Assessing Mobility of Blind and Low-Vision Individuals through a Portable Virtual Reality System and a Comprehensive Questionnaire

Citation for the original published paper (version of record):
Isaksson-Daun, J., Jansson, T., Nilsson, J. (2024). Assessing Mobility of Blind and Low-Vision Individuals through a Portable Virtual Reality
System and a Comprehensive Questionnaire. IEEE Access, 12: 146089-146106.
http://dx.doi.org/10.1109/ACCESS.2024.3471177

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 3 September 2024, accepted 27 September 2024, date of publication 1 October 2024, date of current version 15 October 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3471177

METHODS

Assessing Mobility of Blind and Low-Vision Individuals Through a Portable Virtual Reality System and a Comprehensive Questionnaire

JOHAN ISAKSSON-DAUN[1,2], TOMAS JANSSON[2,3,4], (Member, IEEE), AND JOHAN NILSSON[2], (Member, IEEE)

[1]Division Design and Human Factors, Department of Industrial and Materials Science, Chalmers University of Technology, 412 96 Gothenburg, Sweden
[2]Department of Biomedical Engineering, Lund University, 223 63 Lund, Sweden
[3]Department of Clinical Sciences Lund/Biomedical Engineering, Lund University, 221 84 Lund, Sweden
[4]Clinical Engineering Skåne, Skåne University Hospital, 221 85 Lund, Sweden

Corresponding author: Johan Isaksson-Daun (johan.isaksson-daun@chalmers.se)

This research was funded by the Promobilia Foundation Grant F-17510, and hardware procurement was also funded by the Sten K Johnson Foundation Grant 20190063.

ABSTRACT Blind or low-vision (BLV) individuals often have reduced independent mobility, yet new aids fails in increasing it, are not adopted enough, or both. A major cause is a severe deficiency in how mobility aids are assessed in the field: there are no established methods or measures and those used often have poor relevancy, insight affordances, and reproducibility; probing how actual BLV participants regard a proposed aid and how they compare to current aids is rare; and crucially, tests feature too few BLV participants. In this work two tools are introduced to alleviate this: a portable, large-scale-exploration, virtual reality (VR) system; and a comprehensive, aid-agnostic questionnaire focused on BLV mobility. The questionnaire has been validated once with eight orientation and mobility experts and six BLV respondents. Further, both it and the VR system have been applied in aid assessment with 19 BLV participants in a separate study. The VR system is to our knowledge the first in the field designed for portable evaluation, helping considerably in recruiting adequate numbers of BLV participants, for instance by allowing for testing in participants' homes; while also supporting reproducible and motivated tests and analyses. The questionnaire provides a systematic method to investigate respondents' views of numerous important facets of a proposed mobility aid, and how they relate to other aids. These tools should assist in achieving a widely adopted aid that meaningfully improves its users' mobility.

INDEX TERMS Blindness, electronic travel aids, low vision, mobility aids, patient-reported outcome measures, sensory substitution, sensory supplementation, virtual environments, virtual reality, visual impairment.

I. MAIN INTRODUCTION

There is an estimated 43.3 million blind, and 34.8 million severely vision impaired, people in the world [1]. Reduced eyesight can dramatically curtail mobility, which in turn decreases independence and quality of life of many blind or low-vision (BLV) individuals [2]. Further, primary aids of today, predominantly long canes and guide dogs, are inadequate in resolving this lack of mobility [3], [4], [5], [6], [7].

This has however not gone unnoticed, with numerous attempts in designing better or complementary electronic travel aids (ETAs) made for at least 70 years [8], resulting in no lack of proposed devices. For a historical outline of ETAs see e.g. [9], and for recent examples see the systematic reviews from the last five years listed in Table 1.

Yet, despite this abundance of attempts, and continuous technological advancement, achieving a widely accepted ETA that significantly increases the mobility of its users remains vexing. Inaccessibility of existing ETAs in the form of e.g. cost, discoverability, and training prospects plays an important role in this [10], [11], [12]; however, a major

The associate editor coordinating the review of this manuscript and approving it for publication was M. Sabarimalai Manikandan.

TABLE 1. Recent systematic reviews with various focus on electronic travel aids (ETAs) for blind and low-vision (BLV) individuals.

Year	Reference	Review focus (as per content organization in article)
2023	Xu *et al.* [13]	Wearable obstacle avoidance ETAs ($n = 89$): study selection, study description, description of devices, characteristics of trials.
2022	Budrionis *et al.* [20]	Smartphone-based ETAs ($n = 15$): purpose of the solution and functionality; input/output; data processing and performance; algorithms used in data processing; evaluation of the proposed solutions.
2022	Walle *et al.* [22]	Deep- or machine-learning-powered ETAs ($n = 33$): human-machine interface; artificial intelligence techniques; testing methods.
2021	Fernando *et al.* [28]	Indoor ETAs with route planning ($n = 131$): route optimizing factors; addressing user variations; procedures and algorithms; building-data modeling.
2021	Khan *et al.* [16]	ETAs ($n = 131$): approaches; technologies/tools; mechanisms/applications; applicability and reliability parameters.
2021	Parker *et al.* [21]	ETAs used in real-life wayfinding studies ($n = 35$): participants with visual impairments; technologies and devices; environments, settings, and wayfinding tasks; institutional and financial support; interconnections.
2021	Dos Santos *et al.* [15]	Wearable ETAs ($n = 61$): technologies; feedback interfaces; user evaluation.
2021	Zhang *et al.* [17]	Meta review ($n = 550$): publication outputs; most influential journals; most active and influential authors; key references; primary topics and research hotspots.
2020	Chang *et al.* [18]	Orientation and mobility (O&M) outcome measures in research ($n = 32$): study objectives and characteristics; efficiency, accuracy, and self-reported outcome measures; quality assessment; application in clinical and functional O&M.
2020	Plikynas *et al.* [19]	Indoor ETAs ($n = 27$) (see also: [29]): non-vision-based systems; vision-based-systems; sensors. Computer vision-based smartphone-centered ETAs ($n = 15$): [unsectioned results/discussion].
2019	Žvironas *et al.* [29]	Indoor ETAs ($n = 27$) (see also: [19]): navigation technologies; sensors; computational devices; feedback type.
2018	Araújo *et al.* [30]	ETAs ($n = 16$): [unsectioned results/discussion].
2018	Khan *et al.* [31]	Cane-based ETAs ($n = 36$): operating environment; sensors; system functionality; positioning of sensory unit; sensor sensing range; operation mode; computational device; localization technology; user-system interaction.

underlying reason is likely that no ETA has, to our knowledge and corroborated in [13], shown that it actually improves its users' independent mobility. In [9] we argued that this might be due to that, while the ever-growing body of works offering design guidelines and best practices for ETA development and testing should steadily yield increasingly beneficial grounds for ETA designers, the field routinely fails to utilize them. Our attempt to reconcile this gave rise to the Desire of Use (DoU) model, and its application toward mobility of BLV individuals (DoU-MoB).

Another concern in the field is the severe lack of established methods and measures, as we observed in [9] and [14], and is corroborated in recent systematic reviews [13], [15], [16], [17], [18], [19].

Beside issues in adoption-rate and reproducibility, in both [9] and [14] we leveraged DoU-MoB to argue that measures that frequently are used in the field, e.g. percentage of preferred walking speed (PPWS) or collision tallies, are not sufficient for ETA development due to their high abstraction level. For example, while a high PPWS might indicate great functionality, the more frequent result of a low PPWS yields little information of where an ETA needs improvement. Thus, measures of lower abstractions are needed, as to efficiently target possible issues of a proposed system.

There is also a consistent absence in gauging the end-users' perception of tested ETAs. In their separate reviews of various ETAs, Budrionis et al. noted a low rate of user experience evaluation, and limited scope of such when actually performed [19], [20]. They also found a discrepancy between what researchers evaluate and what users find important [20]. Dos Santos et al. calls for increased employment of qualitative measures of safety following their review of wearable ETAs [15]. Khan et al. seemingly found no assessment of BLV participants' acceptance of proposed devices in their review of ETAs [16]. Zhang et al. concluded that recognized guidelines and benchmarks for usability and functionality (and affordability) are lacking in their meta review of the field [17].

Which measures are feasible depend on the testing methods—yet another aspect of the field found wanting. In [14], we noted poor reproducibility of proposed methods, and that the number of times any specific test procedure has seen repeated use is either low or zero. This is both notable and jarring since indoor mobility courses both are quite popular—it is possible that most tests of novel ETAs are evaluated with such [18]—and that their main advantage over real-life tests are likely their reproducibility affordances. In [18], Chang et al. states that real-world testing both is more generalizable to real life, and that mobility courses are inadequate in testing the full range of needed orientation and mobility (O&M) skills. Parker et al. are proponents of real-world testing as they argue that the uncontrolled for, and frequent, travel distractions they provide are what makes wayfinding complex, and thus better affords evaluation of ETAs [21].

Lastly, but critically, there is a pervasive issue in recruiting enough actual BLV participants for user tests [13], [15], [19], [20], [22], [23], [24], [25]. In many works, possibly predominantly, novel ETAs are only tested with blindfolded sighted people, and when BLV individuals do partake, it tends to be in a too limited extent. In [23], Horton et al. states that five to ten users for usability testing is generally accepted, while statistical tests require at the very least more than ten users. They point out that a reason behind lacking involvement of BLV individuals is the difficulty in recruiting such participants, a view we personally can attest to after [14].

A. AIM OF PAPER

From the above three themes regarding the field emerge: 1) frequently used measures offer little learning, they should focus more on user perception, and they are difficult to compare between aids; 2) there is a lack of standardized, agreed-upon, and well-motivated evaluation methods; and 3) there is a scarcity in employing a sufficient number of blind and low-vision participants.

In this paper, we risk further introduction of new evaluation methods and measures to the field, and argue that these considerable issues can be alleviated by utilizing: 1) a novel, portable virtual reality (VR) system in Section II; and 2) a new patient-reported outcome measure (PROM)[1] in Section III. The paper is concluded with a small note of how these tools have been deployed in the field in an initial study in Section IV, and a main conclusion in Section V.

Preliminary versions of this work has been reported in [26] as well as in [27].

II. PROPOSED PORTABLE VIRTUAL REALITY SYSTEM (PARROT-VR)

A. INTRODUCTION (PARROT-VR)

To address many of the problems raised in Section I, in this section we introduce a novel VR system, Parrot-VR (named after its properties, see Section II-C) to aid in ETA assessment. First, the design considerations of a VR system to evaluate ETAs that takes those problems into account are examined. Next previous works that potentially could conform to these design considerations are outlined. Following this is the introduction of the VR system, as well as a summary and short discussion.

An important side-note is that, for ETA developers, a VR system can be doubly helpful. Both in evaluation, which can illustrate the benefits and shortcomings of a current design, as well as for increasing the rate of design iterations of a proposed ETA. For example, if there is a known situation that an ETA needs improvement in, recreating it in VR can make internal tests quite time efficient, in contrast to revisiting it repeatedly in real-life.

1) PREVIOUS VR SYSTEMS FOR BLV INDIVIDUALS (PARROT-VR)

There are a variety of reviews considering VR systems for BLV individuals in general; for an introduction, we refer to the recent example of Kreimeier and Götzelmann in [32], who also provides an accompanying outline of other reviews. They also introduce a taxonomy of existing VR systems, using the distinguishing aspect of virtual environment (VE) scale. They categorize systems into small, medium, or large scale, and provide common parameters for them, e.g. exploration interface, absolute or relative positioning, and egocentric or exocentric (aka allocentric) perspective. In such terms, a VR system conforming to the considerations above would have to be large scale and egocentric.

Much of the previous work using VR in the field is related to either VE exploration and O&M training, and not ETA assessment, which is the scope of this work. Thus, in Table 2, such works that feature large-scale, egocentric VR systems actually used for assessing ETAs are presented, along with their respective positioning methods and participants. Clearly most work favor absolute rotation, predominately through head-tracking. In contrast, translation (movement or how to walk) is most often relative, mostly by keyboard or joystick. It is also apparent that most works here also struggle with inclusion of BLV participants; almost half have none, and only three have more than ten. To our knowledge, there is no previous work attempting to alleviate this by using VR to test in the vicinity of individual BLV participants.

