create a research laboratory in AAC that had people with disabilities at the leadership level. Co-authors Linda Akagi and Sam Vranizan anchored the AAC Design Council and tirelessly showed up weekly and pushed our group to interact and go deep into important topics to the community.

Research Learning Community

In order to create an environment where people with disabilities can lead and co-design research, we needed to change the way we conceptualized assistive technology and augmentative and alternative communication research. Most saliently, we required adopting a principle that people with the lived experience of disability would be at the leadership level of the project. Additionally, we as a research group had been developing an interest in community-based participatory action research (Cornish et al., 2023; Wallerstein et al., 2018), alongside our major emphasis on single-case design methods in assistive technology.

The participant group engaging in what we called the Research Learning Community

(RLC) were made up of (a) people with communication disabilities at the leadership and participatory levels, (b) research scientists, (c) graduate students, (d) family members and friends. The RLC's purpose was to create a fluid community centered around important life activities and participatory action research. We started the RLC from "out of nothing" as we like to say and used the design thinking steps to guide our creation of a program, AAC Social, and a research agenda, which we eventually decided to use the Photovoice participatory methodology.

We used Design Thinking to guide our creation of what would eventually become the AAC Social. Design Thinking is a five-step process: (a) we empathized, and focused on the people we wanted to design for; (b) we defined the focus of our project; (c) we brainstormed ideas for what it might look like; and (d) we created and (e) tested a prototype. In each phase of this process, we used the Participation Model and the SETT framework to guide our discussion. SETT stands for student, environment, task, and tools, and is a person-centered methodology for determining what supports a person will need for a given task in a given environment. We often used Google's Jamboard tool to record our ideas (See figure 1).

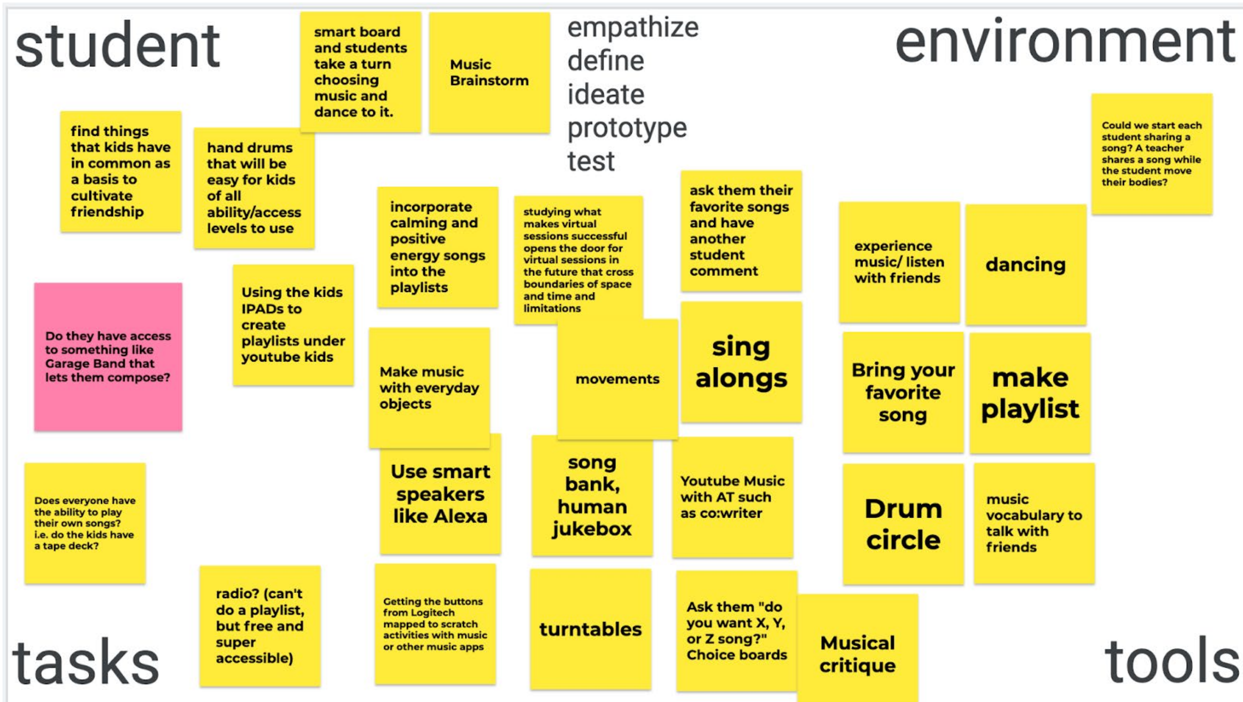


Fig. 1. One of our Google Jamboards from the design thinking process used to create the research learning community's AAC Social.

