Exploring the User Experience of an AI-based Smartphone Navigation Assistant for People with Visual Impairments

Bineeth Kuriakose
Department of Computer Science,
Oslo Metropolitan University
Oslo, Norway
bineethk@oslomet.no

Raju Shrestha
Department of Computer Science,
Oslo Metropolitan University
Oslo, Norway
raju.shrestha@oslomet.no

Frode Eika Sandnes
Department of Computer Science,
Oslo Metropolitan University
Oslo, Norway
frodes@oslomet.no

ABSTRACT

Advancements in artificial intelligence (AI), computer vision, and smartphone technology have spurred the development of navigation assistants aiming to improve the independent mobility experience of people with visual impairments. Still, few studies have explored whether such tools meet this intention. Therefore, this study aimed to examine the perceptions of people with visual impairments on an AI-based navigation assistant and how their perceptions may influence their use of these technologies. A study involving 13 participants was conducted with DeepNAVI, an AI-based smartphone navigation assistant for people with visual impairments in both indoor and outdoor environments. The participants reported mixed reactions towards the navigation assistant, from excitement about the promise of increased independence to a lack of trust and skepticism toward the reliability of this technology. The perceptions from the participants provide some valuable insights about AI-based smartphone navigation assistants and their practicality in real life.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in accessibility**; **Accessibility design and evaluation methods**; **Accessibility systems and tools**; **Usability testing**; **Walkthrough evaluations**.

KEYWORDS

AI, smartphone, portability, deep learning, navigation assistant, visually impaired, user evaluation

ACM Reference Format:
Bineeth Kuriakose, Raju Shrestha, and Frode Eika Sandnes. 2023. Exploring the User Experience of an AI-based Smartphone Navigation Assistant for People with Visual Impairments. In *15th Biannual Conference of the Italian SIGCHI Chapter (CHItaly 2023), September 20–22, 2023, Torino, Italy*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3605390.3605421

1 INTRODUCTION

Numerous solutions utilizing various technologies have been proposed to assist people with visual impairments in navigation [20, 33]. Over the last couple of years, artificial intelligence (AI)-based technologies have been used in navigation systems for people with visual impairments. When portability and convenience become important, AI-based smartphone navigation assistants are proposed to help people with visual impairments navigate their surroundings using their smartphones [4, 14, 30]. Such navigation assistants can be able to provide users with information about their surroundings to help them navigate to desired destinations. Moreover, they can also potentially improve the independence and mobility of people with visual impairments [5, 23, 25]. The portable nature of smartphone-based navigation assistants is complemented by their low costs and social stigma compared to other assistive technologies of similar purposes, such as white canes and smart canes [10, 25, 57]. Previous research [25, 30, 58] has reported the growing popularity of smartphone-based assistive technologies among people with visual impairments as they are easy to use and enhance user interaction. Studies also reported that people with visual impairments use a variety of strategies to adapt to mobile devices and successfully use them to perform everyday tasks and navigate independently. Examples include adjusting the device settings to make it more accessible, wearing headphones only in one ear while keeping the other ear free to sense situational cues, etc. [28].

In this paper, we report the results of a qualitative study involving people with visual impairments (hereby called *participants*) when they use an AI-based smartphone navigation assistant in natural outdoor-indoor environments. The study involved user experience tasks followed by personal interviews to gain insight into participants' experiences with the AI-based smartphone navigation assistant. The aim of this study was to assess the practicability, usability, and users' perceptions of such a smartphone-based navigation assistant.

This paper is organized as follows. Section 2 describes related works, Section 3 gives a brief description of the navigation assistant used in this study, and Section 4 details the methods that provide an overview of the tasks, participants, and data analysis. Section 5 presents the findings from the study and a discussion on Section 6. The conclusion is presented in Section 7.

2 RELATED WORKS

Various smartphone navigation assistants have been proposed to help people with visual impairments. This section briefly describes the systems and methods used to evaluate the assistant.

Sato et al. [50] proposed an indoor navigation assistant called NavCog3, which uses beacon signals to provide semantic information about various obstacles on the pathway. The system was tested on people with visual impairments in a shopping mall. Subjective

CHItaly 2023, September 20–22, 2023, Torino, Italy

ACM ISBN 979-8-4007-0806-0/23/09.
https://doi.org/10.1145/3605390.3605421

responses were solicited, the authors claim that the answers were positive, and participants found the semantic information useful. Similar to NavCog3, Nair et al. [37] proposed a mobile application that leverages beacon signals and an augmented reality (AR) framework to provide navigation instructions. The system was named ASSIST and shares similar challenges to NavCog3 regarding the need for a beacon infrastructure. The evaluation of ASSIST was conducted in a multi-story building, and the authors claim that participants were positive that the app reduced navigation errors.

Kuribayashi et al. [35] recruited 14 participants to evaluate Corridor Walker, a smartphone-based indoor walking assistant. The system uses a LiDAR sensor to construct a 2D map of the surrounding environment. The study showed that the system significantly reduced the number of error contacts and enabled participants to avoid obstacles. The use of LiDAR for navigation assistance for people with visual impairments has also been explored in other works [32]. A smartphone-based navigation assistant using a depth camera for obstacle detection has also been explored [51]. The assistant is controlled using simple voice and gesture commands. An experiment was conducted with five blindfolded participants. However, the capability of the mobile assistant is limited to locating and detecting a general obstacle only and cannot provide additional information such as obstacle type or distance. Rao et al. [47] proposed a navigation solution using Google Glass. The camera embedded in the smart glasses was used to capture images of the surroundings, which were analyzed using Azure Cognitive Services. The system was tested with students with visual impairment in a school environment. The system requires constant network connectivity to process environmental data, and using Google Glass is relatively expensive in developing countries and hence is not easily affordable. In addition, the authors reported battery drainage issues while using Google Glass.