2) DESIGN CONSIDERATIONS OF VR FOR ETA ASSESSMENT (PARROT-VR)

First, ETA evaluation should benefit substantially from being performed in relevant, large-scale environments, in contrast to obstacle courses. The drawbacks of this approach is reduced reproducibility and safety. If a specific location is chosen, another research team will in general not be able to perform the same tests. If instead locations close to individual BLV participants are chosen, the locations will naturally be different between users, and it becomes difficult to ensure that they all exhibit the necessary features for adequate evaluation, as well as to aggregate results between tests. Regardless of location, large-scale urban settings are vulnerable to uncontrollable conditions, e.g. weather and traffic, which also presents major risks to the participants and thus requires significant safety measures or previous validation of the proposed ETA. However, a VE is inherently reproducible in that regard, as well as safe, effectively avoiding both of these considerable issues.

Second, in order to recruit sufficient participants from the target group, it should be greatly advantageous to test near individual participants. Thus, the VR system should be portable in respect to size, weight, setup conditions, and setup time, so that it can be effectively deployed in participants' homes or elsewhere.

---

[1] Note that "patient-reported outcome measure" (PROM), is an established term in many fields, ophthalmology included. Identifying and referring to a tool as such can thus aid in the communication between various stakeholders, i.e. ETA designers, ophthalmologists and O&M experts; as well as should help in discoverability and recognizability. Regrettably, the term could be perceived as alienating, as a respondent may not identify themselves as patients. At this time, it is our opinion that utilizing the established term is still worthwhile due to the aforementioned reasons, especially in a research or medical setting, but terminology such as "user-reported…" or "participant-reported…" could well be used instead.

TABLE 2. **ETA assessments using large-scale, egocentric virtual reality (VR) systems. Publications in the same segment use the same VR system. They evaluate the ETAs summarized at the bottom of each segment; the summary is based on the most recent work in the segment. *Table originally produced for this work, but published first with slight differences in [27].***

Year	Reference	Participants	Translation, method	Rotation, method
2022	Ricci *et al.* [33]	1 sighted*	Absolute, head-tracking	Relative, joystick
Computer vision obstacle detection with multidimensional abdominal haptic feedback.				
2021	Real and Araujo [34]	23 sighted	Absolute, head-tracking	Absolute, open space augmented reality (AR)
2020	Real and Araujo [35]	16 sighted		
The vOICe: head-mounted visuo-auditory sensory supplementation/substitution device (SSD) through greyscale images, see [36]; PVAS/VAS: head-mounted visuo-auditory SSD through distance; hand-held visuo-tactile SSD through distance.				
2020	Lupu *et al.* [37]	15 BLV	Absolute, head-tracking	Relative, keyboard/joystick
2017	Dascalu *et al.* [38]	7 sighted, 8 BLV		
2017	Moldoveanu *et al.* [39]	4 BLV		
Sound of Vision: head-mounted visuo-auditory/tactile SSD through distance and with multidimensional abdominal haptic feedback.				
2019	Maidenbaum and Amedi [40]	40 sighted, 8 BLV	Relative, keyboard	Relative, keyboard
EyeMusic: head-mounted visuo-auditory SSD through RGB images.				
2018	Massiceti *et al.* [41]	18 sighted	Absolute, head-tracking	Absolute, open space AR
Head-mounted visuo-auditory SSD through distance.				
2016	Levy-Tzedek *et al.* [42]	29 sighted	Relative, keyboard	Relative, keyboard
2017	Chebat *et al.* [43]	36 sighted, 20 BLV		
2015	Chebat *et al.* [44]	36 sighted, 20 BLV		
2014	Maidenbaum *et al.* [45]	7 sighted		
2013	Maidenbaum *et al.* [46]	20 sighted, 3 BLV		
EyeCane: hand-held distance perception with auditory feedback.				
2016	Maidenbaum *et al.* [47]	29 sighted, 9 BLV	Relative, keyboard	Relative, keyboard
2015	Maidenbaum *et al.* [48]	10 sighted, 5 BLV		
EyeMusic: head-mounted visuo-auditory SSD through RGB images.				
2014	Sanz textitet al. [49]	13 sighted, 15 BLV	Absolute, head-tracking	Relative, verbal command
Head-mounted visuo-auditory SSD through distance.				
2013	Lun Khoo *et al.* [50]	18 sighted*	Absolute, mouse	Relative, joystick*
Arm-mounted distance perception with haptic feedback on arms.				
2012	Bujacz *et al.* [51]	10 sighted	Absolute, head-tracking	Relative, keyboard*
Naviton: Head-mounted visuo-auditory SSD through distance.				

---

*Not clear from publication but most likely.

B. MATERIALS AND METHODS (PARROT-VR)

Our proposed VR system Parrot-VR is designed to adhere as closely as possible to the previously outlined design considerations—while maintaining a portable format. Below, the design choices to make this possible are presented, as are technical details regarding the currently integrated ETA, implementation and operation, and the VR avatar (i.e. the user-controlled character).

1) DESIGN CHOICES (PARROT-VR)

Large-scale positioning together with portability immediately restricts the possible design space of a VR system. Further, to keep the system as intuitive as possible, how a user controls the VR avatar is essential. This control can be split into two parts, translation (movement) and rotation. It is likely that mirroring a user's real-life movements 1:1 should be both quite natural, and yield a good sense of orientation by leveraging their proprioceptive system. This implies that an absolute translation and rotation is a preferable approach—in contrast to relative dittos. Relative translation could for instance be a button press corresponding to a footstep, and relative rotation another button press to do a 15° turn.

Arguably, a large-scale VR system with absolute translation and rotation can be achieved; by employing AR technology in a large open space, a user might explore a sizable VE as evidenced by e.g. [35], [41], [52]. This is quite an interesting approach, though it has drawbacks in that the VE still will be constrained by the available space, which restricts the possible training or evaluation scenarios; as well as that a suitable open space is needed, preferably near participants and independent of weather and time of day—in effect a portability issue.

Beside such AR-solutions, large-scale translation virtually has to be solved relatively, either with an omnidirectional treadmill, or through some form of abstraction e.g. keyboard, joystick, or video game controller. An omnidirectional treadmill might be more intuitive, and should at least give a sense of how far, and in what direction, a user has walked without other feedback, which a joystick or button-press do not. A limited sense of absolute translation is probably detrimental if macro-navigation ability is to be assessed; however, it is likely of less consequence for micro-navigation tasks.

Fortunately, the arguably more important—regarding both micro- and macro-navigation—absolute sense of direction can be solved with absolute rotation via head-tracking. It might be advantageous to track the user's head and body separately, as they might want to look in another direction than in which they are walking; however this would bring considerable extra complexity through more required sensors.

The portability aspect entails that VR systems requiring desktop computers are prohibitive, as are omnidirectional treadmills, due to size, setup time and weight.

While modern, off-the-shelf, stand-alone VR headsets could feasibly be employed, we saw both limits and opportunities with our earlier proposed ETA, and currently our main object of evaluation, Audomni [14], (see Section II-B2 for a brief overview). Time considerations, along with uncertainties in how involved it would be to either port the Audomni feedback to an existing VR system, or transfer the necessary input from it to Audomni, made us hesitant; especially as the lowest possible latency is desired. Meanwhile, as a head-mounted display should not be critical for ETA evaluation targeting both BLV users, the embedded computer and the inertial measurement unit (IMU) already present in the Audomni prototype provided almost all necessary hardware for a functional VR system with absolute rotation. Further, avoiding a VR system with a head-mounted display, which are often bulky and heavy, could provide for a more reliable evaluation, and greater portability. Lastly, utilizing the existing prototype hardware enables identical feedback as the ETA, and provides the exciting prospect of VR training for users—needing only the ETA.

The only part missing is an input method for large-scale translation. We chose a controller for a popular video game console, taking advantage of the extensive work the company likely has put into reaching a durable and ergonomic product. As such controllers also tend to feature haptic feedback capabilities through vibration, this could also be exploited to give a sense of translational movement as well as to convey any collisions.

For evaluation purposes there are some auxiliary considerations. The test administrator needs visual feedback of the VR avatar movement, and ideally also the ETA feedback that the user receives. Otherwise, it will be quite difficult to ensure that all systems are working as expected, or what a user might need help with if they get confused or stuck. Further, the avatar movement and the feedback should be recorded to allow for subsequent analyses. For the visuals, the avatar data is readable from an external laptop, along with software that can display a copy of the VE and avatar in real-time. As Audomni strictly uses auditory feedback, this is mirrored through an external audio card and outputted in identical headphones as the user's.

TABLE 3. **Hardware of the Parrot-VR system.** ***Table originally produced for this work, but published first with slight differences in [27].***

System proper	
Embedded computer	Jetson Nano Developer Kit, Nvidia.
Inertial measurement unit (IMU)	BNO055 Absolute Orientation Sensor, Adafruit.
Controller	Dualshock 4 V2 Controller, Sony.
Headphones	Sportz Titanium, Aftershokz.
Audio interface	Pmod I2S, Digilent.
Auxiliary	
Laptop	MacBook Pro 13-inch 2017, Apple.
Headphones	Sportz Titanium, Aftershokz.
Audio USB interface	Volt 2, Universal Audio.
Audio recording device	iPhone 6s, Apple.

FIGURE 1. **VR system, Parrot-VR, system hardware.** ***Clockwise from left*:** **bone-conductive headphones; embedded computer and audio interface; headband with IMU; controller.** ***Figure and caption originally produced for this work, but published first with slight differences in [27].***

2) INTEGRATION OF THE AUDOMNI ETA (PARROT-VR)

As we are interested in the evaluation and continued development of our previously proposed ETA, Audomni [14], currently a virtual version of it is integrated in the VR system. Audomni is a visuo-auditory sensory supplementation device, which aims to perform sonification of visual information needed for safe urban mobility. While its feedback has seen fair modification since this previous publication, the basic concept remains the same; in brief,

FIGURE 2. **VR system, Parrot-VR, test setup: *1*) IMU; *2*) bone-conductive headphones; *3*) controller; and *4*) embedded computer. Auxiliary equipment: *5*) audio interface; *6*) laptop; *7*) bone-conductive headphones; and *8*) audio recording device. Note that the embedded computer hosts the software of both Parrot-VR and the Audomni ETA. *Figure and caption originally produced for this work, but published first with slight differences in [27].***

it uses a head-mounted depth camera as input; translates horizontal positions to stereo pan, vertical to pitch, and distance to volume; and finally provides the audio feedback to the user through bone-conductive headphones. Thus, a VR system which integrates it should provide depth images from the VR avatar head, and support audio output.

3) IMPLEMENTATION, WORKFLOW, AND OPERATION (PARROT-VR)

The Parrot-VR hardware is presented in Table 3 and Fig. 1. A typical test setup is shown in Fig. 2, and software details in Table 4. Examples of VEs and test procedures for use with Parrot-VR is beyond the scope of this article, but can be found in [53].