AAC Social

So out of a start from nothing, we began empathizing about what youth with communication disabilities were doing and what they may be missing during the Covid-19 pandemic. Upon completion of initial design sessions with our RLC community, we generated the idea that we could conduct a social group around the topic of music. Music is a powerful means of communication, providing a way to share emotions, intentions, and meanings. It brings people together to identify commonalities and interests so that social connections and relationships can be created. Music is foundational to many social interactions whether friends are listening to, searching for, sharing, making, or dancing to it. The after school social group format engaged students who use AAC in each of these ways and created meaningful interactions and experiences.

The AAC Social, a social communication group, focused on music as the connection. The group did three main activities, including (A) making musical playlists using Spotify music streaming service, (B) creating music using Google's Chrome Music Lab, and (C) purchasing and using djembe drums, which we shipped to participants. The first two weeks of the social group centered on sharing music through collaborating on Spotify playlists crafted around either shared emotions or musical genres. The playlist activity prompted learning through finding new music, conversation, and a digital mixtape of songs that could be shared between participants and outside the group. Next, participants created music through two online tools in Google's Chrome Music Lab, Song Maker and Kandinsky. These digital musical creations also produced a visual representation of the sounds and rhythms, and were saved and shared with others.

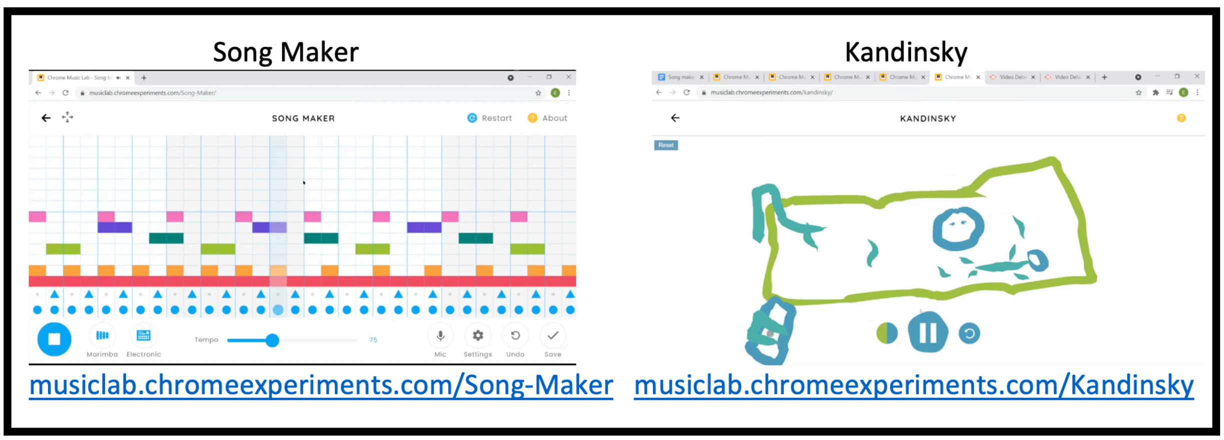


Fig. 2. Google's Chrome Music Lab is a website that makes learning music more accessible through fun, hands-on experiments. Here are two creations from our use during AAC Social using Song Maker and Kandinsky.

In the sessions, we frequently centered experience on the activity of dancing to music. Participants chose the music (and created another short playlist) to let loose and have a virtual dance party together. Finally, the social group engaged in two weeks of drumming and percussion. Participants used shakers and drums made from household objects or received djembe drums to experiment with percussion, create rhythms together, and express their emotions. In reflecting on the skills identified through The Rhythmic Arts Project (Smith et. al. 2019), we socialized, and attended to skills required to make music such as: turn taking, verbalization, motor control, memory, and self-awareness.

