

Beyond the Cane: Describing Urban Scenes to Blind People for Mobility Tasks

KARST M.P. HOOGSTEEEN[†]

Schepens Eye Research Institute, Mass Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts, United States of America, karst_hoogsteen@meei.harvard.edu

SARIT SZPIRO

Department of Special Education, University of Haifa, Haifa, Israel, sarit.szpiro@edu.haifa.ac.il

GABRIEL KREIMAN

Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, Center for Brains, Minds, and Machines, Cambridge, Massachusetts, United States of America, gabriel_kreiman@tch.harvard.edu

ELI PELI

Schepens Eye Research Institute, Mass Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts, United States of America, eli_peli@meei.harvard.edu

Blind people face difficulties with independent mobility, impacting employment prospects, social inclusion, and quality of life. Given the advancements in computer vision, with more efficient and effective automated information extraction from visual scenes, it is important to determine what information is worth conveying to blind travelers, especially since people have a limited capacity to receive and process sensory information. We aimed to investigate *which* objects in a street scene are useful to describe and *how* those objects should be described. Thirteen cane-using participants, five of whom were early blind, took part in two urban walking experiments. In the first experiment, participants were asked to voice their information needs in the form of questions to the experimenter. In the second experiment, participants were asked to score scene descriptions and navigation instructions, provided by the experimenter, in terms of their usefulness. The descriptions included a variety of objects with various annotations per object. Additionally, we asked participants to rank order the objects and the different descriptions per object in terms of priority and explain why the provided information is or is not useful to them. The results reveal differences between early and late blind participants. Late blind participants requested information more frequently and prioritized information about objects' locations. Our results illustrate how different factors, such as the level of detail, relative position, and what type of information is provided when describing an object, affected the usefulness of scene descriptions. Participants explained how they (indirectly) used information, but they were frequently unable to explain their ratings. The results distinguish between various types of travel information, underscore the importance of featuring these types at multiple levels of abstraction, and highlight gaps in current understanding of travel information needs. Elucidating the information needs of blind travelers is critical for the development of more useful assistive technologies.

CCS CONCEPTS • Human-centered computing~Accessibility technologies; • Social and professional objects~People with disabilities;

KEYWORDS: Blindness, impaired vision, scene description, outdoor, independence, mobility, navigation, assistive technologies.

1 INTRODUCTION

Blind people experience difficulties with independent mobility, the act of moving through space in a safe and efficient manner [1], which negatively impacts their employment opportunities, social inclusion, and quality of life [2–4]. These difficulties are especially accentuated in unfamiliar routes [5].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Association for Computing Machinery.

1936-7228/2022/1-ART1 \$15.00

<http://dx.doi.org/10.1145/3522757>

Independent mobility requires both spatial orientation and navigation skills. The term spatial orientation refers to the ability to identify the position or direction of objects or points in space [6]. The act of finding a way from one place to another is often referred to as navigation. Without access to visual information both processes can be very challenging.

Mobility aids such as the long cane and guide dog are effective for obstacle avoidance, but provide limited information about the surrounding environment. This environmental information is difficult to acquire without vision, yet it is important for decision making, spatial learning, and cognitive map development [7].

An exciting possibility to potentially alleviate some of these challenges is the development of Electronic Travel Aids (ETAs). ETAs encompass a wide variety of assistive technology such as electronic canes [8], tactile and digital maps [9, 10], and a variety of systems based on GPS [11–14], Radio Frequency Identification (RFID) tags [15–17], and Bluetooth Low Energy (BLE) beacons [18, 19]. Additionally, computer vision systems may provide obstacle detection [20], collision warnings [21], locate doorways [22], crosswalks [23, 24], pedestrian crossing lights [25], and read signs [26].

In many cases, the development of ETAs has been guided and limited by the capabilities of contemporary technology. This technical focus may divert attention away from the users' needs. The scientific community is recognizing the necessity to gain a more comprehensive understanding of these needs, which are critical for the development of future assistive technologies [27, 28]. Previous research explored the information needs by examining navigation practices of blind individuals and studied: the types and level of information required [29–34], how to convey information [35, 36], personal attributes that affect navigation behavior [37] and information mismatches between navigation instructions provided by sighted people and information needs of blind travelers [38, 39].

Beyond those efforts, it is imperative to understand what types of objects along the route are important for blind travelers to know about and how to describe these objects to the blind traveler. Such environmental information is not only important for immediate mobility needs, but also to promote environmental learning for future travel, which is an important aspect yet unaddressed by current assistive technologies [7, 31]. Lacking from prior work is a direct comparison among the importance of different objects, contrasting different descriptions of the same object, and letting the users dictate their own needs.

To address these gaps, we provided information that is readily available to sighted people to blind participants, and measured what aspects of that information they found useful. While sighted people can effortlessly capture large amounts of visual information about a scene in an instant, blind people can only process a small fraction of this information through other sensory modalities. Thus, we examine what information is worth conveying to blind pedestrians to assist with mobility. We set out to investigate *which* objects are useful to describe and *how* these objects should be described.

We conducted two experiments with blind individuals performing outdoor mobility tasks. In the first experiment, participants were asked to voice their information needs in the form of questions to the experimenter. In the second experiment, participants were asked to score scene descriptions and navigation instructions provided by the experimenter in terms of their usefulness. The scene descriptions included a variety of objects with various annotations per object. Additionally, we asked participants to rank order the objects and the different descriptions per object in terms of priority and explain why the provided information is or is not useful to them.

2 RELATED WORK

2.1 Mobility Without Vision

Giudice [7] highlighted key challenges for blind mobility: insufficient access to navigation-critical information, poor training of spatial skills, and overprotective cultural values. It is possible to gain access to navigation critical information through other sensory modalities such as hearing and touch. In practice, this is often difficult, more

limited due to a relatively low bandwidth and resolution compared to vision, and sometimes not possible at all. With training and the use of mobility aids, blind people can travel effectively, although not without limitations.

Blind pedestrians use two main mobility aids – guide dogs and long canes. Guide dogs are effective in avoiding obstacles and hazards. Guide dog owners report some disadvantages such as the responsibility to care for and exercise the dog [40, 41], which limits their desirability. The most widely adopted mobility aid is the long cane, which warns other pedestrians that the user is blind or visually impaired and, if used correctly, can help detect ground-level obstacles within a one-meter range. Tactile paving, textured ground-level surfaces that are often installed on railway station platforms or near hazardous street crossings [42], can provide important orientation and navigation cues to long cane users. Unfortunately, this structural environmental modification is expensive and, therefore, not widely available.

Listening to the echoes of self-generated sound cues, by one’s own footsteps or tapping a cane on the floor, is another strategy used by blind people. Strelow and Brabyn [43] demonstrated that such cues can be used to guide locomotion but they are limited to large objects like walls. Other sound cues like traffic noise, the footsteps of other pedestrians, or the opening of electric doors can also be used by blind people to aid travel [44].

Altogether, these mobility aids and techniques provide only limited information about the environment outside of the direct path of travel. Hence, they are only of limited use for spatial learning and cognitive map development. More research is needed into technologies that provide real-time access to off-route information and promote environmental learning [7].

2.2 Descriptions to Assist with Blind Mobility

2.2.1 Information mismatch.

Auditory verbal descriptions are a straightforward method to convey visual information about the environment to blind pedestrians. However, providing accurate and useful verbal descriptions remains a critical challenge. Sighted companions provide information that often mismatches the needs of blind people, even when these individuals know each other [39]. For example, long cane users might prefer information that will help them locate certain objects with their cane, since it can serve as a landmark, while sighted people tend to provide information in ways that will avoid such collisions since the same object might be classified as an obstacle from their point of view [39]. Indeed, prior research found that sighted and visually impaired people use different navigation cues for describing a route [45]. The use of a route description based on the cues preferred by sighted people resulted in higher work-load and frustration as compared to one based on the cues preferred by people with visual impairments [46].

The information mismatch between sighted and blind pedestrians can also result from the different perspectives used by these groups while providing route descriptions. In a study by Brambring [38], both sighted and blind individuals were asked to describe familiar routes of 400-1000 meters. The sighted participants gave “environment-oriented descriptions” while blind participants provided “person-oriented” descriptions (e.g., take a right at the pharmacy versus from your current location, take a right in 50 meters). Person-oriented descriptions were also found to be frequently used in written directions exchanged between blind members of a mailing list [47]. An information mismatch may also result from the variability amongst the blind travelers. Prior work found that personal attributes affect navigation behavior [37], and noted navigation instructions provided by ETAs should be customizable in terms of timing and content based on walking pace, navigation skill and style [48].

2.2.2 Description content.

Information preferences of visually impaired pedestrians have been studied in relation to ETAs. Gaunet and Briffault [34] examined route descriptions of blind participants to inform specifications for a verbal wayfinding aid. The authors noted what information should be provided in scenarios involving intersections and street crossings. Golledge *et al.* [33] found that visually impaired pedestrians preferred information about landmarks, streets and street crossings, routes and destinations, buildings, building entrances, and transit options. Aziz *et al.*,

[30] explored preferences for an auditory route description for pre-navigation and found that participants were interested in a wide variety of content like parks, schools, street crossings, distances and directions, and orientation information. However, what exact information should be provided was not studied. Moreover, what is preferred or interesting may not necessarily be the most useful.

Verbal information was also found to be a valuable addition to tactile maps to improve user satisfaction [49] and route understanding [50], but a research gap was noted with regard to the exact spatial and environmental information that should be included [32]. Papadopoulos *et al.*, investigated the usefulness of environment descriptions in audio-tactile maps of campuses. The researchers let visually impaired participants score 213 environment descriptions that were categorized beforehand by the experimenters under the themes: “safety”, “location of services”, and “wayfinding or orientation while travelling”. The authors identified the 30 most useful pieces of information for each category. Descriptions ranged from solely naming an object (e.g., “stairs”), describing the relation to other objects (e.g., “stairs leading to subway”), to concepts such as “dangerous areas” [32]. Due to this large variety of descriptions it remains difficult to tell what information, and in what scenario, makes a description useful. For example, when describing “stairs leading to subway” people may want to avoid the stairs as a hazard, or be interested in knowing there is an entry point to the subway. Missing from the existing literature is a more structured evaluation into what exact information is and is not useful, and in what context this information is useful.

2.2.3 Information framework.

Several frameworks have been proposed that may help to structure the information to be evaluated (see [10, 51]). We adapted a model by Yaagoubi *et al.* [52] that describes a three-level hierarchical structure that focuses on orientation information. The first and lowest level (i.e., the most detailed and least abstract) contains information about objects in the immediate surroundings such as crosswalks, sidewalks, road signs, and imminent landmarks (e.g., a bench on the sidewalk). The second level (at an intermediate level of abstraction) includes information about street sections, intersections, and local landmarks (e.g., a specific building). The third and highest level (the most abstract) contains objects such as: neighborhood, district, and city. Since the present study focuses on information needs of blind pedestrians, and travel at the third level usually involves transportation by vehicles, we focus on the first and second levels.

Orientation and navigation information may be provided to blind pedestrians and can be categorized according to a similar hierarchical structure. Low-level information includes information that can be used to locate objects in the immediate surroundings and directions to maneuver around these objects (e.g., orientation information like “there is a fire hydrant in front of you” or navigation information like “take a step to the left to avoid the fire hydrant”). Intermediate-level information is similar to the information provided by driving GPS systems (e.g., orientation information like “you are on Cambridge Street” or navigation information like “take a right at the next intersection”).

While many of the proposed ETAs focus on some of these types of information, how this information is being provided across different ETAs varies in format, length, and level of detail. For example, intermediate-level orientation and navigation information provided by Google Maps Detailed Voice Guidance: “head west on Cambridge Street, it is about 100 feet to your next turn”, versus information provided by GetThere: “In 105 feet, turn right from Cambridge Street to Staniford Street, 455 feet from there, your destination is on the left”. Some navigation systems even offer extensive customization options to the user, e.g., GPS-talk can provide verbal directions in: right left front back, clockface, compass and headings in degrees.