One of the main observations from recent works reported here is that most navigation assistants have limited operational environments, indoors or outdoors. Even though these systems could bring portability, the cost involved in deploying and generalizing them in different geographical locations is challenging and not practicable. In this situation, an AI-based smartphone navigation assistant that can give information about the environment and obstacles without depending on any external data network or infrastructure becomes significant and useful. The following section describes a brief overview of the navigation assistant used in this study.

3 SYSTEM

This study employed DeepNAVI, an AI-based smartphone navigation assistant for people with visual impairment [34]. The app associated with DeepNAVI is deployed in a smartphone, and the user wears the device by placing it in a vest with its camera facing outwards (see Figure 1). From the captured video frames, various information about obstacles, such as their type, distance to them, position, motion status, and scene information, is delivered to the user through bone-conduction headphones. The DeepNAVI has trained lightweight deep learning models to capture and process data from the environment and detect obstacle information to help users navigate. Because of the use of lightweight deep learning models, the navigation assistant could offer a faster inference time than similar systems reported in the literature [47, 51]. Fast inference time is crucial in a navigation assistant while working in a real-time environment [59, 61]. An illustration of the working of the DeepNAVI is shown in Figure 2. DeepNAVI can be activated with specific voice commands, providing flexibility to users without needing to find buttons on the device or press somewhere on the screen. Unlike other similar deep learning-based navigation assistants [7, 45, 50], the DeepNAVI does not require internet connectivity, and it only requires a smartphone to work fully functional. Thus it offers unyielding portability to the users.

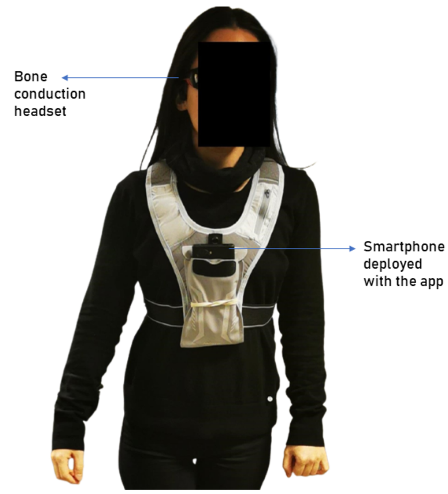

Figure 1: The DeepNAVI setup.

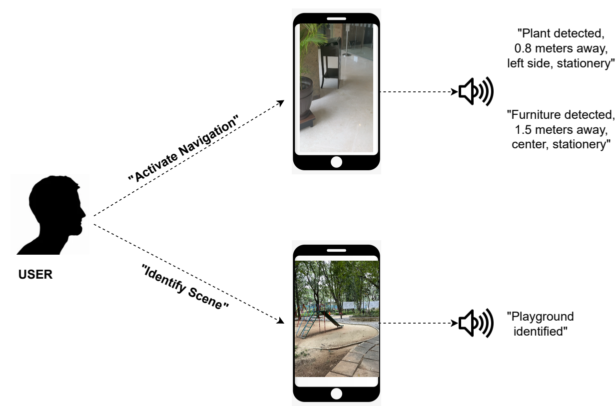

Figure 2: Working example of the DeepNAVI.

4 METHOD

User perceptions could provide more insightful information about a system if the target users themselves test it in its intended natural environments. Considering this, a qualitative study design was chosen. This user-experience study was designed with relevant tasks from the motivation of similar studies reported in the past [35, 37, 47, 51] and also based on interactions with people with visual impairments (not the participants).

Figure 3: Scene from the indoor environment.

Figure 4: Scene from the outdoor environment.

4.1 Navigation tasks and environments

Two navigation tasks were conducted to understand the user experience from the DeepNAVI. The first navigation task was intended for the indoor environment, and the second was for the outdoor environment. Both environments were natural, meaning it was not in a laboratory or simulated setup. The navigation environments were chosen to test the full functionalities of the DeepNAVI that can help the participants to have a complete evaluation. The participants were asked to walk the navigation paths in both environments using the navigation aid for each task condition to reach the respective destination points.

The *indoor environment* consisted of various physical obstacles such as chairs, cabinetry, furniture, stairs, doors, indoor plants, stairs, and trash cans on the path. The participant was asked to reach the destination using the navigation aid and pick an object (a mini coffeemaker) based on each task condition. The path was 70 meters long. Figure 3 shows the indoor environment.

The *outdoor environment* was a playground for children, which was also a traffic simulation park with signs and zebra crossing roads. The path consisted of several obstacles, such as traffic signs, plants, people, fire hydrants, bicycles, trees, benches, and waste containers. 1-2 humans were present in the location as part of the task, which made it possible to test the *human* obstacle recognition functionality of the navigation assistant. The destination point for this task was a vehicle parking area. The path was nearly 150 meters long. Figure 4 shows the outdoor environment.