Photovoice

Our participant recruitment and study activities followed Wang and Burris' (1997) Photovoice methodology, as adapted by Sutton-Brown (2014) and combined with the digital consenting process as outlined and adapted from "The Elements of Informed Consent, A Toolkit". Our total number of participants in the Photovoice presentation was 14. However, we included a number of additional participants, AAC Mentors and graduate students who participated in parts of the program. You can see an example of one of the contributions here in Figure 3.

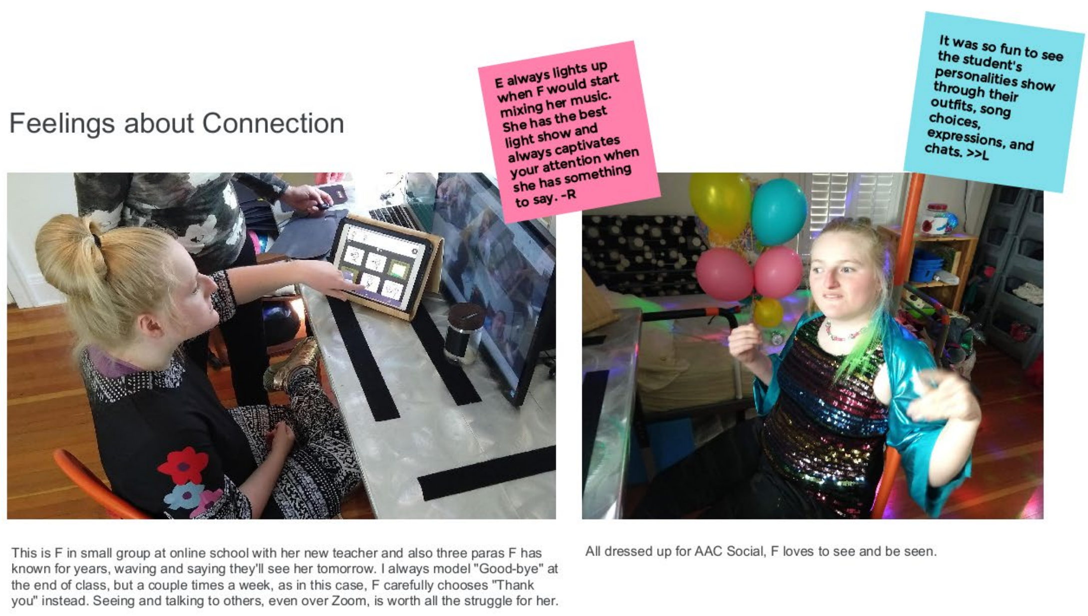


Figure 3. Feelings About Connection Photovoice Presentation Entry.

Wang and Burris outline key steps for a Photovoice project, including:

1. Select and recruit a target audience of policy makers or community leaders.
2. Recruit a group of Photovoice participants.
3. Introduce the Photovoice methodology to participants and facilitate a group discussion.
4. Obtain informed consent.
5. Pose an initial theme for taking pictures.
6. Provide time for participants to take pictures.
7. Meet to discuss photographs.
8. Plan with participants a format to share photographs and stories with policy makers or community leaders.

Overall, our goal was to test out the participatory methodology Photovoice as a way to give opportunity for members of our “pop up” RLC and AAC Social group to voice their thoughts and to have an archive of this dialogue. We believe we conducted our study in line with the three main goals of Photovoice, including: (a) to enable people to record and reflect their community’s strengths and concerns, (b) to promote critical dialogue and knowledge about important community issues through large and small group discussion of photographs, and (c) to reach policymakers (Wang & Burris, 1997).

Research Question

With a community-based participatory action research approach, the AAC RLC invited participation in a special opportunity based around using a Photovoice methodology to engage the community. Specifically, we as a community asked, “How can we reflect on our students with communication disabilities social interactions and participation in the AAC Social music program as described through a Photovoice based research process?

Results

As a way to summarize and interpret the overall AAC Social experience, we as a research learning community presented 14 collages of images and text that made up our Photovoice presentation. Each member presented around these media and they were the result of ongoing dialogue throughout the project. Here in Figure 3 is an example of the mixed media. We as a research learning community gained deep insight from this Photovoice and the overall experience.

