Towards Inclusive External Communication of Autonomous Vehicles for Pedestrians with Vision Impairments

Mark Colley, Marcel Walch, Jan Gugenheimer*, Ali Askari, Enrico Rukzio
Institute of Media Informatics, Ulm University
Ulm, Germany
[firstname.lastname]@uni-ulm.de

ABSTRACT

People with vision impairments (VIP) are among the most vulnerable road users in traffic. Autonomous vehicles are believed to reduce accidents but still demand some form of external communication signaling relevant information to pedestrians. Recent research on the design of vehicle-pedestrian communication (VPC) focuses strongly on concepts for a non-disabled population. Our work presents an inclusive user-centered design for VPC, beneficial for both vision impaired and seeing pedestrians. We conducted a workshop with VIP (*N*=6), discussing current issues in road traffic and comparing communication concepts proposed by literature. A thematic analysis unveiled two important themes: *number of communicating vehicles* and *content* (affecting duration). Subsequently, we investigated these in a second user study in virtual reality (*N*=33, 8 VIP) comparing the VPC between groups of abilities. We found that trust and understanding is enhanced and cognitive load reduced when all relevant vehicles communicate; high content messages also reduce cognitive load.

Author Keywords

Autonomous vehicles; external communication;
inclusive design research; vulnerable road users; accessibility.

CCS Concepts

•**Human-centered computing** → **Accessibility;**

INTRODUCTION

Autonomous vehicles (AVs) are expected to change interaction between pedestrians and vehicles [30]. There is no need for a human driver to be present in AVs. Communication for situations, in which people today rely on eye-contact or gestures [61] will therefore be even more challenging. Recent research projects aim to overcome these upcoming challenges through external communication modalities such as displays [32], LED strips [32,49], movement patterns [93], projections [2] auditory or tactile cues [51], as well as combinations thereof [51] and enhancement of the infrastructure [72].

*now at Télécom Paris/IP-Paris

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376472

Most research projects focus mainly on middle-aged non-disabled people and thereby exclude a majority of the population with disabilities (e.g., currently exist around 1.3 billion people with some, 217 million people with moderate to severe vision impairment and 36 million people are blind [9,59]; 466 million people have disabling hearing loss [60]). One possible approach is applying universal [19] or inclusive design principles [89] to create concepts for external communication that target a larger set of the pedestrian population [2,22,32,51]. To the best of our knowledge, we are the first to actively include people with vision impairments (VIP) into the design process of external communication for future AVs.

Our main contributions are: (1) Unveiling the lack of inclusive design in the field of external communication of AVs through an in-depth literature analysis, evaluating proposed auditory external communication concepts according to Universal Design (UD) guidelines. (2) Identification of relevant themes (*number of communicating vehicles* and *information content*) via a workshop (*N*=6) with VIP exploring their current issues with participating in road traffic and comparing four external communication approaches from literature. (3) Proposal of a novel communication approach: the single "omniscient narrator". (4) Findings of a second study in virtual reality (VR) (*N*=33) showing that high content messages (that also results in longer messages) are clearly preferred and that these, in combination with multiple vehicles communicating, reduce cognitive load. Explicit communication was clearly rated reasonable and necessary. High content messages improve comprehension and two vehicles communicating improve perceived safety.

RELATED WORK

We review approaches and technologies to include VIP in research with special focus on road traffic research.