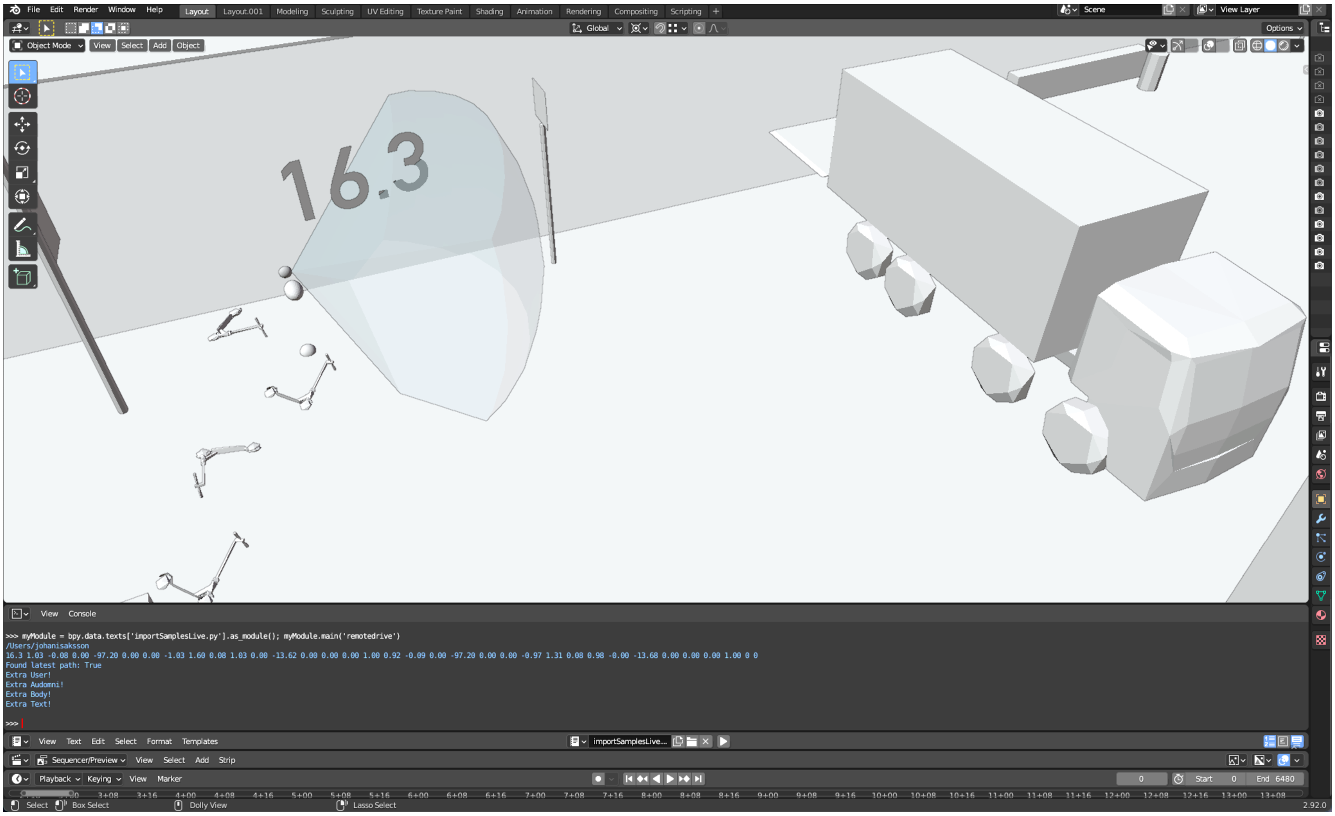
The VEs to be explored are designed in the 3D-creation suite Blender [71]. The Blender add-on Blender-OSM [72], is used to import an accurate 2D-map overlay of a desired urban area, including any user-generated 3D geometries, from OpenStreetMap [73], to Blender. Subsequently using the map overlay, any user-generated geometries, satellite images from Google Maps [74], and Apple Maps [75], as well as physical visits, additional geometries and relevant details are added manually in Blender. Resulting VE examples are included in full in [53], and are partly featured in this work in Figs. 3 and 5.

TABLE 4. **Software of the Parrot-VR system. *Table originally produced for this work, but published first with slight differences in [27].***

Programming and software environment	
Operating system	Ubuntu [54][1]
Programming language	C++ [55]
Compiler	GCC [56]
External libraries	
Inertial measurement unit	BNO055 sensor driver [57]
Window and controller	GLFW [58][2]
Graphics and 3D	OpenGL [59]
Math	GLM [60]
3D model import	Assimp [61]
OpenGL instancing	Glad [62]
Image loading	Stb_image [63]
Electronic travel aid	Audomni [14][3]
Audomni compatability	OpenCV 4.4 [64]
Additional code	
OpenGL, additional	Code by J. de Vries [65]–[69][3]

[1] Slightly modified kernel to allow for controller haptic feedback. [2] Slightly modified source-code to allow for controller haptic feedback as per [70]. [3] Modified.

The 3D geometries of a VE are enough for the VR system to produce feedback for a virtual ETA, and for the VR avatar to walk and look around. However, to restrict avatar movements, e.g. so that it cannot walk through walls,



```
>>> myModule = bpy.data.texts['importSamplesLive.py'].as_module(); myModule.main('remotedrive')
/Users/johanisaksson
16.3 1.03 -0.08 0.00 -97.20 0.00 0.00 -1.03 1.60 0.08 1.03 0.00 -13.62 0.00 0.00 0.00 1.00 0.92 -0.09 0.00 -97.20 0.00 0.00 -0.97 1.31 0.08 0.98 -0.00 -13.68 0.00 0.00 0.00 1.00 0 0
Found latest path: True
Extra User!
Extra Audomni!
Extra Body!
Extra Text!

>>> 
```

FIGURE 3. Screen capture of a live capture of a VR session in Blender. This is what a test administrator might see during a test session. The translucent cone denote the field-of-view (FoV) of the virtual Audomni ETA, and it is attached to the user avatar (which only consists of a head and upper body, see Fig. 4). "16.3" is a text field which follows the avatar, and can be made to display various data points live; here it is simply the current duration of the session in seconds. Note that the camera can be moved during a session to keep the user in view.

collision information of the geometries are needed. This also allows for collision registration as an outcome measure. The information is coded as invisible shapes around relevant objects, such as walls and obstacles, as well as curbs or platform edges if moving over them also should be treated as errors, which is often desirable. The shapes are added manually in Blender through custom helper-scripts written in the Python [76], programming language.

The VEs, collision information, and VR avatar positionings are exported to a USB stick, which is put into the VR system.

The VR system might be started directly on the VR embedded computer, and for a training setting this could be most convenient; however, for a test or development setting, the preferable approach is to bring it up through a secure shell protocol (SSH) connection from a laptop. Since VR can be quite computationally demanding, and latency should be kept as low as possible, the embedded computer should ideally only perform VR-related tasks. In a test setting, visual feedback is all but necessary for the test administrator, and so an SSH setup makes such possible without taxing the embedded computer additionally. An SSH filesystem (SSHFS) can also be used to access real-time data on the VR system from the laptop. This also holds for a development setting, where an integrated development environment can draw unnecessary resources from the embedded computer of the ETA. Furthermore, for a portable test setting, it is more practical to bring and setup a laptop with an integrated keyboard, trackpad, and display, than separate hardware.

During initialization, various settings are read from a text file, including for a VR camera mimicking the field-of-view (FoV), resolution, and noise properties of a real image sensor; which VE to load; the positioning of the sensor on the VR avatar, see Fig. 4; and walking speed.

During operation, the VR camera continually produces an RGB and a grayscale image, as well as a z-buffer (or depth buffer). The z-buffer contains normalized depth from the camera plane for each pixel; by multiplying the inverse of the projection matrix of the VR camera—which is known—with the normalized screen coordinates (including the z-values) of all pixels in the z-buffer, a proper distance (or depth) image giving the euclidean distance from all pixels in the VR camera to their respective sources in the VE can be produced. Any combination of this RGB, grayscale, and depth image can be returned by the VR system. To simulate a real sensor, noise corresponding to its characteristics is added to relevant images. After this, they are fed into the virtual ETA, at a rate of around 40 Hz, whereafter the ETA can produce its corresponding feedback.

Before head-tracking is commenced, the IMU needs to be internally and externally calibrated. The internal calibration is needed for an absolute rotation and is performed continuously on the IMU during movement. If the calibration gets too poor the controller starts to vibrate softly. In such cases the user is to pause their virtual movement and slowly nod or tilt their head until it is adequately calibrated again.The frequency of these occurrences varies on how the user naturally moves, as well as the room the VR system is situated in (as the IMU use Earth's magnetic field for this calibration), but it seems that it seldom happens more than once or twice per VE in test sessions.

The external IMU calibration maps the absolute rotation of the IMU on the user to the VE space. For this, the IMU rotation data of the user when they (with their head) look forward, slightly to the left, and slightly upward is registered. Through this, the orthogonal rotation axes in real-life can be estimated, which are subsequently mapped to the VE. A considerable benefit of this approach is that it removes any dependency of the exact positioning of the IMU on a user's head; as long as it sits still in relation to the head during movement, it will provide adequate head-tracking.

When the head-tracking is running, the IMU rotation is applied to the VR avatar in a split manner; yaw (horizontal, or rotation around the upward axis) is applied to its body, while pitch and roll (rotation around the left and forward axes, respectively) are applied to its head. There are three reasons for this: 1) we want the walking direction to always correspond to the direction of the user's body; 2) we want to preserve full possibility of looking around; and 3) both the user's head and body cannot simultaneously be tracked with the single IMU. This means that the avatar is always walking in the direction of the user's head, and thus they are advised to always turn their whole body whenever they look around, to keep their expected and actual walking direction in line.

Users only need to use the up and down buttons of the controller; up walks the avatar forward with a set speed, and down backward. However, the system also supports translation as well as yaw (horizontal) rotation with the two joysticks, where speed depends on joystick tilt angle. There is also one controller shoulder button that can be used to interact with the virtual ETA, for instance it could be used to signal the ETA to restart its feedback if necessary.

Other buttons on the controller can be used for debugging purposes, for instance to display the different available images in various stages of processing through a SSH terminal window; perform the aforementioned external calibration; print various variables; and shutdown the current VR session.

When walking, the positioning of the avatar is updated accordingly, while a short vibration is produced for every step traversed. If any collisions are detected, the corresponding vibration feedback from the controller is triggered.

During a session, the positioning of both the VR avatar body and head is continuously written to a file at a rate of 5 Hz, along with timestamps and collision flags. Using another custom Python script in Blender running on a laptop connected through SSH and SSHFS, it is possible to read this file, and update the avatar in real-time, in the very same file that the VE was produced in. This provides a test administrator with a live overview, who can thus then ask pertinent questions or otherwise intervene when necessary, see Fig. 3. Any collision flags are also utilized to modify the sensor FoV color to clearly communicate such important events.

After a session, a similar custom Python script can be used in Blender to import all the samples of a recorded file, or of several such files, into the VE file. Then one, or multiple overlaid, sessions can be played back after-the-fact. The collision flags of samples are used identically as in the live playback setting. If any sound recording is available, for instance a recording from the ETA, or of speech from the user and test administrator, these can also be included and synched with the avatar movements. Then, a video file can be produced of this entire collection of test sessions of a given VE, including audio recordings, providing further opportunities for analysis. Such analysis could be detailing the circumstances of collisions or other noteworthy events, see [53]. The custom script can also import such details, and then automatically add color-coded text boxes to the video, with the details appearing in time with the corresponding events; whereafter a new video file including those can be produced. See Fig. 5 for an example still from such a video.

4) THE VR AVATAR (PARROT-VR)

The feedback of an ETA can be affected both by the sensor positioning on a user, and the user's body itself; this needs to be taken into account for the VR avatar. For a head-mounted sensor, both its orientation and position will change with head-movement according to its positioning relative the head pivot point. Further, a user's body shape might interfere with the sensor FoV, for instance might the chest or abdomen occlude a user's immediate ground.

We opt for a standard avatar for all users, as otherwise would be overly time-consuming, possibly sensitive, and of little benefit. However, to make a potential aid useful for the widest possible user distribution, we model the avatar after a hypothetical user whose anthropometric measures would be at the most disadvantageous 95th percentile for sensor occlusion, see Table 5. The anthropometric data is taken from [77], with some additional ad hoc measurements for data that could not be found. The avatar currently consists of only a head and a chest, as more would have been redundant for occlusion regarding a head-mounted sensor placement; and since the avatar model is unrelated to how collisions are identified, which currently is dictated purely by horizontal distance (e.g. within 20 cm) between the avatar body pivot point and any collision objects. The resulting avatar is shown in Fig. 4.