In summary, we identified a deep community need for a social group of this sort and that it appears to be a terrific forum for AAC mentorship work and AAC research and development. We believe this high-level finding is scientifically important, because our community included not only individuals with the lived experience of communication disability, but a whole host of friends, family, and mentors around them dynamically. Having that group at the table for both direct service of the community and for research purposes has the potential to create a rich environment, filled with meaningful feedback and actionable design work.

As a community, we adapted and learned about community-based participatory action research (Wallerstein et al., 2018) and specifically the approach called Photovoice (Wang & Burris, 1997; Catalani & Minkler, 2010). An important insight around this approach is that it helped our project be engaging, fun, and easier on everyone because photographs help set context and make the dialogue more personally meaningful. Additionally, we gained a number of specific insights to share around AAC mentorship, the importance of social experiences, and specifically about the multi-modal nature of the Photovoice approach.

AAC Mentoring

Our results contribute research in the AAC community on the use of mentor programs.

Adults or peer mentors may be trained in teaching others how to use their AAC devices. Research has shown that this type of mentorship may best support individuals in navigating their AAC, especially if they are new to learning about their system or device (Light et al., 2007). One of the parents shared, “It is so powerful for the children and families to see experienced users model what they can do with AAC!” Another parent confirmed this sentiment, “We enjoyed having mentors in our weekly AAC social and appreciated their input into the conversations.” A very engaging moment occurred during our music creation sessions, when one of the guests AAC mentors who is a music producer and rapper visited and surprised the AAC Social with a performance of his song. To add to the excitement, over the course of the session, our resident DJ was able to remix the song into a dance focused mix and we closed the session with a dance party to the remixed song. Additionally, a pioneering singer and musician who uses AAC contributed as a mentor, culminating in presenting a short concert during our final celebration. This role modeling goes beyond just demonstrating use of multi-modal AAC and demonstrates being a musician and a socially valuable activity that can be carried on throughout one’s life.

One promising component of the RLC included mentors extending the work out into the community and this teacher reflection describes the impact that the mentors were able to make:

I teach a medically fragile kindergarten classroom with students that have complex communication needs. I incorporate what the Portland State AAC Social Masters students do on Monday for my Wednesday Social group. My favorite part of this whole process is the AAC mentors and them taking the time out of their day to come to my social group on Wednesdays. I have a student that has the same communication system as Sam V, AAC Mentor, and it really helped this students’ helpers and this student by seeing someone using the device, with success and having to have things added to his talker when he couldn’t find the word for ‘pig’. The quote on Patrick's picture was said in our last social group and it was said at the perfect time and really helped this student and his helpers to see that having a 'voice' matters. - Special Education Teacher

Overall, the AAC mentors and participants reported a very positive experience. This quote exemplifies that, "I enjoyed visiting the classroom and meeting the young students. I am excited for them to build their skills and A.A.C. toolkits. –Patrick"

Fig. 4. The Positive vibes of Mentors Photovoice Presentation Entry.

AAC Social as a Context for Communication and Shared Experience

An important result was the overall success of building the social group around a motivating activity, music. One AAC mentor participant shared about the observed interactions during the AAC social:

> *During my interviews and my observance of the AAC RLC, I noticed that participants in these programs exhibited a pleasant level of camaraderie with each other that was welcome and inviting to everyone that was involved. The AAC RLC online mentoring sessions were*

usually very festive and organized around the listening and the making of music.

In reflection, we recognize that in our design work, our important conversations with AAC mentor and leader Chris Klein helped us understand that gathering around socially valued interests helps create interest and motivation for socializing, setting the context for mentoring and interaction. A mentor shared, *"These activities kept all participants engaged and gave the chance for the students who use AAC a chance to use their devices when expressing their musical choices. There were many smiles and a jubilant spirit as participants had the opportunity to share their musical creativity."*

In summary, a welcoming and accessible environment for socializing was co-constructed by people with disabilities, families, graduate students, and researchers. This important pilot data and exploration of a community system for sharing project and study results using Photovoice prepares our community to present our needs to policymakers.