Low-level orientation and navigation information is used in computer vision and mixed reality systems that use pitch, clicking, and voice commands (e.g., VizWiz LocateIt [53, 54]) via spatial audio (e.g., [55, 56]) to help the user locate and navigate towards objects in the immediate surroundings. These systems help the user by providing the relative position of an object (e.g., “ITEM is 2 feet away, 30 degrees left and 5 inches below the

camera view” [54] and provide navigation instructions to guide the user towards a certain object (e.g., “go forward, up, down, left, right”).

Some of these systems do not just focus on objects in front of the user, but in all directions. Using the soundscape approach, virtual objects emit informative sounds which are presented to users via spatial audio. In order to make such a system comprehensible, May *et al.* [55] used object distance and object importance ratings provided by participants to assign their salience properties, such as the volume falloff and type of information provided (non-speech sounds versus text-to-speech). The authors stress the importance of understanding the information needs of blind people, and argue for their inclusion in the design process to determine object usefulness, and consequently, salience.

While all these studies yielded valuable insights into the information needs of blind people, lacking from prior research is a structured evaluation of information, and understanding of what specific information is preferred and prioritized. We designed two experiments that allowed us to probe information needs in a structured manner while blind people perform outdoor mobility tasks.

3 METHODS

3.1 Participants

Participants were recruited via mailing lists of the blind and low vision community in the Greater Boston Area. Two eligibility criteria were included in the recruitment letter; a minimum age of 18, and a residual sight that at most could only allow for light perception. We excluded legally blind individuals with partial sight who are likely to depend on their vision while performing mobility tasks. We recruited 13 blind individuals (ages 21 to 72, 8 males), 5 of whom were early blind (Table 1). Additional participant characteristics were recorded prior to the experiments. Participants reported how frequently they traveled independently in unfamiliar environments on a weekly basis. Some participants noted they did not travel independently on a weekly basis and reported their travel frequency over multiple weeks, which was then divided by the number of weeks. Participants also reported their confidence during independent travel while only using a cane in unfamiliar environments, and overall experience with the use of a cane. These characteristics were measured using a 6-point scale ranging from “no confidence/experience” to “a lot of confidence/experience”. All participants reported they used a cane for mobility except for one participant who reported having no experience with the use of a cane.

Table 1: Participants characteristics.

Participant	Gender	Age	Age at onset of blindness	Group	Travel frequency*	Travel confidence**	Cane experience***
1	Male	59	Birth	Early blind	3	2	3
2	Male	72	53	Late blind	0.5	0	5
3	Male	68	2	Early blind	5	4	5
4	Female	68	33	Late blind	0	0	0
5	Female	33	Birth	Early blind	5	4	5
6	Female	68	Birth	Early blind	6	3	5
7	Male	66	50	Late blind	0	0	3
8	Male	67	9	Late blind	1	3	5
9	Male	63	60	Late blind	0.25	2	3
10	Male	66	29	Late blind	1	4	4
11	Male	64	42	Late blind	0.25	1	5
12	Female	21	Birth	Early blind	3	3	5
13	Female	27	7	Late blind	0.1	3	5

* Frequency of independent travel in an unfamiliar environment on a weekly basis

** Confidence during independent travel in an unfamiliar environment

*** Experience with the use of a white cane

3.2 Experiment 1 – Participants Asking Questions

In the first experiment, participants completed mobility tasks using only their cane. There were 4 tasks: find the front entrance for a particular building, walk on the sidewalk towards the next intersection, exit a plaza and cross a street (with traffic lights but without an audible signal), and find and enter a train station. Participants were brought to a starting position in the vicinity of their task objective (i.e., approximately 300 meters from their destination). Each task took about 5 - 10 minutes, and experiment 1 lasted approximately 30 minutes.

During the 4 tasks, the participants were accompanied by a sighted investigator, who did *not* guide the participants. The investigator only intervened in potentially dangerous situations. The participants were instructed to request any information they needed, in the form of questions. In turn, the sighted investigator would answer these questions, as asked. Participants were fitted with a microphone to record the questions during the experiment.

All tasks were performed in environments *unfamiliar* to the participants and conducted around the same time of day (noon). However, conditions for each participant varied slightly, due to the dynamic nature of the urban environment which resulted in varying levels of background noise due to traffic and construction, and a different number of potential obstacles like pedestrians and parked vehicles. One participant was not able to complete this experiment and was therefore excluded from the analysis.

3.3 Experiment 2 – Participants Rating of Investigator Descriptions

The second experiment was conducted right after the first experiment with the same 13 participants. In this experiment, the participant was assigned a single mobility task: get to the entrance of a particular building in downtown Boston, which involved a 15-minute walk from the starting point. During this task the sighted investigator provided information to the participant who rated the information on a usefulness scale from 0 (not useful), 1 (slightly useful), to 5 (extremely useful). The sighted investigator described 20 predefined subjects along the route. These included 18 objects, such as a pole or a building, route instructions, and a general description of the participant's current location (Fig. 1, and Table 2). Route instructions and a general description of the current

location were included to test how this type of navigation and orientation information is rated and prioritized compared to the object descriptions. The selection of objects was made based on systems proposed in prior research ([20–26]), expressed information preferences ([32, 33]), and based on objects in the environment in which the experiment took place. Two criteria were considered:

- Information from both a low and intermediate level of abstraction (as described in [52]) was represented in the selection.
- Two instances of each object occurred along the route.

Each object was described using five sentences on average that were rated by the participants as independent pieces of information. The different sentences described various aspects of an object such as the appearance, dimensions, and location.

We exclusively refer here to the 18 objects and not to information about the Route and Current Location, as these could not be described similarly. In the remainder of this work, we make the distinction between subjects (i.e., all 20 subjects described to participants) and objects (a subset of 18 objects out of the 20 subjects).

Information about the route and location descriptions were provided at two levels of precision: imprecise and precise. For the route, the descriptions were formulated as in the following examples: “walk to the end of the block, and take a right” (imprecise), “walk 100 feet towards 11 o’clock, then continue towards 9 o’clock” (precise). The Location descriptions were formulated as in the following examples: “there is a building on your left, close” (imprecise), “there is a building at 9 o’clock, 10 feet away” (precise). All descriptions were formulated from the participant’s point of view since blind people use “person-oriented” descriptions while describing familiar routes [38, 47].



Figure 1: One of the 8 stops, with three example sentences given per object (Experiment 2). The objects described here are highlighted in the image using a color corresponding to the colors of the text-balloons.

Table 2: Examples of sentences scored by the participants.

Object	Aspect	Sentence scored by participant
Street	Detect	There is a street.
	Activity	- *
	Appearance	- *
	Dimension	There is a street, it has 4 lanes.
	Identity	There is a street, it is called Cambridge street.
	Location (Imprecise)	There is a street, to your left, close.
	Location (Precise)	There is a street, 9 o'clock 6ft away.
	Car	
Car	Detect	There is a car.
	Activity	There is a car, it is parked.
	Appearance	There is a car, a grey sedan.
	Dimension	- *
	Identity	There is a car, a BMW with license plate: 0KMP92.
	Location (Imprecise)	There is a car, in front of you, far.
	Location (Precise)	There is a car, 11 o'clock 30 ft away.

* Not all aspects are described for every object. Some were not considered in this experiment due to time constraints or because they were not applicable for a certain object (e.g., describing the identity of a fire hydrant).

Each subject type was discussed twice during the experiment, which accumulated to a total of 40 subjects that were discussed over 8 stops (in groups of 5 per stop). During a stop the participant was asked to imagine the following scenario "you are alone, walking towards the direction you are facing and have access to the following pieces of information:". Then, the sighted investigator presented five different subjects (e.g., [Fig. 1](#), [3](#)). Two instances of the same subject were never presented during a single stop. Descriptions were pre-prepared and read from a script.

As an introduction to each object and to check for answer consistency across the two instances, the experimenter always started describing an object with a *detect* sentence (e.g., instance 1 – description 1: there is a building, instance 2 – description 1: there is a building). In case of more dynamic objects such as a car or pedestrian, a virtual instance of such an object was described to make sure all participants received the same information. Participants were not informed of this, because we did not want to risk a bias in the scoring of these objects. None of the participants remarked missing audio cues in cases where virtual objects were described. This might be due to the background noise at the city center.

For the objects, the two instances were described at a different relative position in terms of distance and angle (e.g., position 1 – there is a building, at 3 o'clock 100 feet away, position 2 – there is a building, at 2 o'clock 30 feet away). In this comparison, we only considered the scores for the *location (precise)* sentences. The objects were divided into 3 groups based on the characteristics of the difference in relative position of the two instances ([Fig. 2](#)). The difference in relative position of objects in each group is summarized in [Table 3](#). We hypothesized that participants would score as more useful the description of objects that are closer but at a similar relative angle (*distance only*), at the similar distance but more in the prospective path of the participant (*angle only*), or both closer and more in the prospective path of the participant (*combined*). In the analysis we excluded objects with an opposing effect of relative distance and angle. For example, position 1 – there is a building, 12 o'clock 200 feet away, position 2 – there is a building, 9 o'clock 20 feet away. The experiment was not designed to investigate this effect; hence, we do not have enough data to examine this relationship. We consider this combination of a

potential positive score effect due to a decreased relative distance, versus a potential negative score effect due to increased relative angle, to be too complex to be adequately evaluated and interpreted.

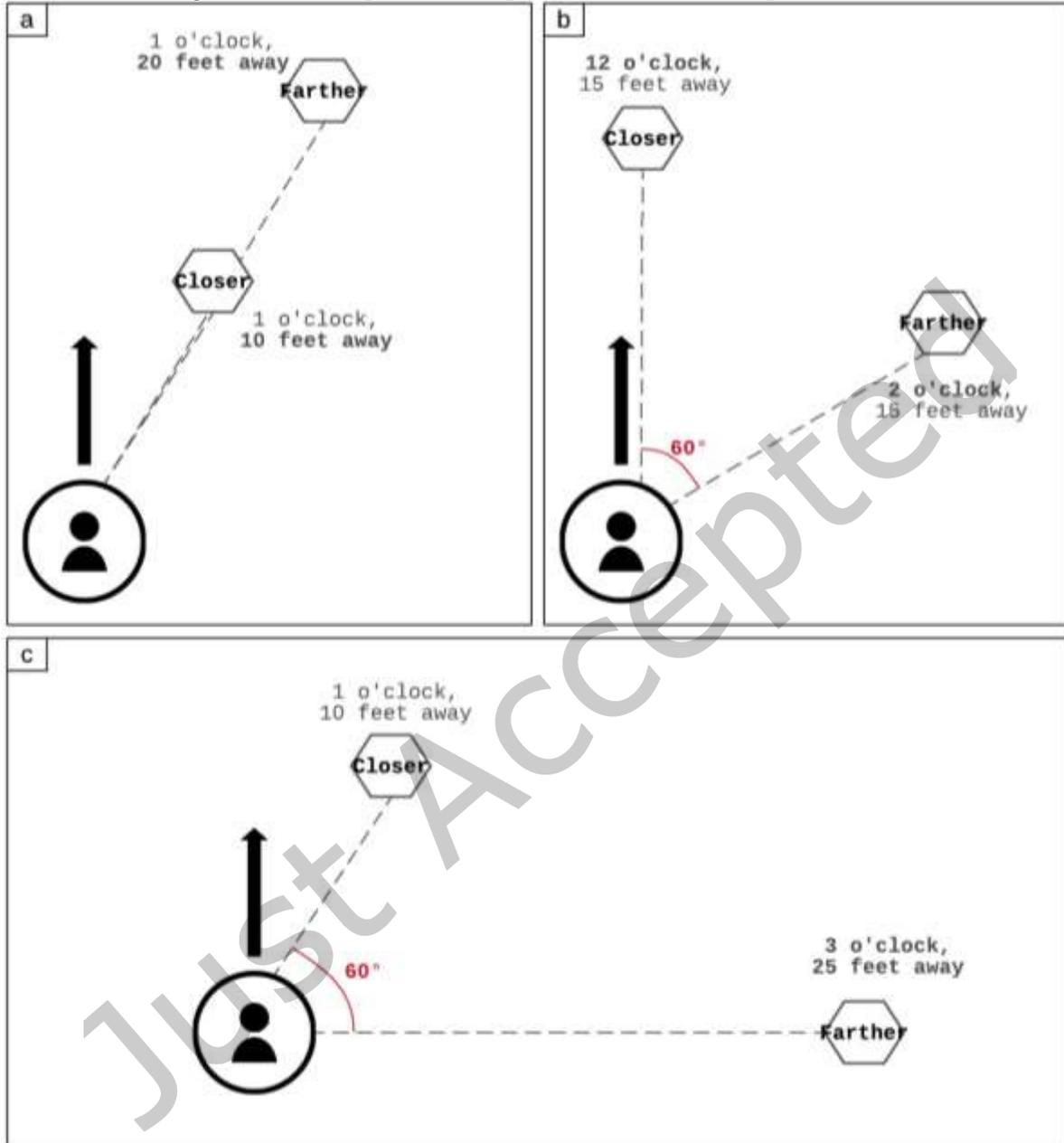


Figure 2: Illustration of the various types of relative position of objects (hexagon) from the participant (circle). These include: difference of relative distance with a similar angular direction (a), difference of relative angle with a similar distance (b), difference of both relative distance and angular directions (c).