Assisting VIP: Methods and Technology

Existing assistive technology (AT) in navigation ranges from low-tech tools such as white canes or guide dogs to high-tech applications such as *Landmark AI* [66]. Research has focused on aiding VIP in traffic either by recognizing the state of the traffic light [43], by creating in- and outdoor navigation systems [39,88] and by exploring obstacle avoidance [35]. While specialist devices are helpful, some of them draw unwanted attention, marking VIP as "different" [71]. These solutions try to target a perceived disability, however, Moser [55] describes that disability is constructed through the social and material environment, therefore allocating responsibility for addressing disability on everybody [11,80]. Recent work questions the goal of independence for VIP [6] and highlights the *interdependence* of *all* people. UD [19] was designed to guide the design of products and infrastructure to be used by any person with any disability. However, disabilities are diverse and a UD based product may not be sufficient or pleasing for some [28]. An inclusive design may therefore be adapted via context or the ability of the user. User Sensitive Inclusive Design [57] aims to include people with disabilities in the design and evaluation process [89]. AT research today seems to focus more on social aspects such as social acceptability [80], accounting for the various methods and skills of VIP [86]. Williams et al. [86] explored navigation of VIP, show differences to seeing people and the inter-connectedness to other pedestrians. They show important properties of communication from other (seeing) pedestrians such as *timing* (e.g., telling VIP to turn "soon") and ambiguousness of communication, missing information or misjudging challenges (e.g., open spaces). Williams et al. report that the "sighted guide" technique is the most preferred mode of walking with sighted people as verbal cues can be omitted.

People with Vision Impairments in Traffic Research

Wiener and Lawson [85] investigated traffic sounds as a means to make crossing decisions. They found that VIP have significant information limitation with auditory information only. While the sound of a vehicle is indicative of its movement, it only allows for a very rudimentary perception of spatial layout structure. Ashmead et al. [4] compare crossing behavior of blind and sighted people at a roundabout. They found that VIP need longer to cross and that they wait approximately three times as long before crossing. VIP also face riskier situations. Arfaoui et al. [3] investigated the role of risk in the daily life of people with impairments. Through semi-structured interviews with seven people of various disabilities, they found that they experience risk at a significant higher level. They all found ways to manage risky situations, but feel the need for safety measures. Hassan [38] found that blind people, compared to sighted and people with minor visual impairments, often overestimate gap times between vehicles. No significant differences were found between normally sighted and people with minor visual impairments.

EXTERNAL COMMUNICATION CONCEPTS

The third pillar of relevant related work is external communication of AVs. We focused on an in-depth analysis of external communication concepts in academia extending the overview on vehicle-pedestrian interaction with both traditional and AVs presented by Rasouli and Tsotsos [62] with special regard to communication modalities and messages. This is followed by an analysis of auditory concepts with special regard on information content and duration. Therefore, we describe our method in a more rigorous manner. Concepts from industry were analyzed by Bazilinskyy et al. [5] with focus on visual features such as color.

Data collection. Relevant publications in the field of external communication are spread over various journals and conferences. Therefore, the publication databases ScienceDirect (SD) and Google Scholar (GS) were screened. A time frame of five years (2014-07/2019) was defined as the field of external communication grew significantly in the last years [2].

Search procedure. In both databases, the terms *external communication/features*, *autonomous vehicle*, and *pedestrian* were combined (search query: `(External Communication OR Features) AND (autonomous vehicle) AND pedestrian)`. In total, 433 publications were found (SD=384, GS=49).

Screening criteria. Four criteria were defined to narrow down the search. The papers had to be about (1) AVs intended for streets (no water based vehicles, e.g. [67]). Additionally, (2) external communication with pedestrians had to be investigated, not, e.g., communication with other vehicles. The publications included for analysis had to be (3) original full papers or Work in Progress (due to the novelty of the field) and had to be (4) written in English or German. If the paper had references that matched the four criteria, these papers were added. Additionally, `researchgate.net` was screened until end of 08/2019 for all first authors found by querying SD and GS. This resulted in a sample of **33** publications (cf. Table 1).

used modalities	**number of publications** + [references]
visual only	**23**: [2, 13–17, 20, 21, 25, 26, 33, 46, 47, 49, 54, 64, 72, 73, 77, 83, 90, 91, 93]
visual + auditory	**7**: [7, 22, 23, 32, 40, 52, 53]
visual + tactile	**1**: [41]
visual + auditory + tactile	**2**: [50, 51]

Table 1: Categorization of publications based on modality.