Each navigation task consisted of two conditions: (1) navigation with DeepNAVI alone and (2) navigation with DeepNAVI along with a white cane. The latter was included to assess how the participants perceive a smartphone navigation assistant when another familiar mobility aid is given as support. The decision to have a white cane as a supplementary aid to DeepNAVI as one task condition was based on interactions with individuals with visual impairments (not the participants) during the study design phase. For the first task condition, the DeepNAVI alone is used, which includes a smartphone, a bone conduction headset, and a vest with a smartphone holder. For the second task condition, a white cane was used in addition. Participants were informed to bring their personal passive white canes without any built-in electronics.

The study was scheduled on different times and days for each participant based on availability. Before starting the tasks, each participant was asked about their demographics, and detailed instructions on the tasks and purpose of the study were given. A briefing session was conducted for 10-15 minutes to familiarize the participant with the navigation assistant. To minimize the learning effect while repeating the task condition in the same environment, the participant was asked to complete an indoor task condition followed by an outdoor task condition. Afterward, it continued with the remaining task condition in indoor and outdoor environments. At the end of each task condition, the participant was probed using a semi-structured interview. The whole process took almost 120 minutes for each participant. Each participant was compensated with a gift voucher for their participation. Before the actual tasks, a pilot test was conducted to validate the feasibility of the procedures.

No private or personal information was collected about the participants. The data handling plan was approved by the Norwegian Agency for Shared Services in Education and Research (Sikt)[1]. The participants gave oral consent before participating in the tasks, and an observer accompanied the participants during the tasks to ensure no harm or injury occurred.

4.2 Participants

Fourteen participants with visual impairments were recruited through several support organizations and social media groups. The observations from one of the participants were discarded as this participant decided to abort one of the tasks. Hence, the results reported herein are based on the 13 participants who completed both tasks. No participants had reduced hearing, and all had experience using white canes. The participant group comprised a balanced mix of men and women with varying smartphone experiences. Table 1 lists demographic details. English was used for communication during the tasks, and due to limited English proficiency, three participants (P6, P9, and P10) communicated using an interpreter while doing the tasks.

[1]www.sikt.no/en/data-management-plan

Table 1: Demographics of participants. Age group: (18-25, 26-35, 36-45, 46-55), Gender: (Male, Female, Others), Visual impairment level: (Moderate, Severe, Blindness), Age of onset (the age at which blindness happened), Usual navigation aids, Smartphone Experience: (Nil, Rare, Moderate, Frequent).

ID	Age group	Gender	Visual impairment level	Age of onset	Usual navigation aids	Smartphone experience
P1	26-35	M	Blindness	0	White cane, Smart cane	Frequent
P2	36-45	M	Blindness	0	White cane	Moderate
P3	26-35	F	Moderate	7	White cane, Smart cane	Frequent
P4	26-35	M	Severe	4	White cane, Smart cane	Moderate
P5	18-25	M	Blindness	0	White cane, Smart cane	Frequent
P6	36-45	F	Severe	4	White cane, Getting help from others	Rare
P7	36-45	M	Severe	10	White cane, Guide dog	Moderate
P8	26-35	M	Moderate	5	White cane, Guide dog	Frequent
P9	46-55	M	Severe	0	White cane	Rare
P10	46-55	F	Severe	0	White cane	Rare
P11	26-35	M	Severe	9	White cane	Moderate
P12	36-45	M	Moderate	16	White cane, Getting help from others	Rare
P13	26-35	M	Blindness	0	White cane, Guide dog	Frequent

4.3 Analysis

The responses from the interview were qualitatively coded using a thematic coding approach that included both inductive and deductive codes [13]. At first, a codebook was developed with an initial set of codes based on the open-ended questionnaire. Then the collected data from the survey was gathered into one document and grouped the excerpts by assigning them to the matching codes. All the data was marked using color coding to differentiate between the different comments of the participants. We took the utmost care to view the collected data objectively, thus avoiding interpreting the participants' responses to give specific meaning to what was said. After all the data had been assigned different codes, we conducted two iterations of thematic coding analysis to find themes or patterns across our data. Once we finished the first iteration, we developed two new codes, "output modality" and "portability and connectivity," and re-assigned related excerpts to that. In the second iteration, we counted how many held the same opinion about the various topics on each code. Then we extracted meaningful quotes that could be interesting to include in the results. Some quotes were paraphrased to avoid revealing the identity of the participants. Furthermore, if two or more participants have a similar opinion, they were merged to form a single quote. After the rounds of coding data, we took the codes and categories and constructed the final narrative. We employed intercoder reliability [15] by ensuring that the multiple authors came to the same conclusions.

5 FINDINGS

This section presents the findings from the interview, which was conducted after the tasks. The findings are categorized according to code labels.

5.1 Portability and connectivity

All participants agreed that DeepNAVI is *truly* a portable navigation assistant. Some participants also emphasized the convenience of carrying the smartphone assistant because of its low weight and small form factor. Participants P3 and P9 conveyed that since the system was attached to the body, they had free hands for carrying various other items. Three participants with experience testing similar navigation assistants shared their views about DeepNAVI.