Table 3: Groups of difference in objects relative positions to the participant, and object locations.

Group	Object	Distance: closer, farther (ft)	Angle: closer, farther (O'clock)*
Distance Only			
	Car	30, 12	11, 12
	Fire Hydrant	20, 6	1, 1
	Grass	30, 3	3, 3
	Intersection	180, 15	12, 11
	Park	300, 18	2, 1
	Tree	18, 6	2, 1
Angle Only			
	Crosswalk	12, 12	9, 12
	Curb	12, 12	9, 12
	Pole	9, 9	2, 12
Combined			
	Pedestrian	45, 30	2, 12
	Stairs	24, 4	2, 12
	Traffic Signal	60, 30	9, 12

* We only considered differences in angular direction of over 30 degrees.

After a subject was described and the descriptions were scored for their usefulness, participants were asked to rank order the described aspects in terms of priority. We included this rank order task to examine which information needs to be prioritized. Subsequently, participants were asked if it is useful to talk about a certain object and to explain why information about this object is or is not useful to them. After all five subjects at a stop were discussed using this procedure, the participants were asked to rank order the subjects in terms of priority for the task performance (Fig. 3). The duration of each stop was 10 - 15 minutes, and experiment 2 lasted a total of 120 minutes.

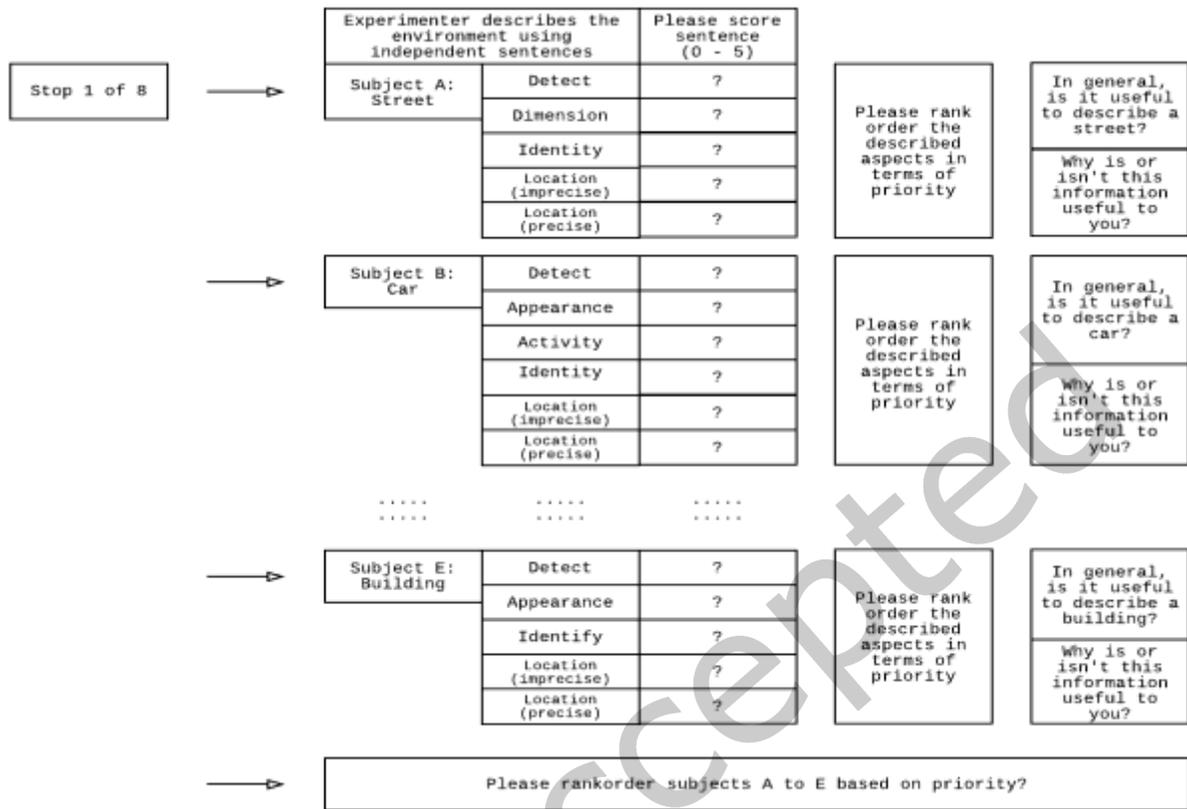


Figure 3: Procedure used in experiment 2. The procedure was executed as follows: for each object, descriptions were read to the participant, the participant was asked to score each description based on usefulness and rank order based on priority. After all sentences were rated and rank ordered the participant was asked if in general, it is useful to describe a given object, and to explain why the information is or is not useful to them. When this process was completed for all 5 objects within the stop, participants were asked to rank order the subjects in terms of priority.

3.4 Data Analyses

Participants' questions recorded in the first experiment were transcribed. The questions were categorized based on subject, description aspect, and on the type of information and level of abstraction as described in section 2.2.

Ratings and rank order data recorded in the second experiment were analyzed using SPSS. The assumption of homogeneity was tested using *Levene's F test*. Further, the assumption of normality was considered to be satisfied if *skew* < |2| and *kurtosis* < |9| [57]. All participant ratings provided during experiment 2 are considered to be non-parametric data. When multiple tests were conducted, a *Benjamini-Hochberg procedure* was used to correct for alpha inflation. Since our analysis involves pairwise comparisons with a relatively large number of tests, we used this procedure instead of the more conservative *Bonferroni correction* to mitigate the risk of type 2 errors (i.e., a failure to reject a false null hypothesis). The alpha level for all tests was set to 0.05.

4 RESULTS

Participants reported their frequency of independent travel in unfamiliar areas on a weekly basis (Fig. 4a). Early blind participants traveled more than late blind participants (4.4 ± 1.34 versus 0.39 ± 0.41 times per week,

respectively; here and throughout mean \pm std; $t(4) = 6.5$, $p = 0.001$, t-test assuming unequal variances). Participants also reported their confidence level during independent travel in unfamiliar areas on a 6-point scale, with a higher score indicating a higher confidence level (Fig. 4b). Early blind participants reported higher confidence (3.2 ± 0.84 versus 1.63 ± 1.6 , $t(11) = 8.06$, $p = 0.034$, t-test assuming equal variances).

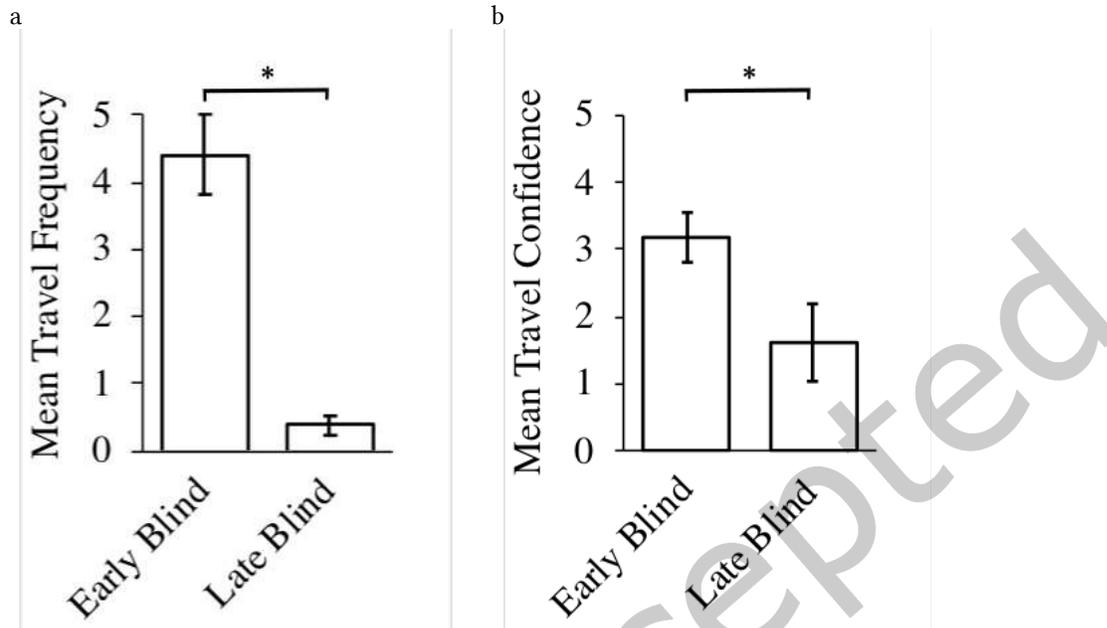


Figure 4: Mean self-reported *travel frequency* in times per week (a) and *travel confidence* (b), for independent travel in unfamiliar areas. The sample is split into early blind participants ($N = 5$) and late blind participants ($N = 8$). Early blind participants reported travelling more frequently and more confidently than late blind participants. The asterisk (*) denotes a significant difference in *travel frequency* and *travel confidence* between the two groups, and the error bars provide the standard errors.

4.1 Experiment 1 – Participants Questions

4.1.1 Number of Questions.

Each participant ($N = 12$) asked on average 25.6 ± 6.6 questions over the four tasks, about one question per minute. Early blind participants on average asked 3 times fewer questions than late blind participants (12 ± 8.37 versus 35.2 ± 25.5 ; $t(8) = 2.25$, $p = 0.027$). Two late blind participants asked a high number of questions (66 and 75). However, due to the small sample size these are not considered to be outliers.

4.1.2 Participants' Questions Content

Only 20% of the questions were open-ended (e.g., where should I go?), while 80% were closed-ended (e.g., should I go left or right?). The question content was categorized based on subject (Fig. 5a). Information about the *route* was most frequently requested. This category included questions like: “where do I go from here?”, “should I go left or right?”, and “am I going in the right direction?”. Questions concerned with maintaining a certain heading were another notable frequent occurrence within this category.

Overall, participants asked questions about 30 different subjects, 24 of these occurred less than 10 times each and were labeled *other*. Within this group the objects: crosswalk, traffic, traffic signal, and wall occurred five times or more. The *other* group also included objects that could be classified as potential hazards, such as obstacle, pole, step, and stairs. However, all of these were asked only once or twice per object. Questions that did not explicitly define a subject were *not considered* in this study. This category included questions like: “what is to my left?”, “is there anything in front of me?”, and “what is that?”. In case such a question occurred, participants were

first asked to specify what information they required. For example, “what is to my left?” was reformulated as “what buildings are to my left?”. If they were not able to specify, the question was *not considered* further, if they did specify, it was included in the appropriate category. The object *street* was the fourth biggest group and included questions such as, “are we still on Staniford street?”, “where is Cambridge street?”, and “does this street have an island in the middle?”.