FIGURE 4. VR avatar in default positioning, from the side and front. White spheres are head (*1*) and body (*3*), and red their pivot points (*2* and *4*), blue cone (*5*) is the sensor FoV (omitted on the right). The body can affect the sound feedback by occluding the FoV, which happens for many possible head rotations. All sizes and relative positions that might affect this are based on a hypothetical female user with relevant features at the, in this regard, most disadvantageous 95th percentile of anthropometric data. The avatar model determines sensor placement, pivot points, and acts as a possible occluding agent for the sensor FoV, but not collisions. Collisions are instead determined by if the horizontal distance between the body pivot point and a collision object is too close, regardless of its height. Note that, since lower torso and extremities are thus both unnecessary for collision detection and redundant for sensor occlusion, they are omitted from the model. *Figure and caption originally produced for this work, but published first with slight differences in [27].*

TABLE 5. Anthropometric data of the VR-avatar in cm. *Table originally produced for this work, but published first in [27].*

Stature	179	**Shoulder (bideltoid) breadth**	47
Eye height	167	**Thorax depth at the nipple**	30
Head length	20	**Chest frontal transl.**[2]	6.5
Head vertical transl.[1,2]	8.8	**Nipple height**[3]	130
Head frontal transl.[2]	1.5		

---

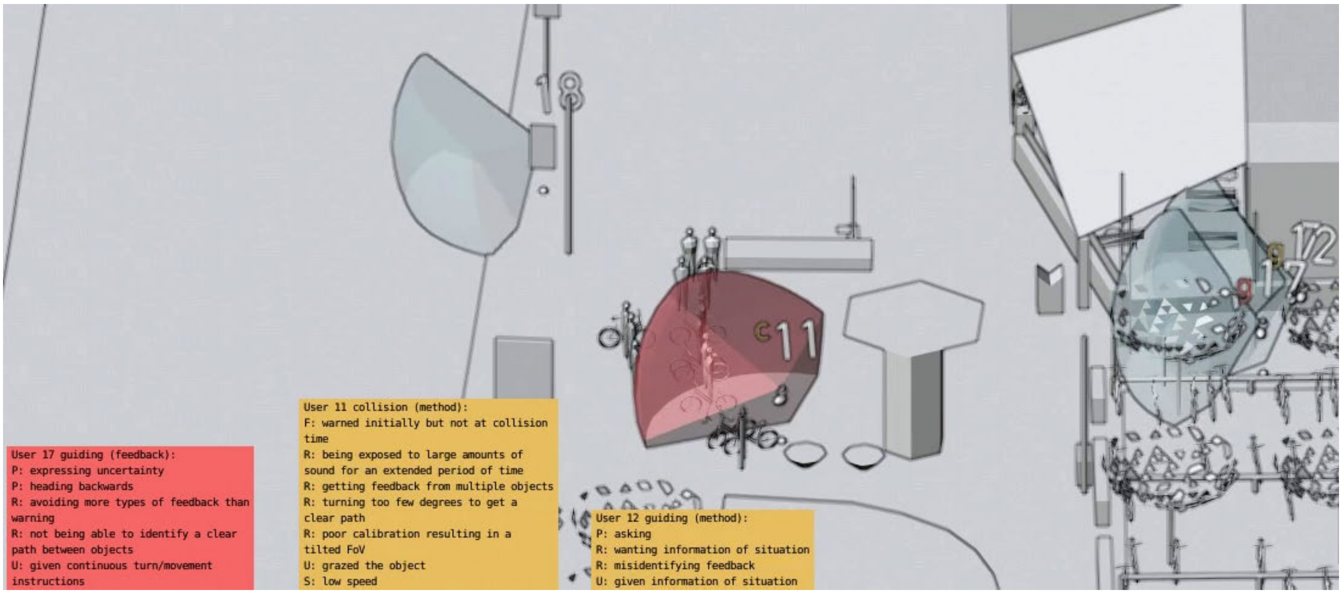
[1] Translation. [2] From its pivot point. Based on ad hoc measurements. [3] Approximated as midpoint between shoulder and elbow height.

C. RESULTS AND DISCUSSION (PARROT-VR)

A Portable, Absolute Rotation, Relative (O) Translation, Virtual Reality system, the Parrot-VR, has been proposed. It is an egocentric and large-scale VR system which uses an IMU and a video-game controller as input. Thus it has a hybrid type of positioning, with absolute rotation and relative translation, of the VR avatar. It can provide audio feedback produced by a virtual ETA, as well as haptic feedback produced by the system or a virtual ETA through controller vibration. Currently, a virtual version of the Audomni ETA is integrated into it.

The Parrot-VR is designed for ETA evaluation in relevant, large-scale, urban VEs. Meanwhile, its portability makes it deployable almost anywhere, allowing tests in e.g. participants' homes. This assists substantially in recruiting an adequate number of BLV participants, as well as can provide motivated and reproducible testing scenarios—all frequent issues in the field as discussed in Section I. In addition, it can be quite helpful in an ETA design process to speed up trials of design proposals for specific mobility scenarios. Thus, while VR testing cannot be seen as a full replacement to the real-life counterpart, which must be considered critical before any ETA can be realized as an end-product, Parrot-VR can assist meaningfully in ETA design and evaluation, and in a safe and resource-efficient manner, as supported by [53].

Even so, Parrot-VR does have room for further improvement. Currently it does not support ambient sounds, which limits its potential for macro-scale navigation assessment. Further, only the VR avatar is moveable during sessions, all obstacles and objects are stationary. Enabling movement of other objects would allow for more scenarios, e.g. crossing busy streets, finding arriving public transport, and avoiding moving pedestrians. Additionally, a more natural walking movement of the avatar could be considered. Currently the body and head are perfectly still vertically at all times; to simulate more realistic ETA behavior, up-and-down movement of body and head during walking could be added. Lastly, presently only the Audomni ETA is integrated into Parrot-VR and naturally other ETA designers might want to evaluate other devices. At this time the Audomni functionality is produced through function calls within the Parrot-VR software code, and thus they are integrated at compilation time. One solution to provide functionality of another ETA could be to replace those function calls to those of the device in question, and alter e.g. sensor placement and properties, and ETA input data, accordingly. This would require that the ETA functionality could be compiled on the Parrot-VR hardware, and that it could handle the feedback modality of the ETA. Another solution would be to provide the necessary input data to the ETA through a physical connection, e.g. SSH or USB. Ultimately, how complex it will be to integrate another ETA will depend heavily on what type of input data it utilizes, how it receives its input data, and what type of hardware it uses for feedback; ETAs making



```
User 17 guiding (feedback):
P: expressing uncertainty
P: heading backwards
R: avoiding more types of feedback than
warning
R: not being able to identify a clear
path between objects
U: given continuous turn/movement
instructions
```

```
User 11 collision (method):
F: warned initially but not at collision
time
R: being exposed to large amounts of
sound for an extended period of time
R: getting feedback from multiple objects
R: turning too few degrees to get a
clear path
R: poor calibration resulting in a
tilted FoV
U: grazed the object
S: low speed
```

```
User 12 guiding (method):
P: asking
R: wanting information of situation
R: misidentifying feedback
U: given information of situation
```

FIGURE 5. **Screen capture of a final video output of a collection of test sessions in a virtual environment (VE), including detailed event analysis. The shown scene is part of a VE reproduced from a real-life train station area and feature five previously and separately recorded user-controlled avatars. (See [53] for details regarding this scene and user test). The text boxes provide details of recent collision or guiding events, and are color-coded based on severity and underlying reason (information also present in the boxes). Here, the numbers above the avatars denote user number, and might feature a smaller and color-coded "c" or "g" denoting that a collision or guiding has just occurred. Note that the red sensor FoV on user 11 denotes a collision occurring at the moment of the capture.**

use of non-physical distance measurements should be a fairly straightforward procedure, while simulating a white cane or guide dog would be quite an undertaking.

The VE production workflow is quite laborious. While it is viable for user tests where each participant is to traverse the same environment, or for training with an ETA in general, it is almost untenable for user-specific testing and training. For such objectives it would be worthwhile to find and develop other, more efficient workflows, possibly by adapting 3D scanning, photogrammetry, or faster Blender tools. Though it is likely that more experience for VE producers with the current workflow would yield noticeable time-savings as well.

III. PROPOSED PATIENT-REPORTED OUTCOME MEASURE (DoUQ-MoB)

A. INTRODUCTION (DoUQ-MoB)

In this section we first review existing PROMs[1] for BLV individuals to support that there is need of a new questionnaire to assess ETAs. Then, current best practices for survey design are presented; followed by design choices, validations, and iterations of the questionnaire. Lastly, the resulting questionnaire is discussed.

1) PROMS FOR BLV INDIVIDUALS (DoUQ-MoB)

A large number of PROMs used within ophthalmology research and practice have been developed, as evidenced in recent reviews, e.g. [78], [79], [80]. These PROMs might target specific diseases or be more generalized. However, most do not focus on O&M, and thus keeps any items regarding it at high abstraction levels. Indeed, to our best knowledge the only exceptions are the Independent Mobility Questionnaire (IMQ) [81], and the Orientation and Mobility Outcomes (OMO) [82], [83], a notion underpinned by [18] and [84].

Whereas the IMQ was designed for respondents with retinitis pigmentosa and only moderate vision loss [81], Fenwick et al. showed that by modifying the instructions to take any mobility aid into account when answering, it can be used for BLV respondents as well [84]. Still, they also proposed substantial modifications based on their Rasch analysis. In another work, they employed the IMQ—seemingly with at least some of their proposed changes—with BLV participants to evaluate a mobility test battery [85]. However, the only instance where the IMQ actually has been used to evaluate a technical intervention might be in [86], where a modified version was used to assess mobility before and after fitting an electronic retinal implant. The IMQ is soundly validated and exhibits mobility tasks at different abstraction levels; regrettably, that it was designed for respondents with moderate vision loss, in addition to scant detail in its original item selection, prompts questioning if it is encompassing and specific enough for BLV individuals, a concern also raised in [85]. Indeed, there are many mobility aspects identified in DoU-MoB which are not included in the IMQ [9], and without specific instructions, many items are highly dependent on various circumstances, e.g. weather, lighting, and if they are in known or unknown environments.

Though the OMO is designed to "compare skills pre–post O&M training" and "evaluate new assistive technologies" [82], as yet there seems to be no published account of it being employed for such. In addition, whereas it affords and promotes qualitative data gathering, the high abstraction level of its items makes it ill-suited to make systematic fine-grained assessments of ETAs. However, the aggregate score it offers might be of interest in long-term studies for instance.

PROMs not specifically focusing on O&M, but including an O&M domain and are relatively well-known are the Impact of Vision Impairment—Very Low Vision (IVI-VLV) with arguably nine items concerning O&M [87], the Ultra-Low Vision Visual Functioning Questionnaire (ULV-VFQ) with ten mobility tasks [88], and the Functional Low-Vision Observer Rated Assessment (FLORA) with eleven visual orientation or mobility tasks [89]. While the O&M items from these instruments might be extracted and used in isolation (ignoring any aggregate scoring scheme), they can all be considered visual functioning questionnaires, pertinently problematic as much of those tend to consist of strictly vision-dependent items, e.g. "When walking down a dimly lit hallway looking for a meeting room, how difficult is it to see if the lights are on in the room?" from ULV-VFQ, or "Locate lights in the environment" from FLORA. On the other hand, the IVI-VLV feature vision independent items, but they lack in detail.

As implied above, and elucidated in our previous work [14], as well as in the works of [13], [15], [16], [17], [18], and [19], it is abundantly clear that no specific questionnaire regarding BLV mobility has enjoyed any notable or consistent use in ETA evaluation. Thus, an ETA designer wanting to take stock of as many relevant aspects as possible of their aid—according to possible end-users—is left with either the IMQ, which still is not fine-grained enough, might not be suitable for BLV individuals, nor offers any mobility aid questions; or designing their own PROM.

2) BEST KNOWN PRACTICES AND OTHER CONSIDERATIONS FOR QUESTIONNAIRE DESIGN (DoUQ-MoB)

Whereas a full review of survey design research is outside the scope of this article, a summary of identified best current practices relevant for physically and verbally administered PROMs are outlined here, and the results are summarized in Table 6. For recent reviews regarding survey design we instead refer to [90], [91], [92], and [93], where for instance [92] and [93] outline various scale design choices that might impact data quality, and the former also provides a number of recommended practices.

From its inception, Likert-type, agree–disagree scales have seen abundant use in many fields; however, recently this format has been subject to increasing scrutiny, with item-specific questions instead gaining in favor. Summarizing the reviews in [90] and [92] (with the same lead authors), item-specific questionnaires is recommended "for most purposes" as such are likelier to be of higher validity and reliability, as corroborated by the review in [91].

Regarding unipolar or bipolar scales there is considerably less research, however [94] suggests using unipolar as it seems to yield more reliable results, possibly due to a lower cognitive load. Unipolar scales also avoids the controversy of offering a neutral middle option or not.

When it comes to labelling, there is reason to believe that verbal response anchors, instead of numerical, increase stability and reliability [92], which might hold especially true for interviewer-administered questionnaires based on our own experience, see Section III-B1. However, in [95], using both numerical and verbal labels (along with a linear graphical scale, though that may not be relevant given a BLV target group) is suggested; a case is made for fully labelled (in contrast to end-labelled) scales; and using "well-researched" verbal anchors is recommended. Schaeffer and Dykema succinctly stated that "Given the continuing confirmations that verbal labels increase reliability and the need for consistency across modes and presentations, five categories for unipolar and possibly bipolar scales is both justifiable and practical at this time" [92], as well as proposed labels for unipolar intensity, frequency and quantity scales (based on [95]), e.g. for intensity: *not at all*, *slightly/a little*, *somewhat/moderately*, *very*, and *extremely*.

Yet that clear recommendation of five response categories might be regarded contentious given that scale length is source of many publications, and of mixed results. In [96] it was found that for uni- and bipolar scales, five and seven categories, respectively, seems favorable for reliability. In the review of [95], it is argued that it makes a difference if the scales are subject- or stimulus-centered, i.e. are the separate items to be aggregated or to be examined separately—the latter of which is of more interest for a PROM for assessing numerous aspects. For stimulus-centered scales it seemed that reliability and validity increases for up to seven responses, though the author makes no distinction between uni- and bipolar scales. Meanwhile, the results in [94] suggests fewer response categories in general to increase reliability.