Photovoice as a Promising Multi-Modal Community Tool

One important result we observed is that our adapted Photovoice methodology appeared to be useful and the photo elicitation is particularly promising given that the community we are centered around includes people who benefit from multi-modal communication tools. While the parents of the participants with communication disabilities presented the photos and engaged in much of the dialogue for this inquiry, all participants were included in the process. One particularly salient reflection demonstrates the overall impression of using Photovoice:

> *As we showed pictures, we, as a group, were able to interact socially and connect with one another through our shared experiences. The desire to connect through photos is natural because we want to share our joys and our little or big moments with others, thus I think through Photovoice we are able to create a forum to reach out to others and build those*

connections. Making connections and friendships can be a struggle especially when you add on any communication difficulties, therefore by creating this safe space we cultivated a place to have these little moments. – Participant reflection.

In summary, we co-designed an interactive social group around music called AAC Social and tried out the Photovoice methodology for the first time.

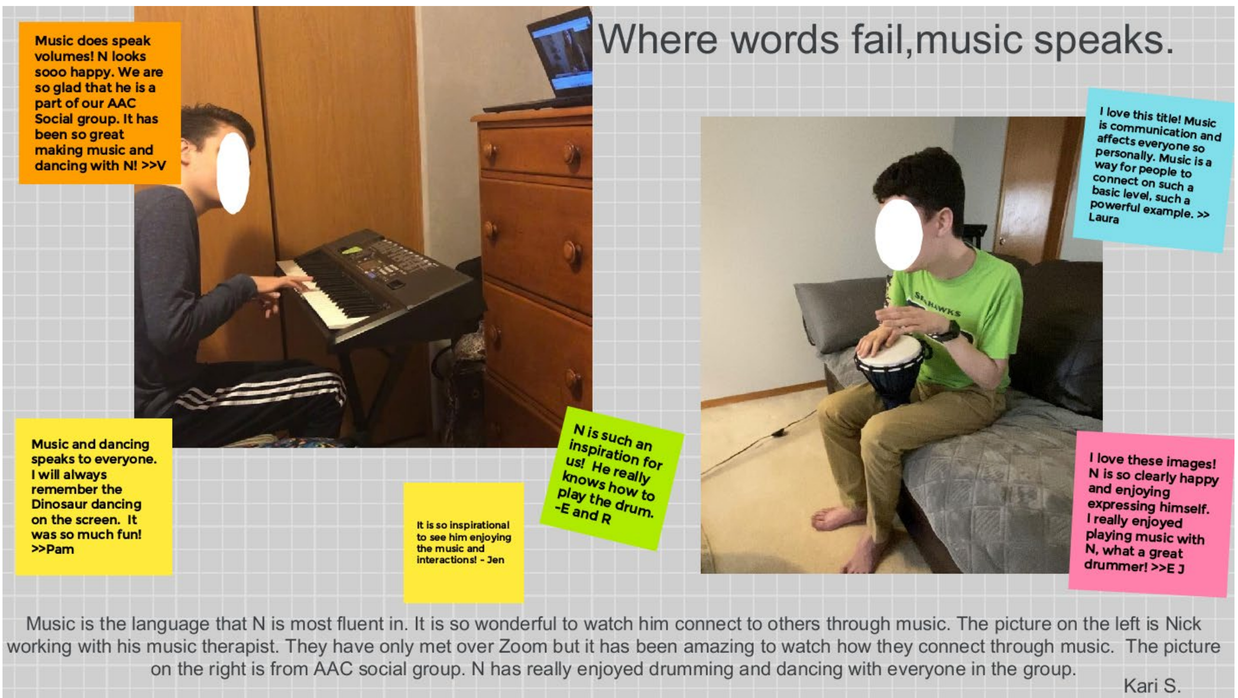


Fig. 5. Where words fail, music speaks Presentation Entry.

Conclusion

In this community based participatory action research, we engaged with important factors that affect a student's relationships, such as shared experiences between students who use AAC and fellow students, environmental adaptation and support provided by staff and fellow students, staff's efforts in building goodwill for students using AAC, and the confidence of students using AAC (Ostvik et al., 2018). Our co-creation of the AAC Social allowed us to create an accessible social group with AAC mentors at the leadership level of the project and to test out the Photovoice methodology. While this is only a first exploratory study, the approach demonstrated a number of key participatory action research principles important to our community and gave us experience as a community.

We confirmed there is a substantial need for future research in the field of inclusion of family and friends in AAC family systems framework with family-centered AAC services for families integrating or using AAC (Mandak et al. 2017). The Photovoice presentation creates a tangible archive of that need, which we can now use as a community to present to policy makers, including seeking sustainable support for this work. Our study and current literature suggest that future research needs to include "more of the voices and perspectives of individuals from culturally and linguistically diverse backgrounds who use AAC devices, and emphasize their differences as well as similarities (Kulkarni, Saili S, & Parmar, Jessica. 2017).