In addition to categorizing the questions according to the subject, the questions’ content was also categorized based on aspect (i.e., the type of information requested about a certain object, Fig. 5b). The aspect labeled *direction* was the largest group and almost entirely overlapped with the category *route*. The aspect *location* was the second-largest, and included questions such as: “where is the building entrance?” “is Staniford street in front of me?” and “how far away is the intersection?”. The aspect labeled *detect* was the third-largest and included questions such as: “any stores here?”, “is that the corner”, and “is that a wall?”. The aspect labeled *identity* included questions such as: “which building entrance is this?”, “is this Government Center?”, and “what is the name of this street?”. All questions that could not be categorized based on aspect were grouped under the label *not considered*. This group included questions like “am I okay now?”, “good enough?”, and “what am I looking for?”. The aspect labeled *appearance* included questions such as: “is this area open?”, and “is that part of the curb or just decoration?”.

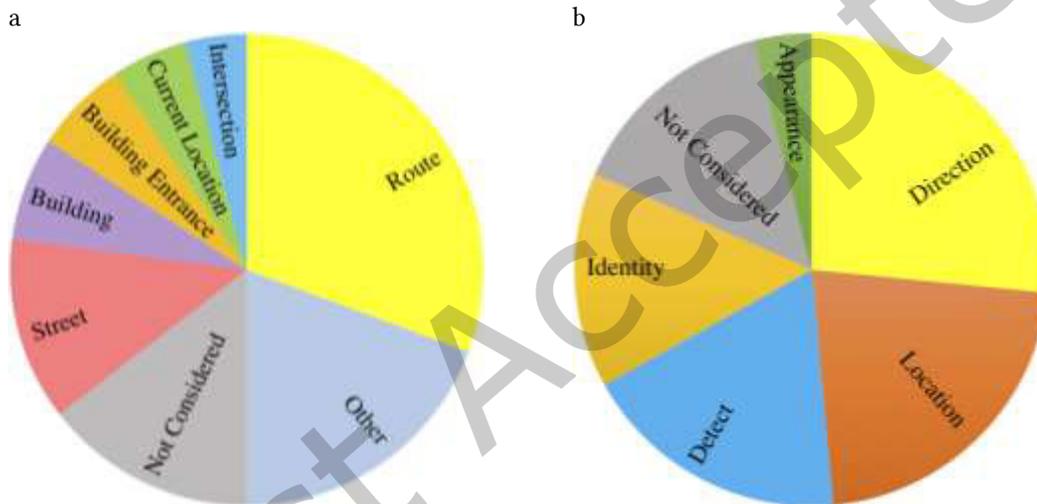


Figure 5: Content of 302 questions categorized based on subject (a) and aspect (b).

When the questions were categorized according to the type of requested travel information, most questions were aimed at gathering *low-level orientation information* (Fig. 6). This category included questions such as: “are there people in front of me?”, “I’m on the sidewalk, is that correct?”, and “am I at the corner?”. The second-largest category includes all questions referring to *low-level navigation information* such as: “if I turn in this direction, is it a clear path?”, “can I just pass here?” and “if I go around this way, am I still on the sidewalk?”. The third-largest category included all questions referring to *intermediate-level orientation information* such as: “what is the name of this street?”, “is Staniford street in front of me?” and “how far away is the intersection?”. The fourth-largest category included all questions referring to *intermediate-level navigation information* such as: “is Staniford street the next right?”, “What direction is Staniford street?” and “will there be a left- or right-hand turn?”.

Not all questions could be categorized according to these four types of travel information, so a fifth category labeled *other* was added. This category contained questions aimed at gathering information that can be used to guide the participants’ behavior in specific scenarios such as: “is the door open?”, “does the traffic light have an audible signal?” and questions that could not be categorized such as “what is that?”.

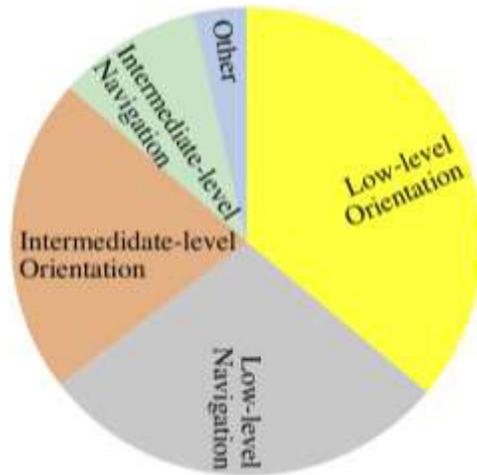


Figure 6: Content of 302 questions categorized according to the type and abstraction level of travel information.

4.2 Experiment 2 - Rating Descriptions

4.2.1 Object Usefulness.

Because the results in experiment 1 showed differences between the early blind and late blind group, we hypothesized that these groups may have different information needs during outdoor mobility tasks, which could result in different mean usefulness scores per object in experiment 2. Yet, no significant differences were found. Hence, we aggregated scores across all participants for the remainder of this section.

All of the described objects were perceived as at least somewhat useful, even when the sighted investigator provided only the *detect* descriptions (e.g., “there is a pole”). Mean usefulness scores for the object *detect* descriptions are illustrated in [Fig. 7](#) and show the least useful description was scored only 1.5 points lower than the most useful description, on a 6-point scale. We sought to examine the difference in participant scores between objects more closely. Since we did not describe the same aspects for all objects, and did not provide the same number of descriptions per object ([Fig. 1](#) and [Table 2](#)), we compared only the *detect* descriptions. This description was similar across all objects (e.g., “there is a pole” or “there is a building”), except for the object itself. To conduct this analysis, we first checked for answer consistency by comparing the usefulness scores for the detect sentences for the two times each object was described. We found no significant difference between the scores and collapsed the two detect scores for each object. Next, we compared detect scores between objects and found a significant difference among the object *detect* scores ($\chi^2(17) = 77.88, p < 0.001, \text{Friedman test}$).

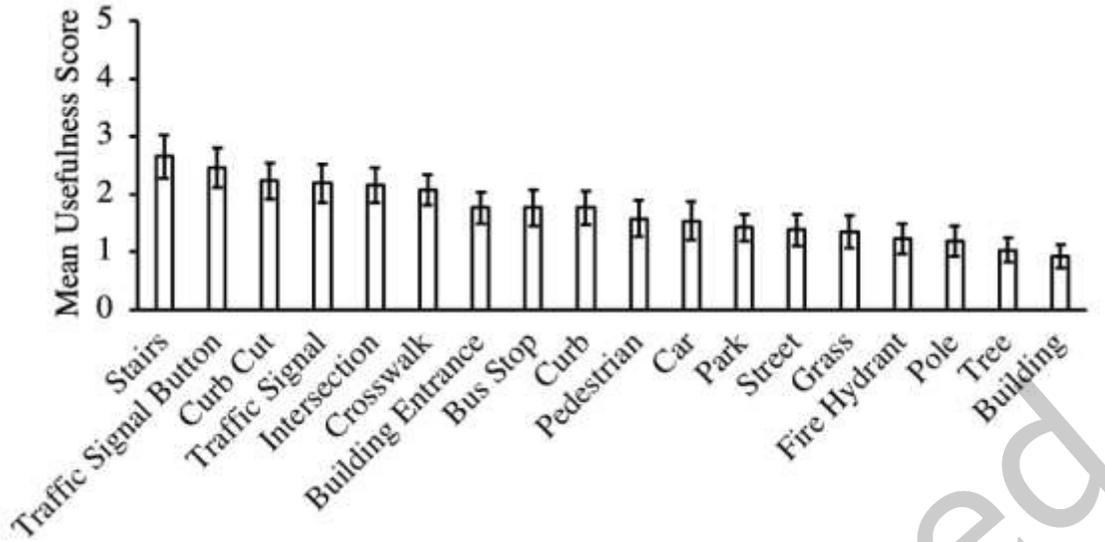


Figure 7: Mean usefulness score per object for the *detect* descriptions ordered from high to low usefulness. The error bars are the standard errors. Due to a high number of significant differences a pairwise comparison table with significant indications is provided in [Appendix A](#).

To examine which objects differed significantly, we conducted a post hoc analysis using *Wilcoxon Signed Ranks tests*. After applying the *Benjamini-Hochberg* procedure to correct for multiple testing we found that 40 out of 153 pairwise comparisons yielded a significant effect, $p < 0.05$. Seventeen out of the 18 objects differed significantly from at least one other object (see [Appendix A](#) for a pairwise comparison table of all 18 objects).

We sought to examine the effect of relative positions of objects on participant scores. From the participant's view point, the two instances of an object were described at a different relative position in terms of distance and angle. Mean usefulness scores for both relative positions for each of these categories are presented in [Fig. 8](#).

The group named *distance only* included object instance pairs where the change in relative position was based on distance (i.e., both instances occurred at similar relative angle but at a different relative distance). We hypothesized that closer instances would be scored as more useful. Indeed, closer object instances were scored as more useful than those positioned farther from the participant ($N = 78$, 3.31 ± 1.24 versus 2.73 ± 1.36 , $z = -3.25$, $p = 0.003$, respectively, *Wilcoxon Signed Ranks Test* with *Benjamini-Hochberg correction* for multiple testing).

The group named *angle only* included those object instance pairs where the change in relative position was based on angle (i.e., both instances had a similar relative distance but at a different relative angle). We hypothesized that instances occurring closer to the prospective path of the participant would be scored as more useful. Object instances in the prospective direction of the participant did not receive higher scores compared to those positioned farther out of the participant's path ($N = 39$, 3.26 ± 1.27 versus 3.44 ± 1.31).

The group named *combined* included object instance pairs where the difference in relative position was based on *both* distance and angle. Object instance pairs only qualified if one of the two instances occurred at both a decreased relative distance and relative angle (i.e., when an instance is relatively closer and closer to the prospective path). The group is named *combined* since a positive effect on the mean usefulness score due to the decreased distance can potentially be complemented by a positive effect on the score due to the decreased angle. Thus, we hypothesized that instances that are closer and occurred closer to the prospective path of the participant would be scored as more useful. Indeed, these instances were, on average, associated with a higher score compared to those that were farther ($N = 39$, 3.36 ± 1.63 versus 2.9 ± 1.55 , $z = -2.6$, $p = 0.014$, *Wilcoxon Signed Ranks Test* with *Benjamini-Hochberg correction* for multiple testing).

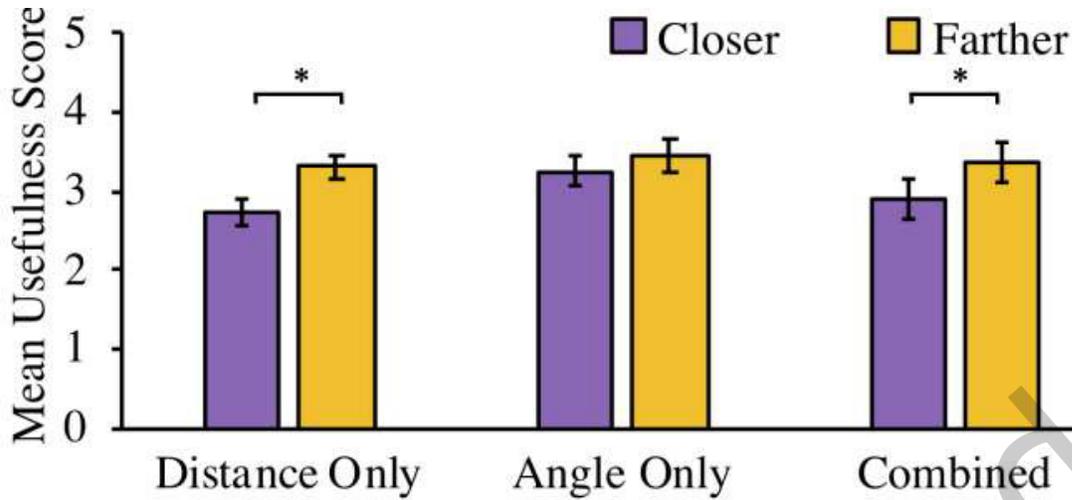


Figure 8: Mean usefulness score for each group of object instances at a closer and farther position in relation to the participant. The figure shows that the description of objects which are: closer but at the same relative angle (*distance only*), at the same distance but closer to the prospective path of the participant (*angle only*), both closer and closer to the prospective path of the participant (*combined*), are evaluated as more useful, with the asterisk (*) indicating a significant difference between the two positions in the *distance only* and *combined* groups. The error bars are the standard errors.