> *"This is much simpler than I thought. During previous testing with other devices, I need to carry a backpack along with a camera on top of my head. I felt so strange and thought about how people could travel with that on public roads. Now with this system, I feel nothing on me and can walk freely. This is much easier."* (P3, P7)

Four participants commented that they liked that the system is integrated into a mobile phone. Three participants conveyed the benefit of not being dependent on any data network, thereby preserving battery life. Participant P3 was curious about what happens when a call comes while DeepNAVI is in use. The participant tested that by calling the smartphone where DeepNAVI was installed. The DeepNAVI was halted due to the incoming call and resumed operation once the call ended. The participant expressed agreement with giving priority over the DeepNAVI app.

5.2 Information details and accuracy

Mixed feedback was received regarding receiving different types of information about obstacles. For instance, a total of 10 participants responded that they liked knowing the type of obstacles and how far away they were.

> *"My favorite thing about the app is that I can activate it through my sound rather than by pressing anywhere. A useful feature of the app is that it informs various information about obstacles on my path, such as type, distance, and location. However, I would say I like the scene identification feature more."* (P7)

The proposal from five participants was to customize information details according to location.

> *"I think information about obstacles is nice to have in a navigation assistant. But sometimes, I feel there is a lot of information. If there is a customizing option*

> *on what information I prefer in one location and what information I want in another would be good."* (P8)

Three participants conveyed that information needs to be clearer and easier to process.

> *"When I get so much information simultaneously, it is hard for me to understand and act accordingly simultaneously. Maybe I am a slow learner and need time to understand and act."* (P12)

Two participants mentioned they would like to have additional information.

> *"The system can identify if there is a bench or people in front of me, how far they are, whether they are moving etc. All are fine. But if there is a bench in a park and I would like to know if people are sitting there or just standing near. If the system says no one is sitting on a bench, I could sit there to rest."* (P11)

Four participants commented on false obstacle detections. For example, the app detected *chairs* as *furniture*, even though they are both considered as different classes in the trained model of the navigation assistant. They were also issues experienced in detecting glass doors while navigating indoors. However, apart from this, participants were happy about the obstacle-detection feature of the navigation assistant.

5.3 Environment preferences

Participants were asked where they would potentially prefer to use DeepNAVI (such as office indoors, public roads, and home-to-office). A total of seven respondents indicated that they would use it at home or in the office, while two others said they might use it in potentially safer indoor locations such as museums and restaurants. Three participants mentioned they might not try it in crowded places or shopping streets.

> *"I like to use DeepNAVI if I am traveling alone and would like to know things around me. But if it is a crowded area, I might have second thoughts."* (P3)

One participant commented that they preferred to use it both indoors and outdoors because of its features and portability. While three participants commented they were unsure about the use. Nine out of thirteen participants expressed their opinion about using DeepNAVI and white cane together.

> *"I am confident to use DeepNAVI anywhere as a complement to my white cane."* (P12)

5.4 Output modality

Though eleven participants expressed satisfaction with the audio output modality, four participants, including the aforementioned, mentioned they would also like additional signals, such as beep tones and vibrations, to get information while navigating. Two participants (P7, P8) conveyed they needed more voice prompts to communicate with the system.

> *"I would prefer the varying intensity of beep tones when an obstacle is approaching. Just getting audio information about obstacles doesn't help me."* (P2)

> *"If there are vibrations on the system when an emergency occurs, or a hazard (such as a fire) is detected from a long distance, it would be nice to be in precaution."* (P4)

5.5 Safety, trust, and privacy

There were some concerning points regarding the safety, trust, and privacy of smartphone-based navigation assistants. Three participants were highly concerned about trusting a smartphone assistant alone.

> *"I am happy with what I have now (white cane), even though it has limitations. I cannot trust technology blindly."* (P10)

Three participants explicitly used the word' safety' when they described their experience.

> *"I am concerned about safety while navigating with an app in a crowded public environment. It takes time to trust the system and needs practice."* (P3)

Two participants expressed concerns about the privacy handling of the app, such as managing visual data when the app captures the images, such as people's faces, from a public environment.

> *"How is privacy maintained if the phone camera captures videos or photos and detects obstacle information using them when I walk around? In one way, it's good for people like us to know things on our path to navigate safely. But what about when people's faces are being captured unintentionally by us?"* (P13)

5.6 Suggestions for additional features

Six participants expressed the need for a button or voice activator in case of emergency to call friends and family. Participant P11 suggested having an emergency voice command feature and giving details about what was happening in sight. This is basically describing all things in a visual scene. Three participants suggested they would like to hear the names of the streets and shops while walking around.

> *"It would be good if the app's detection could work by giving information about items while shopping, information about nearest public transportation stops and timings. Also, I prefer to have an approximate counter to stairs before I start to climb; I can have an estimate on that before climbing."* (P5)

Participant P13 made a similar suggestion to have a step count before starting. Two participants commented that the system should detect more ground-level obstacles, such as pits or dogs sleeping on paths, which are common in the streets where the tasks were conducted.

5.7 DeepNAVI alone or DeepNAVI+White cane?

We asked participants what they prefer to be used when they navigate alone, DeepNAVI alone, or DeepNAVI+White cane. Eight out of thirteen participants indicated they would prefer a combination of DeepNAVI and white cane.