Results from [97] also promote fewer categories to reduce scale direction effects, i.e. bias due to the order of response categories. It also indicates that earlier response categories are more prone to be chosen for earlier items in a questionnaire, due to which the authors suggest randomizing item order between respondents. The author of [98] states that scale order effects are difficult to mitigate, but suggests that if one end of a scale is of more interest, that end should possibly come last. Both of these scale direction effects were also found to be exacerbated in interviewer-administered questionnaires.

Question language is discussed in [92], and the authors particularly recommends to match the responses a question might naturally prompt with the actual response categories.

As our questionnaire is designed to be administered via physical interviews, and for BLV respondents, there are some additional considerations. First, substantially longer questionnaires become possible with physical interviews, compared to self-administration or phone interviews [99]. Second, due to the setting and target group, it becomes most convenient for the administrator to read the items aloud to the respondent and to note the answers, especially also as it allows for speedy changes of previous answers. Third, verbal administration limits the quantity of fully-labelled categories viable, as they take time to read, and must be remembered by the respondent; it also renders numerical and verbal labels (and a graphical scale, but that should likely be omitted regardless given the target group) in conjunction prohibitive. Fourth, in a verbal setting it becomes possible and natural for respondents to ask clarifying questions, and for the administrator to assist; though this can also reduce the consistency between respondents. As summarized in [100], physical interviews are advantageous over other modes for complex subjects, many response categories, complex structure, respondents of differing needs or instructions, longer questionnaires, interpretation needs, immediate control of answers, and open questions; while they are disadvantageous when it comes to questions regarding information from e.g. registers and receipts, as well as for standardized questions.

Since our parallell work [53], was to include solely Swedish participants, and inclusion of such is very likely in future works, the questionnaire is currently constructed in Swedish. This has consequences regarding choice of verbal anchors. Whereas the issue has been studied in English, e.g. [92], the body of research for Swedish counterparts is inadequate at best. A review of noise annoyance surveys found that Swedish translations of the labels *not at all*, *slightly*, *moderately*, *very*, and *extremely*—which corresponds very well with Schaeffer and Dykema's suggestions [92]—had much discrepancy for *moderately* and *extremely* [101]. While there seems to be consensus in translating *not at all* to *inte alls*, *slightly* to *lite*, and *very* to *mycket*; *moderately* either got translated to *måttligt* or *ganska mycket*, which the authors (and we) claim are not equivalent; and *extremely* to *väldigt*, *oerhört mycket*, or *extremt*. The authors suggest the labels *inte alls*, *lite*, *måttligt* (or *ganska mycket*), *mycket*, and *väldigt mycket*; however the only consideration for the uncertain labels seems to have been to reach consensus, with no regard taken to equidistance of labels or translation accuracy. In [102] the authors from the Swedish governmental agency Statistics Sweden, claim that, while not perfect, *måttligt* is likely the closest Swedish equivalent to *moderately*, and less loaded than *ganska mycket*. They made no note of *extremely*.

B. MATERIALS AND METHODS (DoUQ-MoB)

Our proposed PROM, the DoU Questionnaire for Mobility of BLV individuals (DoUQ-MoB), is designed to encompass the most important aspects of mobility and mobility aids for BLV individuals, in an aid-agnostic manner, while following the aforementioned best known practices in survey design.

TABLE 6. Questionnaire design considerations and implementations in the second DoUQ-MoB version. ***Table reproduced with permission from [27].***

Consideration	Implementation	References
Agree–disagree or item-specific questions	Item-specific	[90]–[92]
Uni- or bipolar response scales	Bipolar	[94]
Response anchor labelling: numeric or verbal, end- or fully labelled	Fully-labelled verbal anchors	[92], [95]
Response scale length	Five anchors	[92], [94]–[96]
Scale direction effects	Randomized items	[97], [98]
Response anchor choice (English)	*not at all*, *slightly*, *moderately*, *very*, *extremely*	[92], [95]
Response anchor choice (Swedish)	*inte alls*, *lite*, *måttligt*, *mycket*, and *extremt*	[92], [101], [102]

1) DESIGN CHOICES AND VALIDATION OF THE FIRST DoUQ-MoB VERSION (DoUQ-MoB)

The first iteration of the DoUQ-MoB was heavily inspired by the User Experience Questionnaire [103], [104]. As such, item-specific questions with the numeric, bipolar, and end-labelled scale of [*1..7*] were used. (Such scales are aka semantic differential scales.) The items were derived from the mobility aspects identified in the DoU-MoB [9]. To cover each aspect as comprehensively as possible, multiple items were constructed for some aspects. Mobility aspects lower in the hierarchy and of lower abstraction were generally given more specific items, and aspects higher up more general. Care was taken to employ as simple and unambiguous language as possible.

This questionnaire was subjected to an initial validation and pilot-test with eight O&M experts and six BLV respondents to investigate face and content validity, test–retest reliability, as well as any apparent questionnaire issues. As this validation can be considered quite fundamental, a review of questionnaire validation is out of scope of this work (state of the art and current best practices can be found in [105], [106], [107], [108], [109].) The questionnaire instructions were read to the O&M experts, who then for each item got to rate it as *not necessary*, *useful but not essential*, or *essential* [109], as well as was encouraged to comment the items. Afterward, they had the opportunity to suggest any aspects or items that they felt were missing. The BLV respondents answered the questionnaire as intended, but were like the O&M experts also encouraged to comment the items, as well as the questionnaire in general. After 9–15 days, mean = 13 days, the respondents got to answer the questionnaire again. After each interview, the respondent was asked what any large discrepancies in item responses between the first and second round might have been due to.

After the interviews with the O&M experts and BLV respondents followed various analyses, performed in the numeric computing platform Matlab [110]. The experts' answers were converted to the range [0..2], after which the 95 %-level confidence intervals (CIs) for the mean of each item were computed through bootstrapping. Cohen's weighted kappa were determined for the BLV respondents' answers [111], where each respondent's answers in the first and second round were compared; furthermore, the 95 %-level CIs were computed using both the method of Fleiss et al. [112], and through bootstrapping. The kappa weights were the quadratic normalized numerical distances of the answer confusion matrix. The Pearson correlation coefficients for the BLV respondents' items were also computed, including CIs using Fisher's procedure, e.g. [113, pp. 187–190]. Warnings were then issued for items which various CIs did not include any of the following thresholds (where values in parentheses yielded a more severe warning): mean of expert answers $= 2\,(\geq 1.75)$; kappa $\geq 0.4\,(\geq 0.2)$. Warnings were also issued for suspected ceiling or floor effects, defined as $\geq 50\,\%$ of BLV respondents' answers being in the top or bottom bin; and any absolute item correlation CI not restricted to $\leq 0.4\,(\leq 0.6)$. The warnings and any comments were summarized, and subsequently guided us in item removals, additions or alterations, as well as changes to the questionnaire instructions and answer scales for the next design iteration of the questionnaire. The parameters used in the analysis, along with visualizations of the results, is available in the supplemental material; a post-analysis document summarizing the results for each item, and including suggestions for changes for the second version of the questionnaire, is also available there in the Swedish original.

Most of these results related to individual items, but a recurring theme was seemingly poor test–retest stability, with the majority of items exhibiting CIs that did not include high kappas, and in many cases missed the thresholds (which are set deliberately low for this reason). This might have been due to systematic issues with ambiguous instructions or questions, or confusing response scales—which is expanded on below. However, a major contributor was likely respondent or environment instability, or both. Many respondents reasoned that they scored differently in the second round due to that, in-between rounds, they had been considering the items during walks, as well as that for some respondents heavy snowfall had impacted their walks. Along with the low number of respondents, any conclusions about test–retest stability should not be drawn before the latest version of the DoUQ-MoB has seen such tests, and at a larger scale.

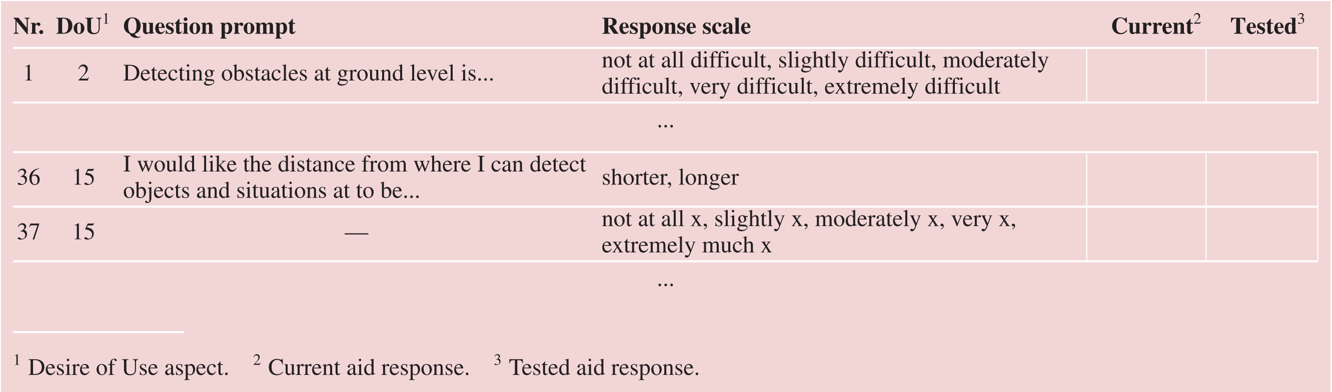
2) DESIGN CHOICES AND MODIFICATIONS OF THE SECOND DoUQ-MoB VERSION (DoUQ-MoB)

Changes to the items between the first and second version of the the DoUQ-MoB included: merging highly correlated items; removal of evidently confusing items, or items the O&M experts thought were not as relevant or important; as well as splitting many items into two, one regarding in known areas and one in unknown, in order to make those items less ambiguous and reduce "it depends…" type of answers. For this reason, the instructions were also modified for respondents to, if relevant for the item, answer as if road conditions are dry; lighting conditions are the worst possible for the respondent; and if other circumstances are relevant, try to make their best balanced judgement.

To reduce systematic scale direction and item order effects, while avoiding to overly increase cognitive load and questionnaire administration time, the order of items for each respondent are semi-randomized in the second version of the DoUQ-MoB. Items are grouped according to which DoU-MoB aspect they derived from. Then the group order is randomized, as well as the item order within groups, except for the known–unknown area pairs, which are never separated and always presented with known area first. Avoiding complete randomization in this manner likely reduce the cognitive load for respondents, who do not have to switch area of questioning as often. They can also often learn that they can answer for a corresponding unknown area item instantly after the known area item, saving a meaningful amount of time.

The response scales were also modified. While the scale of [*1..7*] might work for the User Experience Questionnaire [103], [104]; since the DoUQ-MoB is administered verbally, many respondents had trouble identifying the midpoint (*4*), and often mistakenly answered in the middle or to one side of the scale. Also, there was a substantial number of floor–ceiling effects in the responses; therefore the verbal end-point anchors were changed from *very x* to *extremely x*; and the scale was shifted to get midpoint *0*, and extended from seven to eleven choices, resulting in the range [*-5..5*]. The numerical scale was kept to accommodate subsequent parametrical statistical analyses, as the notion that numerical labels does not necessarily yield interval data was not considered at the time [98].

These scales were used for the first two participants of our parallell work [53]; however, it became clear that they also struggled to answer numerically, often asking what number for instance *very difficult* might be. Thus, it was decided to change the scales again, and this time to, as closely as possible, adhere to the best practices discussed in Section III-A2. We therefore adopted unipolar and fully-labelled verbal scales with five response categories. Though accompanying numerical labels were omitted as they would take considerable extra time to read aloud. During that work, it also transpired that a few questions regarding color and material of an aid were difficult to answer for most respondents, thus they were eventually excluded.



Nr.	DoU[1]	Question prompt	Response scale	Current[2]	Tested[3]
1	2	Detecting obstacles at ground level is...	not at all difficult, slightly difficult, moderately difficult, very difficult, extremely difficult		
			...		
36	15	I would like the distance from where I can detect objects and situations at to be...	shorter, longer		
37	15	—	not at all x, slightly x, moderately x, very x, extremely much x		
			...		

[1] Desire of Use aspect. [2] Current aid response. [3] Tested aid response.