After rating all five subjects at a stop, as to their usefulness, the participant rank ordered those subjects in terms of priority (the higher the rank, the higher the priority, Fig. 9). Each subject was described at two stops (two instances per subject) and ranked against a set of four other subjects described at that stop. Since the distribution of subjects throughout the experiment was based on the environment, each subject was rank ordered against a unique set of other subjects. There was no significant difference between the rankings from early blind and late blind participants.

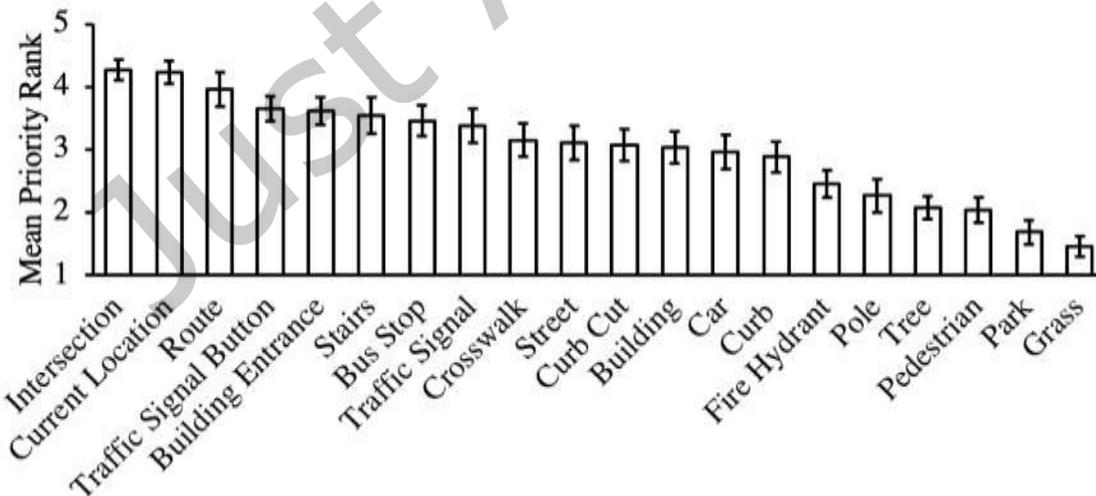


Figure 9: Mean priority rank per subject across participants ordered from high to low priority. All 20 subjects are included here. These scores are based on a total of 415 rank orders across all participants. The error bars are the standard errors.

4.2.2 Imprecise versus Precise Descriptions.

There was no significant difference between the mean usefulness scores provided by early blind and late blind participants for precise and imprecise descriptions. Hence, we aggregated scores across all participants. The mean usefulness score for the *precise location* descriptions ($N = 468, 3.24 \pm 1.36$) was significantly higher than for the *imprecise location* descriptions ($N = 468, 2.69 \pm 1.39$) (Fig. 10, $z = 9.15, p < 0.001$, *Wilcoxon Signed Ranks Test* with *Benjamini-Hochberg correction* for multiple testing). For the subject *route*, there was no difference in the mean usefulness scores between Precise and Imprecise descriptions. For the remainder of the results, we will focus on the Precise Location descriptions.

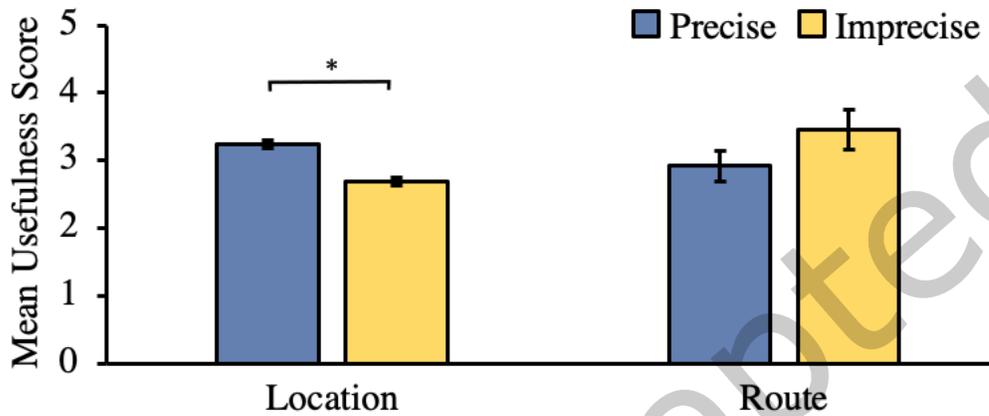


Figure 10: Mean usefulness scores for the two levels of description precision: *precise* and *imprecise* for the *location* aspect and object *route*. Compared to the *imprecise* descriptions, the *precise* descriptions were rated as more useful for the *location* aspect. The asterisk (*) indicates a significant difference. The error bars are the standard errors.

4.2.3 What aspect to describe.

For each aspect description, the mean usefulness scores provided by participants from the early blind and late blind groups are presented in Fig. 11. We hypothesized that these two groups might have different information needs during an urban mobility task, which could result in different mean usefulness scores for various aspects. Yet, only the aspect *identity* (e.g., “there is a building, it is called CVS Pharmacy”) was different between groups and was scored higher by the early blind group ($U = 7625.5, p = 0.012$, *Mann-Whitney U test* with *Benjamini-Hochberg correction* for multiple testing). For the remainder of the analysis on aspect usefulness scores we collapse the results across the two groups.

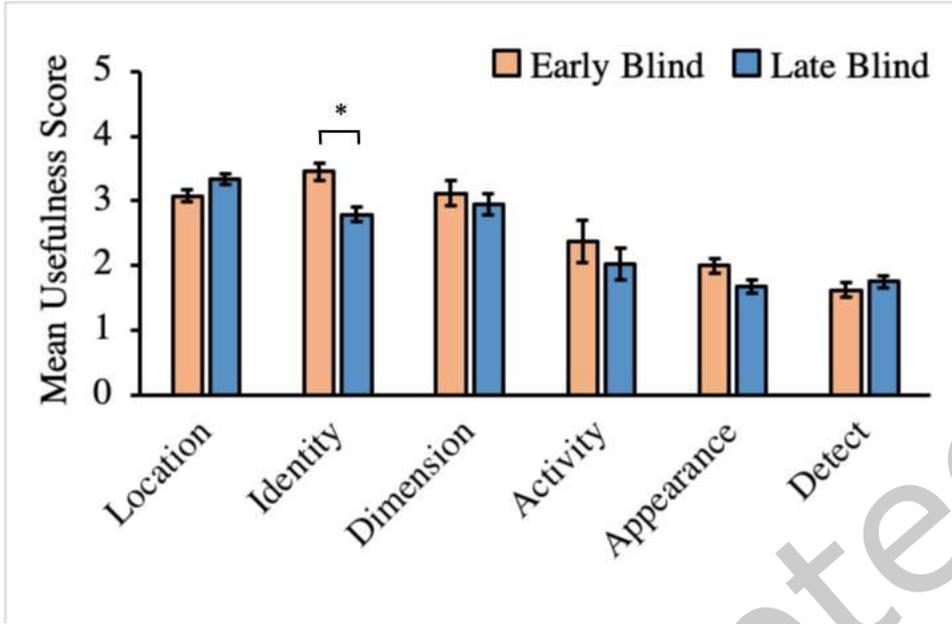


Figure 11: Mean usefulness score per aspect as provided by the early blind and late blind participants. The asterisk (*) indicates a significant difference between the two groups for the *identity* aspect. The error bars are the standard errors.

The analysis in [Section 4.2.1](#) revealed that 40 out of 153 object pairs differed significantly from one another, consequently we did not aggregate scores across all objects. We identified a subset of objects that do not differ significantly from each other in usefulness and used this as an input for the analysis of aspects. We included the objects: *building entrance*, *bus stop*, *car*, *curb*, *fire hydrant*, *grass*, *park*, *pedestrian*, *pole* and *street* (see [Appendix A](#)).

We hypothesized that the various aspects ([Fig. 12](#)) would be scored differently and found a significant overall effect: $\chi^2(5) = 33.32$, $p < 0.001$, *Friedman test*. Post hoc analysis using *Wilcoxon Signed Ranks tests* and the *Benjamini-Hochberg procedure* to correct for multiple testing found that 8 out of 15 pairwise comparisons yielded a significant effect, $p < 0.05$. There was a statistically significant effect for the following aspect combinations: *activity – detect* ($p = 0.009$), *activity – location* ($p = 0.011$), *appearance – dimension* ($p = 0.001$), *appearance – identify* ($p < 0.001$), *appearance – location* ($p < 0.001$), *detect – dimension* ($p = 0.005$), *detect – identify* ($p < 0.001$), and *detect – location* ($p < 0.001$). Hence, we confirmed our hypothesis that some description aspects are scored differently.

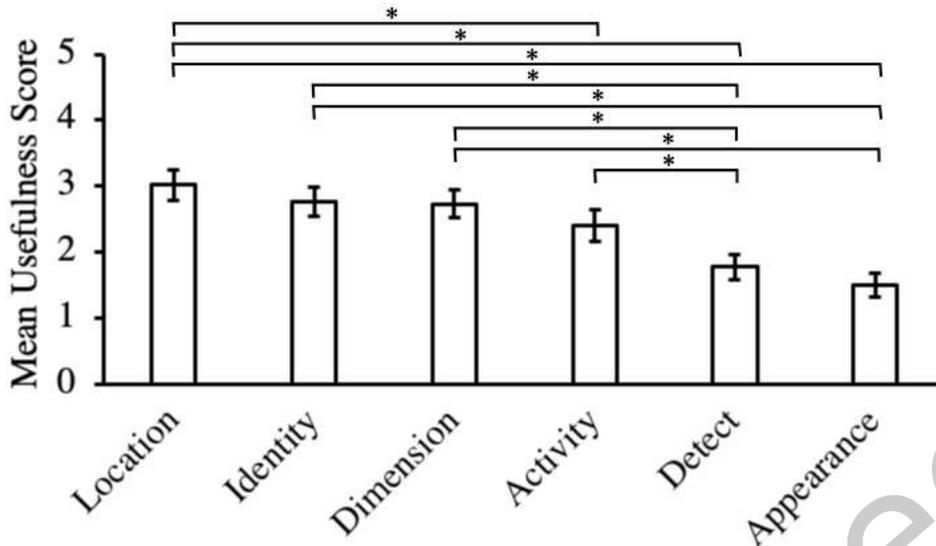


Figure 12: Mean usefulness score per description aspects aggregated across participants and the objects: *building entrance, bus stop, car, curb, grass, park, pedestrian, pole and street*. The asterisk (*) indicates a significant difference between aspects and the error bars are the standard errors.

In addition to scoring aspects based on usefulness, participants rank ordered aspects in terms of priority. As illustrated in [Table 2](#) and [Fig. 1](#), only a few aspects were described for each object. Therefore, not all aspects have been rank ordered for each object and the item scores had to be normalized. When the rankings were separated between the early and late blind groups, we found notable mean priority rank differences for the Location, Identity, and Activity aspects ([Fig. 13](#)). A statistically significant different mean rank was found only for the aspects: Location and Identity ($U = 13776, p < 0.001$ and $U = -3619.5, p < 0.001$, *Mann-Whitney U test* with *Benjamini-Hochberg correction* for multiple testing).