> *"I like apps and stuffs. But when I walk in unfamiliar places, I might need something I am so sure about. So I*

prefer both the white cane and app combination. Maybe my preference could change tomorrow to using an app alone. But now, no." (P3)

"I like the app. But I cannot think about walking alone to an unfamiliar place with the app, especially in a crowded place with heavy traffic and pits in the middle of the road." (P11)

Only one participant was convinced to use DeepNAVI alone. Two participants said they would use DeepNAVI alone if it had additional features such as emergency contact and an option for customizing information flow. Two participants commented they would like to use a white cane and did not want to complicate their lives with the navigation technology.

"I prefer to use the white cane itself; I am not a technology person and cannot get accustomed to it easily. Moreover, it can create additional complications for me to learn something new at this age." (P6)

6 DISCUSSION

The study found that most participants expressed a positive perception of DeepNAVI on its portability and feature support, such as obstacle detection. In previous research on needs-finding studies [42, 44, 56], it can also be seen that people with visual impairments expect navigation assistants to be portable. It has also been reported that users prefer hands-free navigation assistants [41, 43, 62], which aligns well with the findings herein. Moreover, in similar studies, participants expressed support for obstacle-detection features [26, 29]. In most research studies so far, users have been ignored in favor of incorporating the latest technological advancements into their systems [16, 38]. But from the user experience testing results reported, it can be found that most participants felt comfortable learning and using the DeepNAVI.

Information from the navigation assistant helps the users become aware of physical obstacles in the environment. Although DeepNAVI gives many details about obstacles, this may require some cognitive workload to process all the information provided. Therefore, an information moderation filter could be a possible approach for refining the navigation assistant. Some participants expressed that additional information is needed in specific environments, such as unfamiliar places or places of interest [8, 9, 19]. Hence, one future enhancement could be to provide customization options to users on what information they need while navigating an environment. This user-centric focus on the system design could motivate users to control the information they wish to perceive [44]. In addition, the participant's preference for a stand-alone smartphone app requiring no additional peripherals and network connectivity could be advantageous in the cities where data traffic is high and in rural areas where network coverage is low [36].

Regarding user preferences, most participants were skeptical about using DeepNAVI alone outdoors than in familiar indoor places such as offices or homes. Studies also reported that physical environment characteristics could affect the successful navigation of people with visual impairments [27]. Also, most participants suggested having an emergency button or option to connect with family or friends and information about nearby transportation points [17]. This could be easily implemented in systems such as DeepNAVI, but such a service needs data connectivity at least for some time when the user prefers to use such an option. In support of studies done in [2, 31], it is visible that some participants preferred a multimodal output from the navigation assistant. It is quite understandable that users' preferences vary, and this could be solved through personalized settings. Some research also reported that users prefer personalized interaction experiences in navigation assistants [24, 39].

One common theme among the participants was the importance of trust in the navigation assistant [12]. Some participants commented that if they rely on a smartphone-based navigation assistant, they need to trust that the information the assistants provide is accurate and up-to-date. Many participants with little or no smartphone experience commented that they do not want to consider an alternative to a white cane. Three participants expressed their opinion about privacy on smartphone-based navigation assistants. Several studies have indicated the challenges and needs of people with visual impairments in understanding the privacy-related information of various digital technologies [3, 6, 21]. However, it is also worth focusing on the related issues in smartphone-based navigation assistants for people with visual impairments.

A number of studies have found that adopting and using assistive technology involves social dimensions [1, 18, 40]. Furthermore, some assistive devices may draw unnecessary attention to the user's disability based on the social context in which they are used [54]. Previous research has also shown that social context impacts an individual's decision to use an assistive technology device [52, 53]. These findings show that navigation assistants should be designed to be accepted in social settings [55].

This study provided further insight into the experiences and opinions of real users regarding navigation assistants. Some participants found navigation assistants helpful but also expressed concerns about the accuracy and reliability of the information provided by these assistants [22]. It is a fact that there is a trade-off between high accuracy and low latency in lightweight deep-learning models [46]. Since DeepNAVI uses such models, the performance could be lesser than larger and more computationally-intensive general-purpose models. In alignment with that, previous research reports potential drawbacks using a smartphone-based navigation assistant, such as the limited capability to provide information with high accuracy and reliability compared to systems with high processing power [30, 39, 48]. On the other hand, the DeepNAVI could inform the user about the presence of an obstacle without delay, even though it may not be an accurate detection. This could help the user to avoid obstacles without knowing their type.

It is also noticeable that most participants preferred a combination of DeepNAVI and white cane rather than using DeepNAVI alone. It could be because participants are interested in knowing what is happening in their surroundings and be cautious about obstacles by knowing their type. At the same time, they want something they can trust (such as a white cane) and feel things by scanning and getting satisfied with the sound they hear when it hits an obstacle. By doing so, they can avoid obstacles and be sure of what is ahead [60]. There were also some voices from the participants that this perception might change to solely relying on a smartphone for navigation in the future. However, the results also uncovered that some participants reported negative experiences with navigation assistants. These individuals said smartphone-based assistants as often challenging and expressed reluctance to trust and rely on technology-based navigation solutions. According to them, technology-based aids are unreliable and sometimes fail to interpret instructions accurately and adequately. It has been reported that such assistants are helpful to some, while others consider them a hindrance [49]. Moreover, previous research [11] suggests that instead of replacing human assistance, technology should help people with visual impairments connect and interact more effectively in person. Hence, incorporating varying perceptions of people about technology-aided navigation is an exciting direction to explore.