Overview: Visual, Auditory & Tactile Concepts

Visual concepts: All concepts included some visual cue, in fact, most work (23 publications) has focused solely on visual cues in their external communication concept. The communication has been simulated or implemented using various devices such as LED strips [32, 49], displays [32] or projections [2]. Further proposed concepts use an enhanced street as a kind of display [72] or suggest to shape-shift the hood of the vehicle [26]. Some concepts also include the smartphone of the pedestrian [51] or physical objects such as a waving hand mounted on the vehicle [51]. These communication cues, according to the design space by Mahadevan et al. [51], are located mostly on the vehicle. In these concepts, different types of messages are conveyed. *Intent (e.g., "I 'm about to yield")* [46,50], *Awareness (e.g., "I see you")* [50], and *Directive (e.g., "Cross")* [51] are reported most often. Other concepts provide the pedestrian with information about the status of the vehicle (automation mode or current driving speed [17]) or warn the traffic participants of perils (e.g., a child running onto a street) [72]. Mostly, the proposed concept included a way to signal that the vehicle will yield, that the pedestrian can or cannot cross and that the vehicle has detected the pedestrian [51]. Published work differs in their findings. *Text* was reported as least ambiguous [14, 21]. Comparing projections with text and symbols at various locations and different content, Ackermann et al. [2] found that large scale text projections with advice on the windshield were preferred.

Auditory concepts: Nine concepts included an auditory component (cf. Table 1). The auditory component was used to gain attention of the pedestrian via music [32], to indicate that the vehicle will start driving [7] or that it is safe to cross via a horn sound [7] or spoken text [50, 51]. Florentine et al. found that pedestrian engagement was increased with audio cues [32]. Deb et al. [22] report that the verbal message was rated best in comparison to a horn sound, music or no sound by the participants. This is supported by work from Hudson et al. [40] who found that verbal messages are preferred over music. Merat et al. [53] performed a study in three European cities with the CityMobil2. There were five messages under investigation: whether it is stopping or turning, its pace, if it has the intent to start moving, and whether it has detected the pedestrian. There was no clear agreement on the preferred modality across sites. Information about the vehicles' speed was rated least important (by the seeing participants).

Tactile concepts: Mahadevan et al. [50, 51] included tactile feedback in some of their concepts using a smartphone. Vibrating indicated that it is safe to cross. The work seems not to have focused on tactile feedback, however, the authors ask how many modalities are too many but are not able to give a definite answer [51].

Analysis: Information Content & Duration

Based on the UD principles [19], we evaluated the information content of the auditory concepts *[Principle 4a: Use different modes (pictorial, verbal, tactile) for redundant presentation of essential information]* [19]. Being an important factor, we also focused on the duration of messages [51] (repetition being a factor to duration). Work by Matthews et al. [52] is excluded from this analysis because they only report a mounted speaker but no information on sounds or messages.

Information content: Florentine et al. [32] used music to gain the attention of pedestrians. In their visual component, they showed destination and state. The auditory information content was therefore lower. Hudson et al. [40] used music as well as a verbal message ("safe to cross"). This verbal message was preferred. As the preferred visual cue was the text "Walk", the information in the auditory cue is comparable to the visual cue. Deb et al. [22] also used this message as these projects were conducted within the same research group. Merat et al. [53] report that only participants in Trikala, Greece rated signals and spoken words equally. In contrast, spoken text was preferred in La Rochelle and Lausanne. Each message in [53] was given as text, a light signal, a spoken word or an auditory signal. We assume that the information content was therefore equal. Deb et al. [23] compared five external communication features, two of which were audible (mild horn sound, three beeps). Visual features were preferred. The features differed in their meaning ("walk" sign, break light). Audible signals were used as warning signals. Therefore, they differed from the visual cues. Böckle et al. [7] used LED columns to indicate four messages: *not stopping*, *stopping*, *waiting*, and *start driving*. Only the "start driving" message was indicated through a bell sound. The information content is therefore significantly less in this work when comparing auditory and visual cues. Mahadevan et al. [50, 51] used spoken text. In their first design, they used a four times repeated one word message (e.g., "cross" [51]) which was then shortened to a message played once. This message was also used in their second work [50] with a clearer voice. The information content was therefore equal to the used visual cue, which was e.g., the same text on a display.
Of the eight investigated concepts, five provided equal information content in both the auditory and the visual component. The information mostly given was that a crossing was possible.