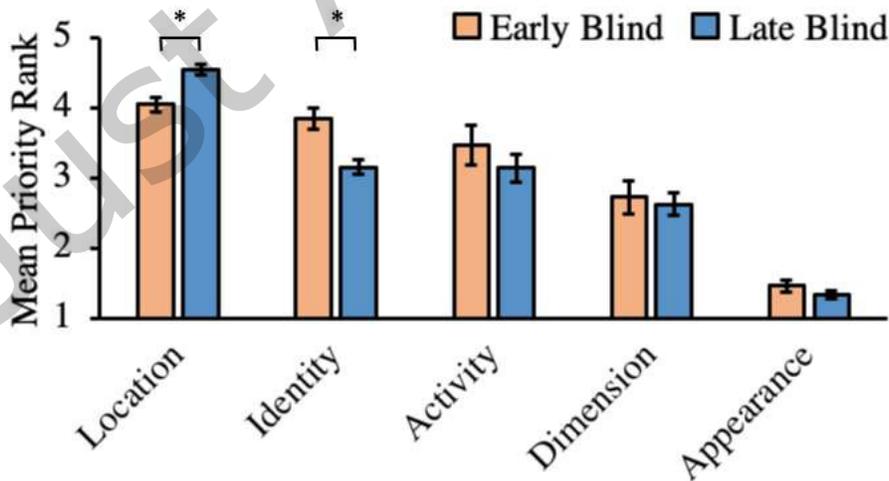


Figure 13: Mean priority rank per aspect as provided by the early blind and late blind participants. A higher rank is an indication of higher priority and the asterisk (*) indicates a significant difference between the two groups for the aspects: *location* and *identity*. *Detect* sentences were not ranked against the other aspects, so they are not included in this graph. The error bars are the standard errors.

4.2.4 Behind the Ratings.

To better understand the ratings of participants we also asked participants to verbally describe why the information is or is not useful to them on an object-by-object basis (Fig. 3). Nine out of the 13 participants found it difficult to answer this question. For 99 out of the 520 described objects, participants from both the early and late blind groups stated they were not able to answer the question. For 95 out of the 99 times this occurred participants were not able to explain why they deemed an object as useful. In cases where participants were able to explain why information was useful to them, many explanations contained concepts such as: *landmarks* (51), *orientation* (40), *safety* (36), *directions* (32), *obstacles* (25), *context* (24), and *hazards* (11). While many of these explanations were self-evident, two notable uses of information stand out. First, participants explained they need information to guide their behavior in certain scenarios. For example, the type of building entrance (e.g., standard versus automatic sliding doors) or knowing if a staircase goes up or down in order for them to appropriately adapt their walking speed so they have more time to find it with their cane. Another scenario that was frequently mentioned was crossing an intersection. In particular, information about the type of intersection is useful as it determines how they would approach crossing it. For example, one of the participants mentioned: “If you are not familiar with the area you don’t precisely know what all the sounds mean, it really depends on what kind of intersection it is. If it is two way, you simply start listening for cars on the left and right, if it is four way you listen for longer until you recognize the traffic pattern.” Second, participants described that information is useful indirectly. For example, one of the participants noted that a curb can indicate a parked vehicle, that in turn, can be a secondary obstacle next to the curb itself. This indirect use of information was not just limited to describing an object in general, but also mentioned with regard to the aspect being described. For example, one of the participants noted that describing the color of a building is useful because he can use this information to describe where he is or needs to go when asking (remote) others for help. This indirect use of information was mentioned by four participants.

5 DISCUSSION

We aimed to get initial insights into the information needs of blind pedestrians performing urban mobility tasks. We set out to investigate *which* objects are useful to describe and *how* those objects should be described.

Few questions were asked by the participants during the first experiment, with the exception of two participants. Questions were mostly close-ended, seeking confirmation of the participant’s assumptions. The low number and restricted nature of these questions may imply that the participants found it difficult to come up with questions about what they do not see or perceive, which is a limitation of our methodology.

Differences between early blind and late blind individuals could affect information needs during travel. Several studies found that early blind participants have a higher likelihood of impaired performance during tasks that required certain spatial skills such as inferring spatial relationships (for a review, see [58]). On the other hand, no significant difference in performance was found for wayfinding tasks that involved orientation and navigation [59].

Giudice [7], noted that it is more informative to use an approach in which information access requirements of the group being investigated is used instead of a categorization based on the type and nature of blindness. In the present study, late blind participants reported travelling less frequently and less confidently than early blind participants. These characteristics may result in a higher need for information and explain why late blind participants asked more questions compared to the early blind participants. Moreover, orientation and mobility (O&M) proficiency was identified as a factor that influenced the type and amount of information that blind travelers required for spatial learning [31]. A recent study by Pigeon *et al.* [60] found few significant age effects when measuring cognitive load of blind people while walking. However, the onset of blindness was not considered in their analysis. In the present study, there was no significant difference in average age between early and late blind participant groups. All participants reported they had at least a moderate, to a lot of experience

with the use of a cane with the exception of one participant. However, no irregularities were found in the data of this participant.

5.1 Which Objects to Describe

We examined which objects are useful to describe to early and late blind travelers. The results from the second experiment suggest that the early and late blind information needs in terms of objects are similar.

All of the described objects were perceived as at least somewhat useful, even when the sighted investigator provided very limited information per object (e.g., “there is a bus stop” which was scored as 1.769 – just over slightly useful). Notably, mean scores for such descriptions are lower than reported in prior work by Papadopoulos *et al.*, [32] who found the description: “bus stop” to be one of the most useful pieces of information (4.763 - very useful, on a 5-point scale). These discrepancies likely resulted from the difference in experimental procedures. Papadopoulos *et al.*, let participants score the usefulness of information in relation to 3 reference themes: “safety”, “location of services” and “wayfinding or orientation while travelling”. These predetermined themes, which were not part of our experiment likely have impacted the results. For example, “stairs leading to subway” was categorized under “Safety”. However, this description could also have been categorized under “Location of services”, the former focusing on the stairs as a hazard while the latter focusing on the availability of the subway. In the present study we provided the context of a mobility task instead of a reference theme. Also, information gathered from a questionnaire administered in a lab, office or even home setting may lead to different results as compared to information gathered whilst performing outdoor mobility tasks. Being at the location you would use the information allows for a more direct evaluation. Furthermore, the number of sensory inputs is different, and thereby cognitive load which can affect the ability to learn information about the environment [61]. Nonetheless, the discrepancies between our findings and those of prior work may suggest that information requirements of blind travelers may vary considerably between mobility related activities (i.e., information required for pre-navigation is different from information required while performing actual mobility tasks such as crossing a street. In this regard, the use of multi-method approaches to probe information needs may be required before the findings can be generalized beyond their specific use case.

We found statistically significantly different usefulness ratings between many objects. This implies that it might be beneficial to build a hierarchical model to prioritize scene descriptions based on usefulness. In terms of usefulness based on the object itself, a relatively higher score was given to information that can be used to cross intersections, such as information about curb cuts, crosswalks and traffic signals. This finding complements prior work in which participants noted they would like a navigation system to provide information about street crossings [30, 33]. A relatively lower score was given to information about potential collision hazards such as a tree, fire hydrant, pole and pedestrian. This low score may seem surprising, but this is in line with the results of the first experiment, where participants asked almost no questions about these kinds of objects. It might be that our participants are less interested in this kind of information since they can acquire it with their cane. In case of the pedestrians, the participants remarked that there are obstacles that get out of the way themselves. Indeed, the use of a cane may warn other pedestrians to get out of the way although this was not always the case during the experiment (e.g., when pedestrians were watching their phone).

Besides the usefulness rating based on the object itself, the indirect use of information is an important factor that could be considered, as it can highlight underlying needs. For example, as one of the participants noted: “a building means people, whom I can ask for directions”. Information can also be used to predict the existence of other objects (e.g., a curb may indicate a parked vehicle). These results are consistent with findings by May *et al.* [55] who noted that blind participants used certain objects to predict the surrounding geometry during an indoor navigation task.

The effect of relative position (e.g., objects on path) translated to a significant difference in mean usefulness scores for the *distance only* and *combined* groups. These results indicate that descriptions of objects which are only closer, or closer and more in the prospective path of the participant are rated as more useful, although the

effect is small. No statistically significant difference was found in the mean usefulness scores of the *angle only* group. This might have resulted from the inability of participants to estimate the relative position and thereby a collision risk based only on relative angle which can be quite complex, especially as some blind individuals may have less developed spatial skills [62, 63]. Another option is that nearby objects are important for blind travelers regardless of relative angle.

No significant difference was found between the priority rankings of subjects provided by the early and late blind participants in this study. Across participants, information about the *route* and *current location* was prioritized over information about objects, with the exception of the object *intersection*. Participants gave a relatively high priority to all information associated with intersections which reflects the usefulness scores for these objects. Potential collision hazards such as a *fire hydrant*, *pole* and *tree* got a relatively low priority. This is consistent with the relatively low usefulness scores for these objects and the results of the first experiment, where participants asked almost no questions about such objects.

Our results suggest that while a lot of research is focused on the development of assistive devices that help to avoid obstacles (e.g., [20, 21, 64–66]), in fact long cane users prioritize other types of information. While this does not undermine the usefulness of information about upcoming obstacles, these observations highlight opportunities to address more prominent needs. Simultaneously, our results underscore a point stressed by Williams *et al.*, [39], who noted that by encountering objects, rather than avoiding them, a cane user receives important tactile feedback that is used to perceive the environment. This approach could be used to repurpose systems that are used for obstacle avoidance (e.g., [20]) in which all detected planes besides the ground plane are treated as an obstacle. Such a system could guide the user to useful objects like a wall to be used as a guide for low-level navigation.

5.2 How to Describe Objects

The results from the second experiment suggest that early and late blind people have similar information needs in terms of aspects. Only describing the *identity* of objects was rated significantly different by these two groups.

All aspects described in the second experiment were perceived as at least somewhat useful by the participants. Notably, the sole detection of an object (e.g., “there is a building”) is perceived as useful as a sentence describing the appearance of an object (e.g., “there is a building, with a grey concrete facade”). This is an indication that more information is not always more useful. This finding confirms prior work that found the availability of more spatial information about an indoor environment did not lead to significantly higher performance of blind participants on a wayfinding task [29]. Moreover, since environmental sound cues and echoes can be used effectively by blind people [67–69], verbal transmission of any travel information should be limited in scope and volume, as much as possible, to mitigate the compromise of the use of these strategies.

The *detect* and *appearance* sentences were rated as significantly less useful compared to *dimension*, *identity* and *location* sentences. This implies that sentences such as “there is a building, it has five floors”, “there is a building, called CVS Pharmacy”, or “there is a building at 9 o’clock, 10 feet away” are perceived as significantly more useful than the *detect* and *appearance* descriptions. These results suggest that it is important to consider what aspect is being described as some are perceived as less useful. More work is required to determine if there is a hierarchical order between the various aspects.

For orientation information, such as describing the location of objects in the environment, *precise* descriptions are perceived as more useful than *imprecise* descriptions (i.e., “there is a building at 9 o’clock, 10 feet away” is perceived as more useful than “there is a building on your left, close). For navigation instructions, *imprecise* descriptions were perceived as more useful than *precise* descriptions (i.e., “walk to the end of the block, and take a right” versus “walk 100 feet towards 11 o’clock, then continue towards 9 o’clock”). However, this effect was not statistically significant. Hence we found no evidence to support prior work, which suggested distance measures in navigation instructions for visually impaired travelers to be expressed in more imprecise concepts like “steps” and

“city blocks” [47]. Future research is needed to examine the usefulness of the various formats’ in which navigation information can be conveyed.

The results suggest that the early and late blind participants prioritize the described aspects differently. The aspects: *location* and *identity* were ranked significantly different by these two groups which could relate to the role of certain spatial skills during mobility. For example, early blind people could prioritize information that helps them to identify a building and identify its location indirectly through an address, while late blind people prioritize information that helps them to locate it directly from their own perspective. Overall participants gave a higher priority to *location* and *identity* descriptions which is consistent with the mean usefulness scores for these aspects. The *appearance* aspect had low priority, which again, reflects the low mean usefulness score for this aspect.

5.3 Travel Information Framework

In order to examine and structure the information needs of blind individuals we build upon a hierarchical semantic model proposed by Yaagoubi *et al.* [52]. In the first experiment, the majority of questions asked by participants were aimed at acquiring *low-level orientation information* (detecting, identifying and locating various objects in the direct surroundings). While ETAs that feature this kind of information for urban travel have been proposed, this was mostly limited to specific objects like traffic lights [25] or crosswalks [23, 24]. Although these specific objects are undoubtedly important, the results from both of our experiments indicate that blind travelers are interested in a greater variety of objects, even when these objects are outside of the direct path of travel. ETAs such as accessible GPS-systems provide such information about relatively large objects like a building or a park. However, our results show that information about smaller objects are also perceived as useful (e.g., describing grass to the right). Blind people may use such information to develop a more detailed cognitive map of the environment which they can use to aid future travel in the same space.