7 CONCLUSION

The study finds that perceptions of navigation assistants among people with visual impairments are mixed. While these tools are widely considered valuable resources for independent navigation, they also need improved reliability and accuracy to build trust and confidence in their use. Most participants reported positive perceptions about the navigation assistant, such as portability and capability to provide obstacle details. However, a small percentage are still unconvinced to depend on a technology-based navigation assistant and are satisfied with what they are using now (such as white canes). It is possible that perceptions of navigation assistants have an impact on their use, with those with positive perceptions more likely to use them frequently. The study also suggests that participants prefer a customization option for choosing information details and output modality based on location. Moreover, most participants felt comfortable using a combination of the navigation assistant and the white cane. However, some participants also indicated that in the future, this perception might change to relying solely on smartphones to navigate. We think more research is needed to gain a deeper insight into the experiences of people with visual impairments to improve such assistants further.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to all the participants who took part in this study and shared their valuable experiences.

REFERENCES

[1] Ali Abdolrahmani, William Easley, Michele Williams, Stacy Branham, and Amy Hurst. 2017. Embracing errors: Examining how context of use impacts blind individuals' acceptance of navigation aid errors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 4158–4169.

[2] Ali Abdolrahmani, Maya Howes Gupta, Mei-Lian Vader, Ravi Kuber, and Stacy Branham. 2021. Towards More Transactional Voice Assistants: Investigating the Potential for a Multimodal Voice-Activated Indoor Navigation Assistant for Blind and Sighted Travelers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–16.

[3] Tousif Ahmed, Roberto Hoyle, Kay Connelly, David Crandall, and Apu Kapadia. 2015. Privacy concerns and behaviors of people with visual impairments. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 3523–3532.

[4] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: a navigational cognitive assistant for the blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, New York, NY, USA, 90–99.

[5] Dragan Ahmetovic, Masayuki Murata, Cole Gleason, Erin Brady, Hironobu Takagi, Kris Kitani, and Chieko Asakawa. 2017. Achieving practical and accurate indoor navigation for people with visual impairments. In *Proceedings of the 14th International Web for All Conference*. ACM, New York, NY, USA, 1–10.

[6] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan Semaan. 2020. "I am uncomfortable sharing what I can't see" privacy concerns of the visually impaired with camera based assistive applications. In *Proceedings of the 29th USENIX Conference on Security Symposium*. 1929–1948.

[7] Aitor Aladren, Gonzalo López-Nicolás, Luis Puig, and Josechu J Guerrero. 2014. Navigation assistance for the visually impaired using RGB-D sensor with range expansion. *IEEE Systems Journal* 10, 3 (2014), 922–932.

[8] Maryam Bandukda, Aneesha Singh, Nadia Berthouze, and Catherine Holloway. 2019. Understanding experiences of blind individuals in outdoor nature. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–6.

[9] Nikola Banovic, Rachel L Franz, Khai N Truong, Jennifer Mankoff, and Anind K Dey. 2013. Uncovering information needs for independent spatial learning for users who are visually impaired. In *Proceedings of the 15th international ACM SIGACCESS conference on computers and accessibility*. ACM, New York, NY, USA, 1–8.

[10] Alexy Bhowmick and Shyamanta M Hazarika. 2017. An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends. *Journal on Multimodal User Interfaces* 11 (2017), 149–172.

[11] Stacy M Branham, Ali Abdolrahmani, William Easley, Morgan Scheuerman, Erick Ronquillo, and Amy Hurst. 2017. " Is Someone There? Do They Have a Gun" How Visual Information about Others Can Improve Personal Safety Management for Blind Individuals. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY, USA, 260–269.

[12] Stacy M Branham and Antony Rishin Mukkath Roy. 2019. Reading between the guidelines: How commercial voice assistant guidelines hinder accessibility for blind users. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY, USA, 446–458.

[13] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[14] Andrius Budrionis, Darius Plikynas, Povilas Daniušis, and Audrius Indrulionis. 2022. Smartphone-based computer vision travelling aids for blind and visually impaired individuals: A systematic review. *Assistive Technology* 34, 2 (2022), 178–194.

[15] Laila Burla, Birte Knierim, Jurgen Barth, Katharina Liewald, Margreet Duetz, and Thomas Abel. 2008. From text to codings: intercoder reliability assessment in qualitative content analysis. *Nursing research* 57, 2 (2008), 113–117.

[16] Babar Chaudary, Iikka Paajala, Leena Arhippainen, and Petri Pulli. 2021. Studying the navigation assistance system for the visually impaired and blind persons and ICT use by their Caretakers. In *2021 28th Conference of Open Innovations Association (FRUCT)*. IEEE, 55–66.

[17] Hsuan-Eng Chen, Yi-Ying Lin, Chien-Hsing Chen, and I-Fang Wang. 2015. BlindNavi: A navigation app for the visually impaired smartphone user. In *Proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems*. ACM, New York, NY, USA, 19–24.

[18] Aline Darc Piculo dos Santos, Ana Lya Moya Ferrari, Fausto Orsi Medola, and Frode Eika Sandnes. 2022. Aesthetics and the perceived stigma of assistive technology for visual impairment. *Disability and Rehabilitation: Assistive Technology* 17, 2 (2022), 152–158.