Duration: Mahadevan et al. [50, 51] used single words as spoken text. Merat et al. [53] report the auditory signal as "spoken word", therefore we assume that single words were used as an auditory signal. Hudson et al. [40] played music (no duration given) as well as the text "safe to cross" (cf. Deb et al. [22]). Böckle et al. [7] do not provide information about the duration of the bell sound. The mild horn sound and the three short beeps used by Deb et al. [23] are considered as messages of low duration. Attention gaining auditory concepts were (probably) continuously played [32].
Most explicit communication was played once [7, 22, 23, 40, 50, 51, 53] or continuously [32, 40]. Mahadevan et al. [51] repeated the message in their first prototype four times with no pause between each repetition. Deb et al. [23] repeated the short beep three times as a warning. No information on pauses is given. In summary, almost all (except the *three short beeps*) final concepts reported were played once or continuously. Single words were mostly used (three times).

This analysis shows that most work focuses on visual external communication. While some publications actually do mention *accessibility* or *impairment* [2, 46, 49, 51], none addresses this topic. Work on auditory cues seems underrepresented and not systematically investigated.

CONCEPT EVALUATION AND REQUIREMENT ANALYSIS

We conducted a workshop with 6 members of a German supra-regional association for VIP to (1) evaluate concepts of external communication derived from current literature, (2) to explore which metrics are important for VIP, and (3) to gain a better understanding of current challenges VIP face when participating in road traffic.

Apparatus

To evaluate the concepts, we implemented a VR simulation using Unity [82]. A visualization of a vehicle approaching a crosswalk was developed as not all participants of the workshop are fully blind. The Western setting (English texts) with some pedestrians on the sidewalks is depicted in Figure 1. The simulation also provided typical background noise like chatting people and engine sounds. The simulation enabled us to track the position and orientation and therefore allowed a realistic sound experience. As the vehicle arrived at the crosswalk, it produced a sound or made the smartphone vibrate (cf. 'Evaluated Concepts'). The participants heard this sound through a surround sound headphone, i.e., Sony MDR-HW700DS (cf. Figure 2a).

Figure 1: VR simulation used in the workshop and study. (a) The view straight ahead and (b) the view to the left.

Evaluated Concepts

We used eight auditory concepts of the analysis and the concepts with a tactile component [41, 50, 51] as basis for our evaluation with VIP (cf. Table 1). We formed categories of the auditory concepts based on their frequency of use and the classification used by Merat et al. [53] as shown in Table 2. For the evaluation, one representative of each category (bold in Table 2) was chosen. We chose concepts with explicit messages like "I'm stopping", therefore excluding attention seeking concepts like music as these seem of little help for VIP. We aimed to include various types of message, i.e., intent ("I'm stopping") vs. awareness ("I see you") vs. directive ("Walk", "Safe to cross"). As there were only two artificial sounds (bell vs. music) we chose the bell indicating a different kind of message (starting to drive instead of conveying that it is safe to cross) as music is used as an attention seeking concept [22, 32, 40]. We used the original sounds from [50] and [7] as well as the Acoustic Vehicle Alerting System (AVAS) sound [1] provided by BMW [24]. The sound effect on traffic lights was also included as it is a well known communication mechanism and an important cue for wayfinding of VIP [45].