The results of the second experiment suggest that many of these objects are used as landmarks. These results are consistent with the findings from a study by Brambring [38], who noted that blind travelers primarily focused on the descriptions of landmarks while describing a familiar route. However, in the present study participants followed an unfamiliar route. Hence it is not surprising that many of the questions asked by participants in the first experiment were aimed at acquiring navigation information. However, *low-level navigation information* was requested far more frequently than the *intermediate-level navigation information*, typically provided in ETAs for urban use such as GPS systems. In part, this is because *intermediate-level information* remains relevant for longer and, therefore, is not required as often. The results underscore the need for *low-level navigation information* such as very detailed course corrections to prevent stray, and proximal directions to maneuver around obstacles.

Besides the need for orientation and navigation information, participants mentioned they need information to guide their behavior in specific scenarios. For example, information that can inform their approach towards crossing an intersection, entering a building or information they can use to adapt their walking speed in anticipation of objects like stairs or a curb. Like orientation and navigation information this type of information can inform behavior at various levels of abstraction.

Even though the highest level of abstraction was not considered in the present study, it is a vital component for mobility as it contextualizes information provided in the lower levels. In order to provide an overview of these levels and the different types of information needs (i.e., *orientation*, *navigation* and *behavioral*), we present a framework of travel information in Table 4. In this framework we provide example sentences for each type of information on these various levels of abstraction and illustrate how they can build on each other to assist during a mobility task. This framework can provide a more structured insight into the information needs of blind travelers and can help to uncover gaps in our understanding of these needs. While our results indicate that locating, identifying, and describing the dimensions of a variety of objects in the environment is useful and prioritized over describing their Appearance, such insights mostly inform our understanding of *low* and *intermediate-level orientation information*. Our study was not designed to test navigation information and

information used to guide specific behaviors at various levels of abstraction. Hence, more work is required to cover these gaps.

Table 4: Framework of travel information that can assist blind pedestrians with independent mobility in unfamiliar environments. The information can be used for *orientation*, *navigation*, and to inform *behavior* and specific scenarios. Information is provided on three levels of abstraction namely: *low*, *intermediate*, and *high*. The different types and levels of information build on one another and together can assist with independent mobility.

Abstraction level	Orientation	Navigation	Behavioral
Low	There is a fire hydrant in front of you.	Take a step to the left to avoid the fire hydrant.	Caution, step down coming up.
Intermediate	You are on Cambridge Street.	Take a right at the next intersection.	The intersection has traffic lights.
High	You are in downtown Boston.	Head towards building X in the West End district.	Take bus 15 at State street and get off at the third stop.

5.4 Recommendations for Future Studies

Future work is required to test if what participants perceived to be useful is in fact useful when implemented in an ETA. Such an evaluation is especially important before findings can be generalized and applied for the development of commercially available ETAs. In the present study, participants were not able to explain their rating in almost one out of every five ratings. These results highlight the possible limitations of a system design approach where information is curated based only on forced choice user ratings, and underscores the need for additional measures to probe the usefulness of information such as the time required to complete a task or cognitive map development. The cases where participants were able to explain their ratings can result in a more comprehensive understanding of information needs. For example, the use of a building as a resource for gathering intermediate-level navigation information, and the usefulness of a description of its color for *intermediate-level orientation information* provide insight into the underlying needs and the use of information. With this understanding, developers of ETAs may be able to address such needs more directly, leading to more efficient travel overall. Also, we found different results from prior work in which similar information was provided in a different context. This suggests more work is needed to examine needs for different mobility related activities, scenarios and contexts.

The most effective and efficient method for conveying various types of information should be studied. Real-world experiments should be conducted in which information such as presented in the travel information framework is provided in various formats and through (combinations of) modalities, especially given that linguistic-based interfaces that use spatial language require more interpretation by the user, resulting in a higher cognitive load as compared to perceptual interfaces such as spatialized audio, touch or vision [7]. Prior work explored the use of such a perceptual interface for waypoint-to-waypoint guidance and found it scored higher in terms of guidance performance and user preference compared to other hearing based interfaces.[35, 36]. Future work should investigate how such interfaces, possibly supported by inputs through other modalities, can be used to convey other types of information such as the location information of on and off-path objects.

Furthermore, more work is needed to investigate how a residual vision, which the vast majority of blind people have, may be utilized for mobility. Indeed, a lot of studies worked on a closely associated topic, so-called head mounted display systems for people with low vision (for a review see [28]). However, blind people have not received as much attention in this avenue of research as their residual vision is often regarded to be not functional for such use cases. Nonetheless, even vision characterized as no more than light perception may yet be utilized to convey relevant visual cues, although probably in an abstract form. Such work is also relevant for the development of visual prosthesis which currently may only restore a rudimentary form of vision with a limited resolution [70, 71]. These limitations only allow for selective visualizations and thus require image pre-processing

to filter the environment for relevant information; a topic which is receiving more and more research interest in recent years (see [72–74]).

5.5 Limitations

Our experiments were framed using the context of a specific mobility task in an unfamiliar environment: get from point A to point B. Although this is a common challenge experienced by many blind travelers, it touches upon limited number of aspects of mobility. A different context such as taking a walk for leisure, or a task involving public transport, is likely to bring in a different set of scenario specific challenges and thereby, information needs. Moreover, discrepancies between findings of the present study and prior work suggests the results may not be generalizable beyond their intended use case (i.e., assistance during mobility tasks using only a cane). Hence, it is important to note, that information needs of blind individuals that use a guide dog or partially sighted individuals that do not use a cane are likely to be different. Environmental factors such as the infrastructure, and more specifically the street layout, are also likely to impact information needs, which in turn, limits the generalizability of the present research. Further, we did not consider all types of information that could be provided for assistance with mobility tasks, such as information about objects that were not selected or did not occur on our route, and high-level information which was shown to be utilized by blind travelers [31].

6 CONCLUSION

Building a more comprehensive understanding of information needs of blind travelers is an important component for the development of future assistive technologies. In the present study we explored these needs by examining *which* objects are useful to describe and *how* these objects should be described. We addressed these questions by asking blind pedestrians to voice their information needs in the form of questions and by letting them rate and rank order scene descriptions based on usefulness and priority.

We identified factors that affect the usefulness of scene descriptions such as the object being described, its distance and the types of descriptions. Additionally, our results reveal differences between early blind and late blind participants, namely the number of times information was requested and prioritization of specific information. In the present study, late blind participants reported travelling less frequently and less confidently which may have resulted in a higher or different need for information. This illustrates the importance of user-specific adjustment of information in mobility tasks. Finally, participants were frequently unable to explain their ratings which could suggest limitations of a system designed solely based on forced choice user ratings.

We present a framework of travel information that could help structure information needs during independent mobility. We distinguish between various types of information and levels of abstraction. While a lot of ETAs focus exclusively on information at one level of abstraction the results of both experiments underscore the importance of providing information at multiple levels.

Our results show a way towards a more thorough understanding of the information needs of blind pedestrians. Beyond such understanding, more work is required to investigate the most effective and efficient method for conveying information, which in combination with the former, can be used to inform the development of future ETAs.

ACKNOWLEDGMENTS

We thank our participants, the Massachusetts Association for the Blind and Visually Impaired (MABVI), the National Federation of the Blind (NFB), and the Visually Impaired and Blind User Group (VIBUG) for their time and efforts. This work was supported in part by a US Department of Defense grant (W81XWH-16-1-0033), NIH core grant P30EY003790, The Promobilia Foundation (Sweden), The Netherland-American Foundation, and a gift from The Margaret and Leo Meyer and Hans M. Hirsch Foundation.

REFERENCES

- [1] P. R. Kosior, "Foundations of Orientation and Mobility," *Int. J. Orientat. Mobil.*, 2010, doi: 10.21307/ijom-2010-008.
- [2] A. McClimens, N. Partridge, and E. Sexton, "How do people with learning disability experience the city centre? A Sheffield case study," *Heal. Place*, 2014, doi: 10.1016/j.healthplace.2014.02.014.
- [3] A. Lubin and D. Deka, "Role of Public Transportation as Job Access Mode: Lessons from Survey of People with Disabilities in New Jersey," *Transp. Res. Rec. J. Transp. Res. Board*, 2012.
- [4] G. Townley, B. Kloos, and P. A. Wright, "Understanding the experience of place: Expanding methods to conceptualize and measure community integration of persons with serious mental illness," *Heal. Place*, 2009, doi: 10.1016/j.healthplace.2008.08.011.
- [5] E. Foulke, "The perceptual basis for mobility.," *Am. Found. Blind. Res. Bull.*, 1971.
- [6] A. Benton and D. Tranel, "Visuo-perceptual, visuo-spatial, and visuo-constructive disorders.," 1993.
- [7] N. A. Giudice, "Navigating without vision: Principles of blind spatial cognition," in *Handbook of behavioral and cognitive geography*, Edward Elgar Publishing, 2018.
- [8] S. Bhatlawande, M. Mahadevappa, J. Mukherjee, M. Biswas, D. Das, and S. Gupta, "Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 6, pp. 1148–1159, 2014.
- [9] M. A. Espinosa and E. Ochaita, "Using tactile maps to improve the practical spatial knowledge of adults who are blind," *J. Vis. Impair. Blind.*, vol. 92, no. 5, pp. 338–345, 1998.
- [10] M. Chen, H. Lin, D. Liu, H. Zhang, and S. Yue, "An object-oriented data model built for blind navigation in outdoor space," *Appl. Geogr.*, vol. 60, pp. 84–94, 2015.
- [11] H. Makino, I. Ishii, and M. Nakashizuka, "Development of navigation system for the blind using GPS and mobile phone combination," in *Proceedings of 18th annual International Conference of the IEEE Engineering in Medicine and Biology society*, 1996, vol. 2, pp. 506–507.
- [12] W. Balachandran, F. Cecelja, and P. Ptasiński, "A GPS based navigation aid for the blind," in *17th International Conference on Applied Electromagnetics and Communications, 2003. ICECom 2003.*, 2003, pp. 34–36.
- [13] H. Petrie, V. Johnson, T. Strothotte, A. Raab, S. Fritz, and R. Michel, "MoBIC: Designing a travel aid for blind and elderly people," *J. Navig.*, vol. 49, no. 1, pp. 45–52, 1996.
- [14] L. Ran, S. Helal, and S. Moore, "Drishti: an integrated indoor/outdoor blind navigation system and service," in *Second IEEE Annual Conference on Pervasive Computing and Communications, 2004. Proceedings of the*, 2004, pp. 23–30.
- [15] S. Chumkamon, P. Tuvaphanthaphiphat, and P. Keeratiwintakorn, "A blind navigation system using RFID for indoor environments," in *2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 2008, vol. 2, pp. 765–768.
- [16] T. Amemiya, J. Yamashita, K. Hirota, and M. Hirose, "Virtual leading blocks for the deaf-blind: A real-time way-finder by verbal-nonverbal hybrid interface and high-density RFID tag space," in *IEEE Virtual Reality 2004*, 2004, pp. 165–287.
- [17] J. Faria, S. Lopes, H. Fernandes, P. Martins, and J. Barroso, "Electronic white cane for blind people navigation assistance," in *2010 World Automation Congress*, 2010, pp. 1–7.
- [18] D. Ahmetovic, C. Gleason, C. Ruan, K. Kitani, H. Takagi, and C. Asakawa, "NavCog: a navigational cognitive assistant for the blind," in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2016, pp. 90–99.
- [19] K. Duarte, J. Cecilio, and P. Furtado, "Easily guiding of blind: Providing information and navigation-smartnav," in *International Wireless Internet Conference*, 2014, pp. 129–134.
- [20] G. Presti *et al.*, "Iterative Design of Sonification Techniques to Support People with Visual Impairments in Obstacle Avoidance," *ACM Trans. Access. Comput.*, vol. 14, no. 4, pp. 1–27, 2021.
- [21] S. Pundlik, M. Tomasi, M. Moharrer, A. R. Bowers, and G. Luo, "Preliminary Evaluation of a Wearable Camera-based Collision Warning Device for Blind Individuals," *Optom. Vis. Sci.*, vol. 95, no. 9, pp. 747–756, 2018.
- [22] A. Fiannaca, I. Apostolopoulous, and E. Folmer, "Headlock: a wearable navigation aid that helps blind cane users traverse large open spaces," in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, 2014, pp. 19–26.
- [23] D. Ahmetovic, C. Bernareggi, A. Gerino, and S. Mascetti, "Zebrarecognizer: Efficient and precise localization of pedestrian crossings," in *2014 22nd International Conference on Pattern Recognition*, 2014, pp. 2566–2571.
- [24] V. Ivanchenko, J. Coughlan, and H. Shen, "Crosswatch: a camera phone system for orienting visually impaired pedestrians at traffic intersections," in *International Conference on Computers for Handicapped Persons*, 2008, pp. 1122–1128.
- [25] S. Mascetti, D. Ahmetovic, A. Gerino, C. Bernareggi, M. Busso, and A. Rizzi, "Robust traffic lights detection on mobile devices for pedestrians with visual impairment," *Comput. Vis. Image Underst.*, vol. 148, pp. 123–135, 2016.