FIGURE 6. **Excerpt from and typical formatting of the Desire of Use Questionnaire for Mobility of Blind and Low-Vision Individuals (DoUQ-MoB). Note that the questions are tentative translations of the Swedish questions. Also note the filter question *36*, in which the respondent first choose response polarity, and then in question *37* choose the corresponding strength. *Figure and caption originally produced for this work, but published first in [27].***

Awaiting any more work regarding label choice in Swedish, see the outline above, for the unipolar items we opted for a scale of *inte x alls*, *lite x*, *måttligt x*, *mycket x*, and *extremt x*, *x* denoting a specific answer for a given item. Whereas the choice of the other labels should be clear from above, we chose *extremt x* in lieu of *väldigt mycket x*, as the latter only works for quantities (compare the English *very much*), and at face value, the English *extremely* and the Swedish *extremt* should be close to equivalent. This was also linguistically sound for the vast majority of items, but for *more* or *less* quantifiers, as well as for *time*-responses, the last category was changed to *extremt mycket x* [*extremely much x*] in order to keep the scale as similar as possible throughout the questionnaire. For the *time* items, also the first category was changed to *inte lång tid alls* [*not much time at all*].

Awaiting more work regarding Swedish label choices, see the outline in Section III-A2, we arrived at the following. For the unipolar items we opted for a scale of *inte x alls*, *lite x*, *måttligt x*, *mycket x*, and *extremt x*. Whereas the choice of the other labels should be clear from Section III-A2, we chose *extremt x* in lieu of *väldigt mycket x*, as the latter only works for quantities (compare to the English *very much*), and at face value, the English *extremely* and the Swedish *extremt* should be close to equivalent. This was also linguistically sound for the vast majority of items, but for *more* or *less* quantifiers, as well as for *time*-responses, the last category was changed to *extremt mycket x* [*extremely much x*] in order to keep the scale as similar as possible throughout the questionnaire. For the *time* items, also the first category was changed to *inte lång tid alls* [*not much time at all*].

Most items could be constructed with unipolar response scales, e.g. "To detect obstacles at ground level is…" [*not at all difficult..extremely difficult*]. However, some items were constructed with first a filter question, and then the scale, e.g. "I would have wanted that the distance at which I can discover objects and situations at were…"[*shorter* or *longer*], were an answer of *shorter* for instance would prompt the scale [*not at all shorter..extremely much shorter*]. This also helped in keeping the answer style consistent between items. See Fig. 6 for example questions and format of the current DoUQ-MoB version, and Table 6 for a summary of how the questionnaire design considerations were implemented.

We regard the current DoUQ-MoB considerably improved over the pre-validation version; however, given the issues with test–retest stability of the former version, employing the DoUQ-MoB as an outcome metric over time should be avoided until the new version has been validated for that. However, the current method of answering an item for both aids simultaneously, and only comparing the answers for the specific items, should yield reliable intra-user comparisons.

C. RESULTS AND DISCUSSION (DoUQ-MoB)

To summarize, the questionnaire DoUQ-MoB is an aid-agnostic PROM[1] to systematically assess mobility of BLV individuals using a mobility aid. It especially facilitates the comparison of such aids for important mobility and mobility aid aspects, where each item and aspect can be compared between aids. It is administered through physical interviews, and takes around one hour to complete. It consists of 106 item-specific questions with unipolar and fully-labelled scales with five response categories, and strictly verbal labels using the closest to us known Swedish equivalents of *not at all*, *slightly/a little*, *somewhat/moderately*, *very*, and *extremely*.

The questionnaire is available in its entirety in the supplemental material, both in the original Swedish version and in a tentatively translated English version.

When comparing aids, for instance a proposed ETA with a respondent's currently used aid, each item is answered once for each device before moving on to the next item. In addition, the last seven items of the questionnaire asks the respondent to directly compare the aids. This final part also encourages follow-up questions as a semi-structured interview, as to probe what the participants found most interesting, beneficial, or problematic with an ETA proposal.

Efforts have been made to keep the language as simple and unambiguous as possible; and instructions as well as items have been adapted to mitigate how external circumstances such as weather and lighting conditions, as well as known or unknown areas, could affect an answer. As yet, it has seen good results for face and content validity in a validation round with O&M experts and BLV respondents. How DoUQ-MoB results can be analysed is demonstrated in our other work [53].

The comprehensive set of items of the DoUQ-MoB, provides fine-grained measures of what is likely most important mobility and mobility aid aspects, thus elucidating strengths and weaknesses of an aid in both a broad and a specific sense. It achieves this through probing insight from BLV respondents, the group that needs to find an ETA valuable, lest it will not be used. In addition, it examines the respondents' attitudes of the proposed aid in relation to their current aids—a critical test for any proposal to reach an adequate end-product. These attributes are valuable in and of themselves, but especially so as the field of ETA development is severely lacking in all of them, as shown in Section I.

In future work it would be interesting to offer an aggregate DoU score for a respondent with a given aid through a Rasch analysis. It is also possible that a review from survey design experts could increase the validity and reliability of the DoUQ-MoB.

IV. PARROT-VR AND DoUQ-MoB DEPLOYMENT RESULTS

The Parrot-VR and DoUQ-MoB have been successfully deployed for assessing the Audomni ETA, see [53]. For that work, VEs depicting a real-life train station as well as a user test procedure were developed for use with the Parrot-VR and DoUQ-MoB, so we refer to it for an example of such practical details. In it 19 BLV participants, well-distributed in regards to e.g. gender, age, and education, used the Parrot-VR to explore the VEs; they then assessed the aid and their current aid using the DoUQ-MoB. The work then presents a detailed movement analysis utilizing the VR recordings, as well as summarizes and analyzes the questionnaire results and interviews. A discussion regarding the strengths and limitations of the respective tools also follows. As it allows for considerably more astute observations after a tool has seen real use, we refer to that work also for further discussion of the Parrot-VR and DoUQ-MoB.

V. MAIN CONCLUSION

By allowing for systematic and fine-grained user assessment focused on numerous aspects of mobility and mobility aids, the DoUQ-MoB fills an important gap in the field of evaluating ETAs. The DoUQ-MoB follows the best known practices of survey design. In addition, as it is aid-agnostic, it offers the possibility to—and actively promotes—comparison between BLV individuals' current mobility aids and a proposed aid.

Meanwhile, the Parrot-VR system enables reproducible testing and rich movement analysis. Its control scheme allows for large-scale VE exploration, which in turn makes it possible to evaluate ETAs in relevant surroundings. The absolute rotation and minimal use of the controller keeps it relatively intuitive and easy to use for beginners. In addition, to our knowledge this is the first VR system in the field designed to be portable, a feature which greatly facilitates the recruitment of BLV participants, as tests can be performed close to users.

Together, the DoUQ-MoB and Parrot-VR aids considerably in ETA assessment, by allowing for and encouraging: well-motivated testing scenarios, reproducibility, systematic measures at various abstraction levels, BLV participant recruitment, and user evaluation—increasing the chances of reaching a user-accepted ETA, and ultimately increasing its users' mobility.

ACKNOWLEDGMENT

The O&M professionals and the questionnaire respondents have their sincere appreciation for their time and effort in their respective parts of the questionnaire validation.

ABBREVIATIONS

AR	Augmented Reality.
BLV	Blind and Low-Vision.
CI	Confidence Interval.
DoU	Desire of Use.
DoU-MoB	Desire of Use for Mobility of Blind and low-vision individuals.
DoUQ-MoB	Desire of Use Questionnaire for Mobility of Blind and low-vision individuals.
ETA	Electronic Travel Aid.
FLORA	Functional Low-Vision Observer Rated Assessment.
FoV	Field-of-View.
IMQ	Independent Mobility Questionnaire.
IMU	Inertial Measurement Unit.
IVI-VLV	Vision Impairment—Very Low Vision.
O&M	Orientation and Mobility.
OMO	Orientation and Mobility Outcomes.
PPWS	Percentage of Preferred Walking Speed.
PROM	Patient-Reported Outcome Measure.
SSD	Sensory Supplementation Device.
SSH	Secure Shell protocol.
SSHFS	SSH Filesystem.
ULV-VFQ	Ultra-Low Vision Visual Functioning Questionnaire.
VE	Virtual Environment.
VR	Virtual Reality

REFERENCES

[1] R. Bourne et al., "Trends in prevalence of blindness and distance and near vision impairment over 30 years: An analysis for the global burden of disease study," *Lancet Global Health*, vol. 9, no. 2, pp. e130–e143, 2021.

[2] *World Report on Vision*, World Health Org., Geneva, Switzerland, 2019.

[3] M. Goldschmidt, "Orientation and mobility training to people with visual impairments," in *Mobility of Visually Impaired People*, E. Pissaloux and R. Velázquez, Eds., Cham, Switzerland: Springer, 2018, pp. 237–261.

[4] R. Manduchi and S. Kurniawan, "Mobility-related accidents experienced by people with visual impairment," *AER J., Res. Pract. Vis. Impairment Blindness*, vol. 4, no. 2, pp. 44–54, Feb. 2011.

[5] R. Velázquez, "Wearable assistive devices for the blind," in *Wearable and Autonomous Biomedical Devices and Systems for Smart Environment*, vol. 75, A. Lay-Ekuakille and S. Mukhopadhyay, Eds. Heidelberg, Germany: Springer, 2010, pp. 331–349.

[6] M. Hersh and M. Johnson, "Mobility: An overview," in *Assistive Technology for Visually Impaired and Blind People*, M. Hersh and M. Johnson, Eds., London, U.K.: Springer, 2008, pp. 167–208.

[7] J. Loomis, R. Golledge, R. L. Klatzky, and J. Marston, "Assisting wayfinding in visually impaired travelers," in *Applied Spatial Cognition: From Research to Cognitive Technology*. London, U.K.: Psychology Press, 2007, ch. 7, pp. 179–202.

[8] F. S. Cooper, "Guidance devices for the blind," *Phys. Today*, vol. 3, no. 7, pp. 6–14, Jul. 1950.

[9] J. Isaksson, T. Jansson, and J. Nilsson, "Desire of use: A hierarchical decomposition of activities and its application on mobility of by blind and low-vision individuals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1146–1156, May 2020.

[10] D.-R. Chebat, V. Harrar, R. Kupers, S. Maidenbaum, A. Amedi, and M. Ptito, "Sensory substitution and the neural correlates of navigation in blindness," in *Mobility of Visually Impaired People*, E. Pissaloux and R. Velázquez, Eds., Cham, Switzerland: Springer, 2018, pp. 167–200.

[11] Á. Kristjánsson, A. Moldoveanu, Ó. I. Jóhannesson, O. Balan, S. Spagnol, V. V. Valgeirsdóttir, and R. Unnthorsson, "Designing sensory-substitution devices: Principles, pitfalls and potential," *Restorative Neurol. Neurosci.*, vol. 34, no. 5, pp. 769–787, Sep. 2016.

[12] G. V. Elli, S. Benetti, and O. Collignon, "Is there a future for sensory substitution outside academic laboratories?" *Multisensory Res.*, vol. 27, nos. 5–6, pp. 271–291, 2014.

[13] P. Xu, G. A. Kennedy, F.-Y. Zhao, W.-J. Zhang, and R. Van Schyndel, "Wearable obstacle avoidance electronic travel aids for blind and visually impaired individuals: A systematic review," *IEEE Access*, vol. 11, pp. 66587–66613, 2023.

[14] J. Isaksson, T. Jansson, and J. Nilsson, "Audomni: Super-scale sensory supplementation to increase the mobility of blind and low-vision individuals—A pilot study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1187–1197, May 2020.

[15] A. D. P. D. Santos, A. H. G. Suzuki, F. O. Medola, and A. Vaezipour, "A systematic review of wearable devices for orientation and mobility of adults with visual impairment and blindness," *IEEE Access*, vol. 9, pp. 162306–162324, 2021.

[16] S. Khan, S. Nazir, and H. U. Khan, "Analysis of navigation assistants for blind and visually impaired people: A systematic review," *IEEE Access*, vol. 9, pp. 26712–26734, 2021.

[17] X. Zhang, X. Yao, L. Hui, F. Song, and F. Hu, "A bibliometric narrative review on modern navigation aids for people with visual impairment," *Sustainability*, vol. 13, no. 16, p. 8795, Aug. 2021.

[18] K. J. Chang, L. L. Dillon, L. Deverell, M. Y. Boon, and L. Keay, "Orientation and mobility outcome measures," *Clin. Exp. Optometry*, vol. 103, no. 4, pp. 434–448, Jul. 2020.