Procedure

The workshop consisted of three phases: *introduction*, *evaluation* and an open *discussion*. During the whole session, audio was recorded. The workshop started with an *introduction* of the participants and the research field of the workshop organizers. The goal of the workshop, getting to know the difficulties of VIP in traffic and possibilities of communication from AVs to overcome these, was explained. UD [19] was then introduced. Following this, one participant at a time was separated from the rest to experience and evaluate the chosen concepts in the VR simulation (cf. Figure 2a), presented in counter-balanced order. Participants were instructed to imagine standing at a street waiting to cross. Participants were asked about associations, perceived trust, and whether they would cross the street if presented with this signal. The remaining participants meanwhile were engaged in a group discussion (cf. Figure 2b). A walk-through of a typical journey via foot was described by two participants. Demanding situations in today's journey were discussed afterwards. Based on these scenarios, communication possibilities were discussed, especially how AVs could enhance these. After all participants had experienced the concepts, we initiated a group discussion. The workshop lasted about 3.5h, 20min for the *introduction*, 2.0h for the *evaluation* and 1.0h for the *discussion*. Experiencing the communication concepts lasted about 20min per participant.

Participants

ID	Gender	Age	Vision	Since	In %	Aids
P1	F	53	PB	birth	2	tandem bike
P2	M	55	FB	1991	0	guide dog, private vehicle
P3	M	56	FB	1985	0	
P4	M	50	FB	1974	0	guide dog
P5	M	45	FB	birth	2	
P6	M	n.a.	S	n.a.	n.a.	wheelchair

Table 3: Demographic information of workshop participants who are seeing (S), partially blind (PB), and fully blind (FB). All but [P6] use a *walking cane* and *short-range public transportation*.

We recruited N=6 participants (cf. Table 3). Participants were on average M=51.8 (SD=4.44; range: 45 to 56) years old and were, with the exception of the seeing participant [P6], with a visual impairment for at least 28 years. [P6] is an architect with special focus on accessibility both for wheelchair users and for VIP. All participants are experts on accessibility.

Analysis

The discussions of the workshop and videos captured in the evaluation part of the workshop were transcribed by the first author. The

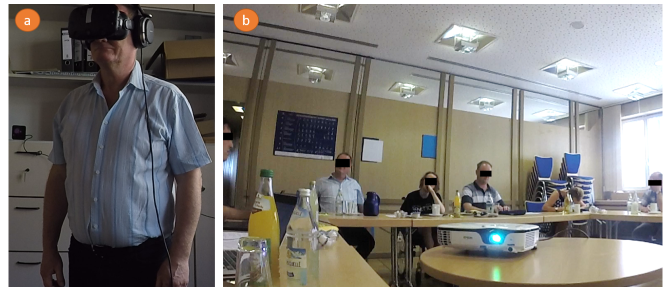

Figure 2: (a) A participant experiencing our setup for the evaluation in VR and (b) the setting of the group discussion.

timing of the participants' statements, interruptions as well as special focus on topics e.g., a special emphasize, were written down. According to Stol et al. [76], we used Straussian grounded theory. Following Furniss et al. [34], we differed from the original approach by conducting a thorough literature review (cf. 'External Communication Concepts'), which led to the definition of the research gap. Therefore, the first author already had extensive knowledge of the problem. However, the second researcher present was unknowing and therefore not biased. As the participants were split for the discussion and the evaluation, we believe that this method enabled us to have the benefits of both an unbiased look at the problem as well as needed knowledge about proposed concepts. We used open and axial coding for the evaluation and open coding for the group discussion. The coding was done by the first, second and fourth author (in accordance to work of Saldaña [68]). First, the transcription of the evaluation of concepts of the first participant was coded independently by three coders with in-vivo codes, as proposed in [34]. Subsequently, an initial code book was defined. Next, the second evaluation was again independently coded and, based on this coding, the code book was refined. Finally, the first and second author coded all interviews and the group discussion together. Conflicts were resolved via discussions among the researchers.