[19] William Easley, Michele A Williams, Ali Abdolrahmani, Caroline Galbraith, Stacy M Branham, Amy Hurst, and Shaun K Kane. 2016. Let's get lost: Exploring social norms in predominately blind environments. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, USA, 2034–2040.

[20] Fatma El-Zahraa El-Taher, Ayman Taha, Jane Courtney, and Susan Mckeever. 2021. A systematic review of urban navigation systems for visually impaired people. *Sensors* 21, 9 (2021), 3103.

[21] Yuanyuan Feng, Abhilasha Ravichander, Yaxing Yao, Shikun Zhang, and Norman Sadeh. 2022. Exploring and Improving the Accessibility of Data Privacy-related Information for People Who Are Blind or Low-vision. *arXiv preprint arXiv:2208.09959* (2022).

[22] Nicholas A. Giudice, Benjamin A. Guenther, Toni M. Kaplan, Shane M. Anderson, Robert J. Knuesel, and Joseph F. Cioffi. 2020. Use of an Indoor Navigation System by Sighted and Blind Travelers: Performance Similarities across Visual Status and Age. *ACM Trans. Access. Comput.* 13, 3 (2020), 1–27.

[23] João Guerreiro, Dragan Ahmetovic, Daisuke Sato, Kris Kitani, and Chieko Asakawa. 2019. Airport accessibility and navigation assistance for people with visual impairments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14.

[24] João Guerreiro, Eshed Ohn-Bar, Dragan Ahmetovic, Kris Kitani, and Chieko Asakawa. 2018. How context and user behavior affect indoor navigation assistance for blind people. In *Proceedings of the 15th International Web for All Conference*. ACM, New York, NY, USA, 1–4.

[25] Lilit Hakobyan, Jo Lumsden, Dympna O'Sullivan, and Hannah Bartlett. 2013. Mobile assistive technologies for the visually impaired. *Survey of ophthalmology* 58, 6 (2013), 513–528.

[26] Andreas Hub, Joachim Diepstraten, and Thomas Ertl. 2003. Design and development of an indoor navigation and object identification system for the blind. *ACM Sigaccess Accessibility and Computing* 77–78, 6 (2003), 147–152.

[27] Hernisa Kacorri, Eshed Ohn-Bar, Kris M Kitani, and Chieko Asakawa. 2018. Environmental factors in indoor navigation based on real-world trajectories of blind users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–12.

[28] Shaun K Kane, Chandrika Jayant, Jacob O Wobbrock, and Richard E Ladner. 2009. Freedom to roam: a study of mobile device adoption and accessibility for people with visual and motor disabilities. In *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. ACM, New York, NY, USA, 115–122.

[29] Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2020. Guiding blind pedestrians in public spaces by understanding walking behavior of nearby pedestrians. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–22.

[30] Akif Khan and Shah Khusro. 2021. An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities. *Universal Access in the Information Society* 20, 2 (2021), 265–298.

[31] Jee-Eun Kim, Masahiro Bessho, Shinsuke Kobayashi, Noboru Koshizuka, and Ken Sakamura. 2016. Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing*. ACM, New York, NY, USA, 604–611.

[32] Bineeth Kuriakose, Raju Shrestha, and Frode Eika Sandnes. 2022. LiDAR-Based Obstacle Detection and Distance Estimation in Navigation Assistance for Visually Impaired. In *Universal Access in Human-Computer Interaction. User and Context Diversity: 16th International Conference, UAHCI 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings, Part II*. Springer, 479–491.

[33] Bineeth Kuriakose, Raju Shrestha, and Frode Eika Sandnes. 2022. Tools and technologies for blind and visually impaired navigation support: A Review. *IETE Technical Review* 39, 1 (2022), 3–18.

[34] Bineeth Kuriakose, Raju Shrestha, and Frode Eika Sandnes. 2023. DeepNAVI: A deep learning-based smartphone navigation assistant for people with visual impairments. *Expert Systems with Applications* 212 (2023), 118720.

[35] Masaki Kuribayashi, Seita Kayukawa, Jayakorn Vongkulbhisal, Chieko Asakawa, Daisuke Sato, Hironobu Takagi, and Shigeo Morishima. 2022. Corridor-Walker: Mobile Indoor Walking Assistance for Blind People to Avoid Obstacles and Recognize Intersections. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (2022), 1–22.

[36] Anirudh Nagraj, Ravi Kuber, Foad Hamidi, and Raghavendra SG Prasad. 2021. Investigating the Navigational Habits of People who are Blind in India. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY, USA, 1–10.

[37] Vishnu Nair, Greg Olmschenk, William H Seiple, and Zhigang Zhu. 2020. ASSIST: Evaluating the usability and performance of an indoor navigation assistant for blind and visually impaired people. *Assistive Technology* 34, 3 (2020), 289–299.

[38] Charis Ntakolia, George Dimas, and Dimitris K Iakovidis. 2020. User-centered system design for assisted navigation of visually impaired individuals in outdoor cultural environments. *Universal Access in the Information Society* 21 (2020), 249–274.

[39] Eshed Ohn-Bar, João Guerreiro, Kris Kitani, and Chieko Asakawa. 2018. Variability in reactions to instructional guidance during smartphone-based assisted navigation of blind users. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 2, 3 (2018), 1–25.