[19] D. Plikynas, A. Žvironas, M. Gudauskis, A. Budrionis, P. Daniušis, and I. Sliesoraitytė, "Research advances of indoor navigation for blind people: A brief review of technological instrumentation," *IEEE Instrum. Meas. Mag.*, vol. 23, no. 4, pp. 22–32, Jun. 2020.

[20] A. Budrionis, D. Plikynas, P. Daniušis, and A. Indrulionis, "Smartphone-based computer vision travelling aids for blind and visually impaired individuals: A systematic review," *Assistive Technol.*, vol. 34, no. 2, pp. 178–194, Mar. 2022.

[21] A. T. Parker, M. Swobodzinski, J. D. Wright, K. Hansen, B. Morton, and E. Schaller, "Wayfinding tools for people with visual impairments in real-world settings: A literature review of recent studies," *Frontiers Educ.*, vol. 6, pp. 1–23, Oct. 2021, Art. no. 723816.

[22] H. Walle, C. De Runz, B. Serres, and G. Venturini, "A survey on recent advances in AI and vision-based methods for helping and guiding visually impaired people," *Appl. Sci.*, vol. 12, no. 5, p. 2308, Feb. 2022.

[23] E. L. Horton, R. Renganathan, B. N. Toth, A. J. Cohen, A. V. Bajcsy, A. Bateman, M. C. Jennings, A. Khattar, R. S. Kuo, F. A. Lee, M. K. Lim, L. W. Migasiuk, A. Zhang, O. K. Zhao, and M. A. Oliveira, "A review of principles in design and usability testing of tactile technology for individuals with visual impairments," *Assistive Technol.*, vol. 29, no. 1, pp. 28–36, Jan. 2017.

[24] L. F. Cuturi, E. Aggius-Vella, C. Campus, A. Parmiggiani, and M. Gori, "From science to technology: Orientation and mobility in blind children and adults," *Neurosci. Biobehav. Rev.*, vol. 71, pp. 240–251, Dec. 2016.

[25] U. R. Roentgen, G. J. Gelderblom, M. Soede, and L. P. De Witte, "The impact of electronic mobility devices for persons who are visually impaired: A systematic review of effects and effectiveness," *J. Vis. Impairment Blindness*, vol. 103, no. 11, pp. 743–753, Dec. 2009.

[26] J. Isaksson, T. Jansson, and J. Nilsson, "A VR-approach to assess mobility of blind and low-vision individuals using the Audomni sensory supplementation device," in *Proc. Medicinteknikdagarna*, 2021, p. 1.

[27] J. Isaksson-Daun, "A sound approach toward a mobility aid for blind and low-vision individuals," Ph.D. dissertation, Dept. Biomed. Eng., Lund Univ., Lund, Sweden, 2023.

[28] N. Fernando, D. A. Mcmeekin, and I. Murray, "Route planning methods in indoor navigation tools for vision impaired persons: A systematic review," *Disab. Rehabil., Assistive Technol.*, vol. 18, no. 6, pp. 763–782, Aug. 2023.

[29] A. Zvironas, M. Gudauskis, and D. Plikynas, "Indoor electronic traveling aids for visually impaired: Systemic review," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2019, pp. 936–942.

[30] E. R. Araújo, C. Rodrigues, and L. B. Martins, "Facilitators for safety of visually impaired on the displacements in external environments: A systematic review," in *Occupational Safety and Hygiene VI*, P. M. Arezes, et al., Eds., Boca Raton, FL, USA: CRC Press, 2018, pp. 371–375.

[31] I. Khan, S. Khusro, and I. Ullah, "Technology-assisted white cane: Evaluation and future directions," *PeerJ*, vol. 6, Dec. 2018, Art. no. e6058.

[32] J. Kreimeier and T. Götzelmann, "Two decades of touchable and walkable virtual reality for blind and visually impaired people: A high-level taxonomy," *Multimodal Technol. Interact.*, vol. 4, no. 4, p. 79, Nov. 2020.

[33] F. S. Ricci, A. Boldini, J.-R. Rizzo, and M. Porfiri, "Learning to use electronic travel aids for visually impaired in virtual reality," *Proc. SPIE*, vol. 12045, Apr. 2022, Art. no. 1204504.

[34] S. Real and A. Araujo, "VES: A mixed-reality development platform of navigation systems for blind and visually impaired," *Sensors*, vol. 21, no. 18, p. 6275, Sep. 2021.

[35] S. Real and A. Araujo, "VES: A mixed-reality system to assist multisensory spatial perception and cognition for blind and visually impaired people," *Appl. Sci.*, vol. 10, no. 2, p. 523, Jan. 2020.

[36] P. B. L. Meijer, "An experimental system for auditory image representations," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 2, pp. 112–121, Feb. 1992.

[37] R.-G. Lupu, O. Mitruţ, A. Stan, F. Ungureanu, K. Kalimeri, and A. Moldoveanu, "Cognitive and affective assessment of navigation and mobility tasks for the visually impaired via electroencephalography and behavioral signals," *Sensors*, vol. 20, no. 20, p. 5821, Oct. 2020.

[38] M. Dascalu, A. Moldoveanu, O. Balan, R. G. Lupu, F. Ungureanu, and S. Caraiman, "Usability assessment of assistive technology for blind and visually impaired," in *Proc. E-Health Bioeng. Conf. (EHB)*, Jun. 2017, pp. 523–526.

[39] A. D. B. Moldoveanu, S. Ivascu, I. Stanica, M.-I. Dascalu, R. Lupu, G. Ivanica, O. Balan, S. Caraiman, F. Ungureanu, F. Moldoveanu, and A. Morar, "Mastering an advanced sensory substitution device for visually impaired through innovative virtual training," in *Proc. IEEE 7th Int. Conf. Consum. Electron. Berlin (ICCE-Berlin)*, Sep. 2017, pp. 120–125.

[40] S. Maidenbaum and A. Amedi, "Standardizing visual rehabilitation using simple virtual tests," in *Proc. Int. Conf. Virtual Rehabil. (ICVR)*, Jul. 2019, pp. 1–8.

[41] D. Massiceti, S. L. Hicks, and J. J. van Rheede, "Stereosonic vision: Exploring visual-to-auditory sensory substitution mappings in an immersive virtual reality navigation paradigm," *PLoS ONE*, vol. 13, no. 7, Jul. 2018, Art. no. e0199389.

[42] S. Levy-Tzedek, S. Maidenbaum, A. Amedi, and J. Lackner, "Aging and sensory substitution in a virtual navigation task," *PLoS ONE*, vol. 11, no. 3, Mar. 2016, Art. no. e0151593.

[43] D.-R. Chebat, S. Maidenbaum, and A. Amedi, "The transfer of non-visual spatial knowledge between real and virtual Mazes via sensory substitution," in *Proc. Int. Conf. Virtual Rehabil. (ICVR)*, Jun. 2017, pp. 1–7.

[44] D.-R. Chebat, S. Maidenbaum, and A. Amedi, "Navigation using sensory substitution in real and virtual Mazes," *PLoS ONE*, vol. 10, no. 6, Jun. 2015, Art. no. e0126307.

[45] S. Maidenbaum, D. R. Chebat, S. Levy-Tzedek, and A. Amedi, "Depth-to-audio sensory substitution for increasing the accessibility of virtual environments," in *Universal Access in Human-Computer Interaction. Design and Development Methods for Universal Access* (Lecture Notes in Computer Science), vol. 8513, C. Stephanidis and M. Antona, Eds., Cham, Switzerland: Springer, 2014, pp. 398–406.

[46] S. Maidenbaum, S. Levy-Tzedek, D.-R. Chebat, and A. Amedi, "Increasing accessibility to the blind of virtual environments, using a virtual mobility aid based on the 'EyeCane': Feasibility study," *PLoS ONE*, vol. 8, no. 8, Aug. 2013, Art. no. e72555.

[47] S. Maidenbaum, G. Buchs, S. Abboud, O. Lavi-Rotbain, and A. Amedi, "Perception of graphical virtual environments by blind users via sensory substitution," *PLoS ONE*, vol. 11, no. 2, Feb. 2016, Art. no. e0147501.

[48] S. Maidenbaum, S. Abboud, G. Buchs, and A. Amedi, "Blind in a virtual world: Using sensory substitution for generically increasing the accessibility of graphical virtual environments," in *Proc. IEEE Virtual Reality (VR)*, Mar. 2015, pp. 233–234.

[49] P. R. Sanz, B. R. Mezcua, J. M. S. Pena, and B. N. Walker, "Evaluation of a new mobility assistive product for the visually impaired," *Int. J. Comput. Technol.*, vol. 13, no. 2, pp. 4191–4205, Mar. 2014.

[50] W. L. Khoo, J. Knapp, F. Palmer, T. Ro, and Z. Zhu, "Designing and testing wearable range-vibrotactile devices," *J. Assistive Technol.*, vol. 7, no. 2, pp. 102–117, Jun. 2013.

[51] M. Bujacz, P. Skulimowski, and P. Strumiłło, "Naviton—A prototype mobility aid for auditory presentation of three-dimensional scenes to the visually impaired," *J. Audio Eng. Soc.*, vol. 60, no. 9, pp. 696–708, 2012.

[52] L. Thevin, C. Briant, and A. M. Brock, "X-road: Virtual reality glasses for orientation and mobility training of people with visual impairments," *ACM Trans. Accessible Comput.*, vol. 13, no. 2, pp. 1–47, Jun. 2020.

[53] J. Isaksson-Daun, T. Jansson, and J. Nilsson, "Using portable virtual reality to assess mobility of blind and low-vision individuals with the audomni sensory supplementation feedback," *IEEE Access*, vol. 12, pp. 26222–26241, 2024.

[54] *Ubuntu, 18.04.5 LTS, GNU/Linux 4.9.140-tegra*, Canonical, London, U.K., 2019.

[55] *ISO/IEC 14882:2017 Programming Languages—C++*, Standard ISO/IEC 14882:2017, Dec. 2017.

[56] (2023). *GNU Compiler Collection (GCC)*. Accessed: Jan. 5, 2023. [Online]. Available: https://gcc.gnu.org

[57] *BNO055 Sensor Driver*, Bosch Sensortec, Kusterdingen, Germany, 2020.

[58] *GLFW*. Accessed: Jan. 5, 2023. [Online]. Available: https://www.glfw.org

[59] *OpenGL (3.3)*, Khronos Group, Beaverton, OR, USA, 2010.

[60] *OpenGL Mathematics (GLM). (4.2). G-Truc Creation*. Accessed: Jan. 5, 2023. [Online]. Available: https://glm.g-truc.net

[61] *Open Asset Import Library (Assimp)*. Accessed: Jan. 5, 2023. [Online]. Available: https://github.com/assimp/assimp

[62] D. Herberth. *Glad*. Accessed: Feb. 7, 2023. [Online]. Available: https://github.com/Dav1dde/glad/tree/master

[63] S. Barret. *STB_Image*. Accessed: Jan. 5, 2023. [Online]. Available: https://github.com/nothings/stb

[64] *OpenCV (4.4)*, OpenCV Team, Palo Alto, CA, USA, 2020.

[65] J. de Vries. *Hello Window*. Accessed: Jan. 4, 2023. [Online]. Available: https://learnopengl.com/Getting-started/Hello-Window

[66] *Shaders*. Accessed: Jan. 4, 2023. [Online]. Available: https://learnopengl.com/Getting-started/Shaders

[67] *Camera*. Accessed: Jan. 4, 2023. [Online]. Available: https://learnopengl.com/Getting-started/Camera

[68] *Mesh*. Accessed: Jan. 4, 2023. [Online]. Available: https://learnopengl.com/Model-Loading/Mesh

[69] *Model*. Accessed: Jan. 4, 2023. [Online]. Available: https://learnopengl.com/Model-Loading/Model

[70] GitHub. *Gamepad Rumble Effects #1678*. Accessed: Jan. 4, 2023. [Online]. Available: https://github.com/glfw/glfw/pull/1678/files

[71] *Blender (2.92)*, Blender Found., Amsterdam, The Netherlands, 2021.

[72] *Blender-OSM*. Accessed: Jan. 5, 2023. [Online]. Available: https://github.com/vvoovv/blender-osm

[73] *OpenStreetMap Homepage*. Accessed: Jan. 5, 2023. [Online]. Available: https://www.openstreetmap.org

[74] *Google Maps Homepage*. Accessed: Jan. 5, 2023. [Online]. Available: https://www.google.com/maps

[75] *Apple Maps*, Apple, Cupertino, CA, USA, 2021.