- [26] H. Shen and J. M. Coughlan, "Towards a real-time system for finding and reading signs for visually impaired users," in *International Conference on Computers for Handicapped Persons*, 2012, pp. 41–47.
- [27] B. J. Nguyen *et al.*, "Large-scale assessment of needs in low vision individuals using the Aira assistive technology," *Clin. Ophthalmol. (Auckland, NZ)*, vol. 13, p. 1853, 2019.
- [28] H. M. Htike, T. H. Margrain, Y.-K. Lai, and P. Eslambolchilar, "Ability of head-mounted display technology to improve mobility in people with low vision: A systematic review," *Transl. Vis. Sci. Technol.*, vol. 9, no. 10, p. 26, 2020.
- [29] N. A. Giudice, J. Z. Bakdash, and G. E. Legge, "Wayfinding with words: spatial learning and navigation using dynamically updated verbal descriptions," *Psychol. Res.*, vol. 71, no. 3, pp. 347–358, 2007.
- [30] N. Aziz, T. Stockman, and R. Stewart, "An investigation into customisable automatically generated auditory route overviews for pre-navigation," 2019.
- [31] N. Banovic, R. L. Franz, K. N. Truong, J. Mankoff, and A. K. Dey, "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*, 2013, pp. 1–8.
- [32] K. Papadopoulos *et al.*, "Environmental information required by individuals with visual impairments who use orientation and mobility aids to navigate campuses," *J. Vis. Impair. Blind.*, vol. 114, no. 4, pp. 263–276, 2020.
- [33] R. G. Golledge, J. R. Marston, J. M. Loomis, and R. L. Klatzky, "Stated preferences for components of a personal guidance system for nonvisual navigation," *J. Vis. Impair. Blind.*, vol. 98, no. 3, pp. 135–147, 2004.
- [34] F. Gaunet and X. Briffault, "Exploring the functional specifications of a localized wayfinding verbal aid for blind pedestrians: Simple and structured urban areas," *Human-Computer Interact.*, vol. 20, no. 3, pp. 267–314, 2005.
- [35] J. M. Loomis, R. G. Golledge, and R. L. Klatzky, "Navigation system for the blind: Auditory display modes and guidance," *Presence*, vol. 7, no. 2, pp. 193–203, 1998.
- [36] J. M. Loomis, J. R. Marston, R. G. Golledge, and R. L. Klatzky, "Personal guidance system for people with visual impairment: A comparison of spatial displays for route guidance," *J. Vis. Impair. Blind.*, vol. 99, no. 4, pp. 219–232, 2005.
- [37] M. A. Williams, A. Hurst, and S. K. Kane, "Pray before you step out," in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '13*, Oct. 2013, pp. 1–8, doi: 10.1145/2513383.2513449.
- [38] M. Brambring, "Mobility and orientation processes of the blind," in *Electronic spatial sensing for the blind*, Springer, 1985, pp. 493–508.
- [39] M. A. Williams, C. Galbraith, S. K. Kane, and A. Hurst, "'Just let the cane hit it': How the blind and sighted see navigation differently," 2014, doi: 10.1145/2661334.2661380.
- [40] J. K. F. Lloyd, R. C. Budge, K. J. Stafford, and S. J. La Grow, "A focus group discussion on using guide dogs," *Int. J. Orientat. Mobil.*, vol. 2, no. 1, pp. 52–64, 2009.
- [41] R. J.-T. Miner, "The experience of living with and using a dog guide," *RE view*, vol. 32, no. 4, p. 183, 2001.
- [42] "Tactile paving - Wikipedia." https://en.wikipedia.org/wiki/Tactile_paving (accessed Aug. 01, 2021).
- [43] E. R. Strelow and J. A. Brabyn, "Locomotion of the blind controlled by natural sound cues," *Perception*, vol. 11, no. 6, pp. 635–640, 1982.
- [44] A. Koutsoklenis and K. Papadopoulos, "Auditory cues used for wayfinding in urban environments by individuals with visual impairments," *J. Vis. Impair. Blind.*, vol. 105, no. 10, pp. 703–714, 2011.
- [45] N. A. Bradley and M. D. Dunlop, "Investigating context-aware clues to assist navigation for visually impaired people," 2002.
- [46] N. A. Bradley and M. D. Dunlop, "An experimental investigation into wayfinding directions for visually impaired people," *Pers. Ubiquitous Comput.*, vol. 9, no. 6, pp. 395–403, 2005.
- [47] M. K. Scheuerman, W. Easley, A. Abdolrahmani, A. Hurst, and S. Branham, "Learning the language: The importance of studying written directions in designing navigational technologies for the blind," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017, pp. 2922–2928.
- [48] E. Ohn-Bar, J. Guerreiro, K. Kitani, and C. Asakawa, "Variability in reactions to instructional guidance during smartphone-based assisted navigation of blind users," *Proc. ACM interactive, mobile, wearable ubiquitous Technol.*, vol. 2, no. 3, pp. 1–25, 2018.
- [49] M. Barouti and K. Papadopoulos, "Satisfaction of individuals with blindness from use of audio-tactile maps, tactile maps and walking experience as means for spatial knowledge of a city route," *TOJET Turkish Online J. Educ. Technol. Spec. Issue August*, pp. 447–452, 2015.
- [50] C. Bringhammar, G. Jansson, and G. Douglas, *The usefulness of a tactile map before and during travel without sight: A research report*. University of Birmingham, Research Centre for the Education of the Visually ..., 1997.
- [51] V. R. Schinazi, *Representing space: the development, content and accuracy of mental representations by the blind and visually impaired*. University of London, University College London (United Kingdom), 2007.

- [52] R. Yaagoubi, G. Edwards, and T. Badard, "Standards and Spatial Data Infrastructures to help the navigation of blind pedestrian in urban areas," *Urban Reg. Data Manag. UDMS 2009 Annu.*, pp. 139–150, 2009.
- [53] J. P. Bigham *et al.*, "VizWiz: nearly real-time answers to visual questions," in *Proceedings of the 23rd annual ACM symposium on User interface software and technology - UIST '10*, 2010, p. 333, doi: 10.1145/1866029.1866080.
- [54] N. D. Troncoso Aldas, S. Lee, C. Lee, M. B. Rosson, J. M. Carroll, and V. Narayanan, "AIGuide: An Augmented Reality Hand Guidance Application for People with Visual Impairments," in *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, 2020, pp. 1–13.
- [55] K. R. May, B. J. Tomlinson, X. Ma, P. Roberts, and B. N. Walker, "Spotlights and Soundscapes: On the Design of Mixed Reality Auditory Environments for Persons with Visual Impairment," *ACM Trans. Access. Comput.*, vol. 13, no. 2, pp. 1–47, 2020.
- [56] M. Eckert, M. Blex, and C. M. Friedrich, "Object detection featuring 3D audio localization for Microsoft HoloLens," in *Proc. 11th Int. Joint Conf. on Biomedical Engineering Systems and Technologies*, 2018, vol. 5, pp. 555–561.
- [57] H. O. Posten, "Robustness of the two-sample t-test," in *Robustness of statistical methods and nonparametric statistics*, Springer, 1984, pp. 92–99.
- [58] C. Thinus-Blanc and F. Gaunet, "Representation of space in blind persons: vision as a spatial sense?," *Psychol. Bull.*, vol. 121, no. 1, p. 20, 1997.
- [59] M. Fortin *et al.*, "Wayfinding in the blind: larger hippocampal volume and supranormal spatial navigation," *Brain*, vol. 131, no. 11, pp. 2995–3005, 2008.
- [60] C. Pigeon, T. Li, F. Moreau, G. Pradel, and C. Marin-Lamellet, "Cognitive load of walking in people who are blind: Subjective and objective measures for assessment," *Gait Posture*, vol. 67, pp. 43–49, 2019.
- [61] K. M. Rand, S. H. Creem-Regehr, and W. B. Thompson, "Spatial learning while navigating with severely degraded viewing: The role of attention and mobility monitoring," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 41, no. 3, p. 649, 2015.
- [62] E. Koustriava and K. Papadopoulos, "Mental rotation ability of individuals with visual impairments," *J. Vis. Impair. Blind.*, vol. 104, no. 9, pp. 570–575, 2010.
- [63] E. Koustriava and K. Papadopoulos, "Are there relationships among different spatial skills of individuals with blindness?," *Res. Dev. Disabil.*, vol. 33, no. 6, pp. 2164–2176, 2012.
- [64] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: A survey," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 40, no. 1, pp. 25–35, 2010, doi: 10.1109/TSMCC.2009.2021255.
- [65] M. Martinez, A. Roitberg, D. Koester, R. Stiefelhagen, and B. Schauerte, "Using technology developed for autonomous cars to help navigate blind people," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 1424–1432.
- [66] B. Li, X. Zhang, J. P. Muñoz, J. Xiao, X. Rong, and Y. Tian, "Assisting blind people to avoid obstacles: An wearable obstacle stereo feedback system based on 3D detection," in *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2015, pp. 2307–2311.
- [67] M. Cotzin and K. M. Dallenbach, "'Facial vision': The role of pitch and loudness in the perception of obstacles by the blind," *Am. J. Psychol.*, pp. 485–515, 1950.
- [68] C. E. Rice, S. H. Feinstein, and R. J. Schusterman, "Echo-detection ability of the blind: size and distance factors.," *J. Exp. Psychol.*, vol. 70, no. 3, p. 246, 1965.
- [69] W. N. Kellogg, "Sonar system of the blind," *Science (80-)*, vol. 137, no. 3528, pp. 399–404, 1962.
- [70] X. Chen, F. Wang, E. Fernandez, and P. R. Roelfsema, "Shape perception via a high-channel-count neuroprosthesis in monkey visual cortex," *Science (80-)*, vol. 370, no. 6521, pp. 1191–1196, 2020.
- [71] E. Fernandez, "Development of visual Neuroprostheses: trends and challenges," *Bioelectron. Med.*, vol. 4, no. 1, pp. 1–8, 2018.
- [72] J. de R. van Steveninck, U. Guclu, R. J. A. van Wezel, and M. A. J. van Gerven, "End-to-end optimization of prosthetic vision," *bioRxiv*, 2020.
- [73] V. Vergnieux, M. J. Macé, and C. Jouffrais, "Simplification of visual rendering in Simulated Prosthetic Vision facilitates navigation," *Artif. Organs*, vol. 41, no. 9, pp. 852–861, 2017.
- [74] A. Lozano *et al.*, "Neurolight: A Deep Learning Neural Interface for Cortical Visual Prostheses.," *Int. J. Neural Syst.*, p. 2050045, 2020.