Results

Current Situation and Relevant Scenarios: [P4] and [P5] described their journey to the workshop location. Navigation, orientation and safety are the main concerns. Navigation starts with the planning process. Usually a tool (e.g., Google Maps) and making arrangements (e.g., [P4] "I call the service of the Deutsche Bahn 24h in advance") are involved. During the travel, orientation is a constant concern. For this, a standardized orientation system for blind people (OSBP) is the most important prerequisite. Still, e.g., after arriving by train, "true orientation" [P4] is missing as the OSBP can't provide the information e.g. about the correct exit. Reliance on seeing people is therefore often inevitable. Receiving such relevant information in situ increases the sense of safety. Infrastructure and traffic often fail to explicitly provide this information (e.g., through announcements). Therefore, "enormously high concentration" [P5] is needed to extract the information from all sounds, leading to arduous and unsafe journeys ("no risk no fun" [P1]; cf. [87]). Therefore, VIP "[...] try to work with their residual vision" [P1].

Need of Explicit Communication: There is agreement (6 of 6) among the participants that additional explicit communication of

Category	Paper	Characteristic	Message	Rating	Would cross
Spoken Text	**[50]** [22,40,51,53]	male voice	"stopping", "I see you"	*M*=3.40, *SD*=1.82	3/5
Conventional Sounds	**[22]** [23,53]	horn sound	"Walk"	*M*=1.80, *SD*=.84	0/5
Artificial Sounds	**[7]** [22,32,40]	bell	start driving	*M*=1.40, *SD*=.55	0/5
Tactile	**[50]** [41,51]	vibration	"safe to cross"	*M*=2.00, *SD*=1.75	1/5
AVAS	-	-	movement characteristic	*M*=3.20, *SD*=2.05	2/5
Traffic light sound	-	-	Safe to cross	*M*=4.00, *SD*=2.00	3/5

Table 2: Concepts under evaluation in workshop. Concepts under the dashed line were chosen with no equivalent in literature.

an AV would be beneficial contrary to some work indicating that movement and sounds are enough to communicate intent [27]. [P1] stated "I'd actually like to have it [explicit communication], be it at a crosswalk or elsewhere. [...] I think it's good." [P3] added "It depends on the ambient noise. What do you notice? If you can hear the cars clearly, it's okay. But there is often a jackhammer, loud music." Based on the journey description and the evaluation of the concepts (cf. Table 2), we derived the needs of VIP for the communication of external vehicles: [P1] stated "Orientation is one thing and road safety is another. Two aspects that we always have to think about at the same time." Road safety refers to situational awareness as defined by Endsley [29]. Perceiving relevant information and being able to derive assumptions for the future increases the perceived safety of VIP. Participants mentioned that details like vehicle size are not important aspects for their awareness: "[...] they're all road users where I'm the weaker one. That would not be decisive for me." [P4]. We therefore refer to the occupancy grid metaphor, meaning that it is important to know that something is there but it does not matter what exactly it is. However, as there is the constant threat of having missed important clues — risk is constantly perceived. Concluding, we found that the two major desires of the participants are *Awareness* and the resulting sense of *Safety*.

Evaluation of Current Concepts: The spoken text was best received. Conventional as well as the artificial sound were, without the visual context, always perceived as warnings. Additionally, these surprising occurrences of short sounds frightened the participants. There was the concern that brief signals could be missed. Regarding the tactile feedback on the smartphone, there were worries that the feedback could be mistaken for an incoming message (cf. *Distinctiveness*).
The traffic light sound was rated as a "good idea" [P4] but there are concerns about possible misunderstandings regarding the source of the signal: did a vehicle or a traffic light emit it, where latter means that it's "dead certain that you can cross" [P3].