[76] *Python (3.7.7)*, Python Softw. Found., Wilmington, DE, USA, 2020.

[77] L. Hanson, L. Sperling, G. Gard, S. Ipsen, and C. O. Vergara, "Swedish anthropometrics for product and workplace design," *Appl. Ergonom.*, vol. 40, no. 4, pp. 797–806, Jul. 2009.

[78] G. Gazzard, M. Kolko, M. Iester, and D. P. Crabb, "A scoping review of quality of life questionnaires in glaucoma patients," *J. Glaucoma*, vol. 30, no. 8, pp. 732–743, 2021.

[79] W. Luu, M. Kalloniatis, E. Bartley, M. Tu, L. Dillon, B. Zangerl, and A. Ly, "A holistic model of low vision care for improving vision-related quality of life," *Clin. Exp. Optometry*, vol. 103, no. 6, pp. 733–741, Nov. 2020.

[80] T. Braithwaite, M. Calvert, A. Gray, K. Pesudovs, and A. Denniston, "The use of patient-reported outcome research in modern ophthalmology: Impact on clinical trials and routine clinical practice," *Patient Rel. Outcome Measures*, vol. 10, pp. 9–24, Jan. 2019.

[81] K. Turano, D. Geruschat, J. Stahl, and R. Massof, "Perceived visual ability for independent mobility in persons with retinitis pigmentosa," *Invest. Ophthalmol. Vis. Sci.*, vol. 40, no. 5, pp. 865–877, 1999.

[82] L. Deverell, D. Meyer, B. T. Lau, A. Al Mahmud, S. Sukunesan, J. Bhowmik, A. Chai, C. McCarthy, P. Zheng, A. Pipingas, and F. M. A. Islam, "Optimising technology to measure functional vision, mobility and service outcomes for people with low vision or blindness: Protocol for a prospective cohort study in Australia and Malaysia," *BMJ Open*, vol. 7, no. 12, Dec. 2017, Art. no. e018140.

[83] L. Deverell, J. Bradley, P. Foote, M. Bowden, and D. Meyer, "Measuring the benefits of guide dog mobility with the orientation and mobility outcomes (OMO) tool," *Anthrozoös*, vol. 32, no. 6, pp. 741–755, Nov. 2019.

[84] E. K. Fenwick, F. O'Hare, L. Deverell, L. N. Ayton, C. D. Luu, S. McSweeney, S. A. Bentley, R. H. Guymer, and R. P. Finger, "Rasch analysis of the independent mobility questionnaire," *Optometry Vis. Sci.*, vol. 93, no. 2, pp. 181–187, 2016.

[85] R. P. Finger, L. N. Ayton, L. Deverell, F. O'Hare, S. C. McSweeney, C. D. Luu, E. K. Fenwick, J. E. Keeffe, R. H. Guymer, and S. A. Bentley, "Developing a very low vision orientation and mobility test battery (O&M-VLV)," *Optometry Vis. Sci.*, vol. 93, no. 9, pp. 1127–1136, 2016.

[86] T. L. Edwards, C. L. Cottriall, K. Xue, M. P. Simunovic, J. D. Ramsden, E. Zrenner, and R. E. MacLaren, "Assessment of the electronic retinal implant alpha AMS in restoring vision to blind patients with end-stage retinitis pigmentosa," *Ophthalmology*, vol. 125, no. 3, pp. 432–443, Mar. 2018.

[87] R. P. Finger, B. Tellis, J. Crewe, J. E. Keeffe, L. N. Ayton, and R. H. Guymer, "Developing the impact of vision impairment–very low vision (IVI-VLV) questionnaire as part of the Lovada protocol," *Invest. Ophthalmol. Vis. Sci.*, vol. 55, no. 10, pp. 6150–6158, 2014.

[88] P. E. Jeter, C. Rozanski, R. Massof, O. Adeyemo, and G. Dagnelie, "Development of the ultra-low vision visual functioning questionnaire (ULV-VFQ)," *Transl. Vis. Sci. Technol.*, vol. 6, no. 3, p. 11, Jun. 2017.

[89] D. R. Geruschat, M. Flax, N. Tanna, M. Bianchi, A. Fisher, M. Goldschmidt, L. Fisher, G. Dagnelie, J. Deremeik, A. Smith, F. Anaflous, and J. Dorn, "FLORA^{TM}: Phase I development of a functional vision assessment for prosthetic vision users," *Clin. Exp. Optometry*, vol. 98, no. 4, pp. 342–347, Jul. 2015.

[90] J. Dykema, N. C. Schaeffer, D. Garbarski, N. Assad, and S. Blixt, "Towards a reconsideration of the use of agree-disagree questions in measuring subjective evaluations," *Res. Social Administ. Pharmacy*, vol. 18, no. 2, pp. 2335–2344, Feb. 2022.

[91] J. K. Höhne, D. Krebs, and S.-M. Kühnel, "Measurement properties of completely and end labeled unipolar and bipolar scales in likert-type questions on income (in)equality," *Social Sci. Res.*, vol. 97, Jul. 2021, Art. no. 102544.

[92] N. C. Schaeffer and J. Dykema, "Advances in the science of asking questions," *Annu. Rev. Sociol.*, vol. 46, no. 1, pp. 37–60, Jul. 2020.

[93] A. DeCastellarnau, "A classification of response scale characteristics that affect data quality: A literature review," *Qual. Quantity*, vol. 52, no. 4, pp. 1523–1559, Jul. 2018.

[94] D. F. Alwin, E. M. Baumgartner, and B. A. Beattie, "Number of response categories and reliability in attitude measurement," *J. Surv. Stat. Methodol.*, vol. 6, no. 2, pp. 212–239, 2017.

[95] J. W. Beckstead, "On measurements and their quality. Paper 4: Verbal anchors and the number of response options in rating scales," *Int. J. Nursing Stud.*, vol. 51, no. 5, pp. 807–814, May 2014.

[96] J. Pasek and J. A. Krosnick, "Optimizing survey questionnaire design in political science: Insights from psychology," in *The Oxford Handbook of American Elections and Political Behavior*, J. E. Leighley, Ed., London, U.K.: Oxford Univ. Press, 2010.

[97] T. Yan, F. Keusch, and L. He, "The impact of question and scale characteristics on scale direction effects," *Surv. Pract.*, vol. 11, no. 2, pp. 1–10, Jan. 2018.

[98] A. Persson, "Response categories," in *Questions and Answers: About Question Construction in Questionnaire and Interview Surveys*, A. Persson, Ed. Stockholm, Eds., Stockholm, Sweden: Statistics Sweden, 2016, ch. 3, pp. 81–127.

[99] L. Fjelkegård and A. Persson, "Questionnaire methodology," in *Questions and Answers: About Question Construction in Questionnaire and Interview Surveys*, A. Persson, Ed., Stockholm, Sweden: Statistics Sweden, 2016, ch. 4, pp. 134–159.

[100] A. Persson, "data collection method," in *Questions and Answers: About Question Construction in Questionnaire and Interview Surveys*, A. Persson, Ed., Stockholm, Sweden: Statistics Sweden, 2016, ch. 6, pp. 203–259.

[101] G. Bluhm and S. Nilsson, "Methodology development of noise questionnaires for determination of annoyance," in *Questionnaire Surveys Regarding Noise Annoyance*, 1st ed., Karlskrona, Sweden: Boverket, 2012, pp. 31–45.

[102] F. Scheffer and A. Björnram, "To ask questions regarding noise," in *Questionnaire Surveys Regarding Noise Annoyance*, 1st ed., Karlskrona, Sweden: Boverket, 2012, pp. 47–71.

[103] B. Laugwitz, T. Held, and M. Schrepp, "Construction and evaluation of a user experience questionnaire," in *HCI and Usability for Education and Work*, A. Holzinger, Ed. Berlin, Germany: Springer, 2008, pp. 63–76.

[104] *UEQ Homepage*. Accessed: Jan. 4, 2023. [Online]. Available: https://www.ueq-online.org

[105] G. O. Boateng, T. B. Neilands, E. A. Frongillo, H. R. Melgar-Quiñonez, and S. L. Young, "Best practices for developing and validating scales for health, social, and behavioral research: A primer," *Frontiers Public Health*, vol. 6, pp. 1–18, Jun. 2018.

[106] F. F. R. Morgado, J. F. F. Meireles, C. M. Neves, A. C. S. Amaral, and M. E. C. Ferreira, "Scale development: Ten main limitations and recommendations to improve future research practices," *Psicologia, Reflexão e Crítica*, vol. 30, no. 1, pp. 1–20, 2017.

[107] E. Peter, R. M. Baiju, N. O. Varghese, R. Sivaraman, and D. L. Streiner, "How to develop and validate a questionnaire for orthodontic research," *Eur. J. Dentistry*, vol. 11, no. 3, pp. 411–416, Jul. 2017.

[108] S. Tsang, C. Royse, and A. Terkawi, "Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine," *Saudi J. Anaesthesia*, vol. 11, no. 5, pp. 80–89, 2017.

[109] H. Taherdoost, "Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research," *Int. J. Acad. Res. Manage.*, vol. 5, no. 3, pp. 28–36, 2016.

[110] *Matlab (R2021b Update 2)*, MathWorks, Natick, MA, USA, 2021.

[111] J. Cohen, "Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit," *Psychol. Bull.*, vol. 70, no. 4, pp. 213–220, 1968.

[112] J. L. Fleiss, J. Cohen, and B. S. Everitt, "Large sample standard errors of Kappa and weighted Kappa," *Psychol. Bull.*, vol. 72, no. 5, pp. 323–327, Nov. 1969.

[113] G. W. Snedecor and W. G. Cochran, *Statistical Methods*, 8th ed., Ames, IA, USA: Iowa State Univ. Press, 1989.

JOHAN ISAKSSON-DAUN was born in Jönköping, Sweden, in 1990. He received the M.Sc. and Ph.D. degrees in biomedical engineering from Lund University, Lund, Sweden, in 2016 and 2023, respectively.

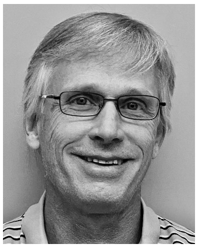
From 2017 to 2018, he was a Project Assistant and a Ph.D. Student with the Department of Biomedical Engineering, Lund University, from 2018 to 2024. Since 2024, he has been a postdoctoral researcher position with the Division of Design and Human Factors, Chalmers University of Technology, Sweden. His research interests include human factors, human–computer interaction, signal and image processing, computer vision, deep learning, and rehabilitation, in particular for assistive device applications, and especially in regards to mobility of users with blindness or low-vision.

Dr. Isaksson-Daun is a member of Swedish Society for Medical Engineering and Föreningen för Synrehabilitering. He has been a main recipient of awards, scholarships, and research grants from eight different organizations, foundations, and agencies for his work with mobility aids for people with blindness or low-vision.

TOMAS JANSSON (Member, IEEE) was born in Hagfors, Sweden, in 1967. He received the M.Sc. and Ph.D. degrees in electrical measurements from Lund University, Lund, Sweden, in 1993 and 1999, respectively.

After periods as a Research Assistant with the University of Rochester, Rochester, NY, USA, and a postdoctoral researcher position with Linköping University, Sweden. In 2010, he became an Associate Professor, and a full professor position with Lund University, in 2018. His research on magnetomotive ultrasound resulted in the spin-off company NanoEcho, where he was the first CEO, until 2019, and then served on the company board, until 2022. His research interests have concerned ultrasound based methods to measure flow, contrast agents, blood vessel characterization, visual aids for blind and visually impaired people, deep learning for interpretation of ultrasound images, and magnetomotive ultrasound.

Prof. Jansson was until recently the treasurer of Swedish Society for Medical Engineering and has been co-organizer of several international and national conferences.



JOHAN NILSSON (Member, IEEE) was born in Malmö, Sweden, in 1961. He received the M.Sc. degree in electrical engineering and the Ph.D. degree in electrical measurements from Lund University, Lund, Sweden, in 1987 and 1993, respectively.

In 1993, he became a Research Fellow with Lund University followed by an assistant professor position, in 2000. In 2003, he got an associate professor position with the Department of Electrical Measurements. Since 2014, he has been with the Department of Biomedical Engineering, Lund University. He is the author of 87 journal publications, more than 100 conference publications, and involved in a number of patent applications. His research interests include found within the areas of microfluidics with a focus on droplet and acoustofluidic applications, sensors for food chain surveillance, and electronic aids for blind and visually impaired people.