Future research can help make the Photovoice methodology adapted to even more include people with disabilities in the dialogue. Research in the future needs to consider individual student autonomy, choice, the perspectives of students who use AAC (Kulkarni, et al. 2017). In addition, we believe there is need to further understand what factors contribute to the rejection and abandonment of AAC prosthetics by the families whose children use these devices and the role that mentors may play in changing those outcomes (Moorcroft, 2020).

Overall, this first AAC Photovoice study by our community demonstrates a specific methodology and approach to participatory research that is designed to democratize important community dialogues. To fulfill the last step in the process, we now go forward to advocate for support through shared grant proposals and additional community engagement.

Works Cited

Batorowicz, B., Mcdougall, S., & Shepherd, T.A. “AAC and community partnerships: The participation path to community inclusion.” *Augmentative and Alternative Communication*, vol. 22, no. 3, 2006, pp. 178–195.

Brady, N. C., et al. “The Communication Bill of Rights. Communication services and supports for individuals with severe disabilities: Guidance for assessment and intervention.” *American Journal on Intellectual and Developmental Disabilitie*s, vol. 121, no. 2, 2016, pp. 121-138.

Catalani, C., & Minkler, M. “Photovoice: A review of the literature in health and public health.” *Health Education & Behavior*, vol. 37, no. 3, 2010, pp. 424–451.

Cornish, F., Breton, N., Moreno-Tabarez, U., Delgado, J., Rua, M., de-Graft Aikins, A., & Hodgetts, D. “Participatory action research.” *Nature Reviews Methods Primers*, vol. 3, no. 1, 2023.

Grace, E., Raghavendra, P., McMillan, J.M., & Gunson, J.S. “Exploring participation experiences of youth who use AAC in social media settings: Impact of an e-mentoring intervention.” *Augmentative and Alternative Communication,* vol. 35, no. 2, 2019, pp. 132-141.

Kulkarni, S.S., & Parmar, J. “Culturally and linguistically diverse student and family perspectives of AAC.” *Augmentative and Alternative Communication,* vol. 33, no. 3, 2017, pp. 170–180.

Light, Janice, et al. “The AAC Mentor Project: Web-based instruction in sociorelational skills and collaborative problem solving for adults who use augmentative and alternative communication.” Augmentative and Alternative Communication, vol. 23, no. 1, 2007, pp. 56–75.

Mandak, K., O’Neill, T., Light, J., & Fosco., G.M. “Bridging the gap from values to actions: A family systems framework for family-centered AAC services.” *Augmentative and Alternative Communication*, vol. 33, no. 1, 2017, pp. 32-41.

McNaughton, D., Light, J., Beukelman, D. R., Klein, C., Nieder, D., & Nazareth, G. “Building capacity in AAC: A person-centered approach to supporting participation by people with complex communication needs.” *Augmentative and Alternative Communication*, vol. 35, no. 1, 2019, pp. 56-68.

Østvick, J., Ytterhus, B., & Balandin, S. “‘So, how does one define a friendship?’: Identifying friendship among students using AAC in inclusive education settings.” *European Journal of Special Needs Education*, vol. 33, no. 3, 2018, pp. 334-348.

Smith, B. A., Tuduri, E., Mostovoy, E., Pannell, D., & Landon, C. “Evaluation of the rhythmic arts project, a multi-modal rhythm-based perception and action intervention, in a school-based setting in children with Autism Spectrum Disorders.” *Journal of the American Academy of Special Education Professionals*, 2019, pp. 171–179.

Sutton-Brown, C. A. (2014). “Photovoice: A methodological guide.” *Photography and Culture*, vol. 7, no. 2, 2014, pp. 169–185.

Wallerstein, N., Duran, B., Oetzel, J., Minkler, M., “Community-based participatory research for health: Advancing social and health equity.” John Wiley & Sons, 2017.

Wang, C., & Burris, M. A. (1997). Photovoice: Concept, methodology, and use for participatory Needs Assessment. *Health Education & Behavior*, vol. 24, no. 3, 1997, pp. 369–387.