[40] Joyojeet Pal, Anandhi Viswanathan, Priyank Chandra, Anisha Nazareth, Vaishnav Kameswaran, Hariharan Subramonyam, Aditya Johri, Mark S. Ackerman, and Sile O'Modhrain. 2017. Agency in Assistive Technology Adoption: Visual Impairment and Smartphone Use in Bangalore. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5929–5940.

[41] Sabrina A Panëels, Adriana Olmos, Jeffrey R Blum, and Jeremy R Cooperstock. 2013. Listen to it yourself! evaluating usability of what's around me? for the blind. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 2107–2116.

[42] Darius Plikynas, Arūnas Žvironas, Andrius Budrionis, and Marius Gudauskis. 2020. Indoor navigation systems for visually impaired persons: Mapping the features of existing technologies to user needs. *Sensors* 20, 3 (2020), 636.

[43] Halley Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K Kane. 2016. The AT effect: how disability affects the perceived social acceptability of head-mounted display use. In *proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, New York, NY, USA, 4884–4895.

[44] Pablo-Alejandro Quinones, Tammy Greene, Rayoung Yang, and Mark Newman. 2011. Supporting visually impaired navigation: a needs-finding study. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1645–1650.

[45] Xukan Ran, Haoliang Chen, Zhenming Liu, and Jiasi Chen. 2017. Delivering deep learning to mobile devices via offloading. In *Proceedings of the Workshop on Virtual Reality and Augmented Reality Network*. ACM, New York, NY, USA, 42–47.

[46] Xukan Ran, Haolianz Chen, Xiaodan Zhu, Zhenming Liu, and Jiasi Chen. 2018. Deepdecision: A mobile deep learning framework for edge video analytics. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 1421–1429.

[47] Sanjeev U Rao, Swaroop Ranganath, TS Ashwin, Guddeti Ram Mohana Reddy, et al. 2021. A Google glass based real-time scene analysis for the visually impaired. *IEEE Access* 9 (2021), 166351–166369.

[48] Santiago Real and Alvaro Araujo. 2019. Navigation systems for the blind and visually impaired: Past work, challenges, and open problems. *Sensors* 19, 15 (2019), 3404.

[49] Manaswi Saha, Alexander J Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. 2019. Closing the gap: Designing for the last-few-meters wayfinding problem for people with visual impairments. In *The 21st international ACM sigaccess conference on computers and accessibility*. ACM, New York, NY, USA, 222–235.

[50] Daisuke Sato, Uran Oh, Kakuya Naito, Hironobu Takagi, Kris Kitani, and Chieko Asakawa. 2017. Navcog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY, USA, 270–279.

[51] Aaron Raymond See, Bien Grenier Sasing, and Welsey Daniel Advincula. 2022. A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind. *Applied Sciences* 12, 6 (2022), 2802.

[52] Kristen Shinohara. 2010. Investigating meaning in uses of assistive devices: implications of social and professional contexts. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*. ACM, New York, NY, USA, 319–320.

[53] Kristen Shinohara and Josh Tenenberg. 2009. A blind person's interactions with technology. *Commun. ACM* 52, 8 (2009), 58–66.

[54] Kristen Shinohara and Jacob O Wobbrock. 2011. In the shadow of misperception: assistive technology use and social interactions. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, New York, NY, USA, 705–714.

[55] Kristen Shinohara and Jacob O Wobbrock. 2016. Self-conscious or self-confident? A diary study conceptualizing the social accessibility of assistive technology. *ACM Transactions on Accessible Computing (TACCESS)* 8, 2 (2016), 1–31.

[56] Paraskevi Theodorou and Apostolos Meliones. 2022. Gaining insight for the design, development, deployment and distribution of assistive navigation systems for blind and visually impaired people through a detailed user requirements elicitation. *Universal Access in the Information Society* (2022), 1–27.

[57] Paraskevi Theodorou, Kleomenis Tsiligkos, Apostolos Meliones, and Costas Filios. 2022. An Extended Usability and UX Evaluation of a Mobile Application for the Navigation of Individuals with Blindness and Visual Impairments Outdoors—An Evaluation Framework Based on Training. *Sensors* 22, 12 (2022), 4538.

[58] Mohammad Moeen Valipoor and Angélica de Antonio. 2022. Recent trends in computer vision-driven scene understanding for VI/blind users: a systematic mapping. *Universal Access in the Information Society* (2022), 1–23.

[59] Wenfu Wang, Yongjian Fu, Zhijie Pan, Xi Li, and Yueting Zhuang. 2020. Real-time driving scene semantic segmentation. *IEEE Access* 8 (2020), 36776–36788.

[60] Michele A Williams, Caroline Galbraith, Shaun K Kane, and Amy Hurst. 2014. " just let the cane hit it" how the blind and sighted see navigation differently. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. ACM, New York, NY, USA, 217–224.

[61] Bichen Wu, Forrest Iandola, Peter H Jin, and Kurt Keutzer. 2017. Squeezedet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. IEEE, 129–137.

[62] Limin Zeng, Markus Simros, and Gerhard Weber. 2017. Camera-based mobile electronic travel aids support for cognitive mapping of unknown spaces. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, New York, NY, USA, 1–10.