Requirements on Implicit Communication: Participants referred to encounters with electric vehicles not yet equipped with an AVAS. [P4] stated "I was so scared when an electric car went past me. It came out of nowhere." Participants independently mentioned that the information that a vehicle is approaching is important to them. This indicates a mental model that represents the world including its moving vehicles. This is in line with experiencing fright when a sudden sound is played. Through the representation of the world, VIP derive expectations about the future, as proposed by Level 3 of Endsley's situational awareness model [29]. In the model in Figure 3, this is represented by *Predictability*. This information is needed *continuously* and has to be *perceptible*, i.e., loud enough.

Requirements on Explicit Communication: In a special situation, e.g., at a crossing or when a bus stop is reached, explicit communication is necessary (cf. Figure 3). The most important aspect of this communication is *Unambiguousness* [86]. Ambiguity is omnipresent in current traffic, e.g., a honk can be a greeting, a warning, or a signal letting you pass. Participants clearly stated that *spoken text* is preferred given its unambiguousness, i.e., a message with a clear indication about what to do. This is in line with related work that found text on a display to be least ambiguous [14, 33]. Spoken text seems to be addressing even more people as illiterate children could understand this spoken text. [P5] requested his mother tongue as he is not capable of understanding English while [P6] argued English to be sufficient. One option mentioned is to announce in alternating languages. *Unambiguousness* was divided into three subitems in the workshop: *Standardization*, *Distinctiveness*, and *Perceptibility*.
Participants fear having to learn multiple communication patterns that different manufacturers might employ. While they do not object variations for example in pitch or voice, they want to have the same *standardized* signal, i.e., message, to guide them.
Distinctiveness refers to the ability to distinguish the auditory cue from others as a message relevant to traffic. This is in contradiction to work by Merat et al. [53] that found seeing individuals to prefer conventional sounds over spoken words emphasizing the differences across abilities. *Perceptibility:* Participants frequently mentioned that signals

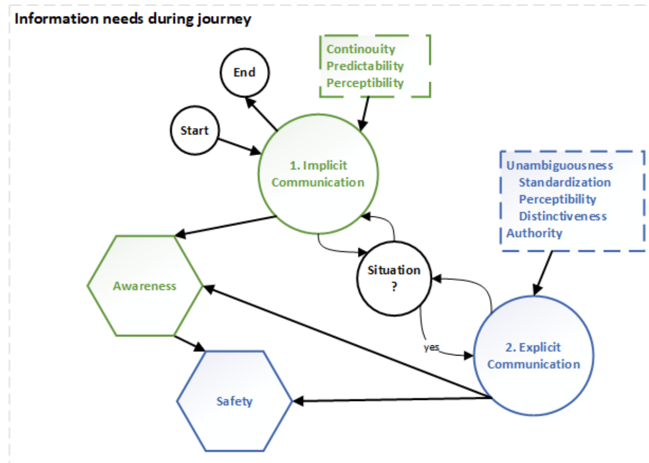

Figure 3: Communication needs during a journey of a VIP. During the entire journey *Implicit Communication* indicating the position and state of the vehicle, thus increasing *Awareness* is necessary. It has to be *continuously* available, *perceptible*, and allow *predicting* near future events. In situations where decisions have to be made, *Explicit Communication* is desired. This should be *unambiguous* and can have the *authority*, i.e., allow a crossing. Explicit communication provides *Safety* to the VIP as well as *Awareness*, which itself provides some (but not enough) *Safety*.

should follow norms such as the two-sense principle [63]: at least two senses have to be stimulated for a signal to be perceptible. Further recommendations for the auditory sense are appropriate *volume* and *duration* of the sound: "one word is not enough for me because you may not be prepared for it, 2-3 words would be better" [P1]. This is in contrast to the findings of Mahadevan et al. [51], where the seeing study participants favored brevity and close to the "safe to cross" message by Hudson et al. [40]. Repeating is possible, but there should be pauses in-between to allow the perception of ambient noise (again in contradiction to [51] where the repeating first prototype had no pauses). Regarding tactile feedback, the smartphone seems inappropriate because it is often carried in a backpack and if not "one hand is no longer free as we need it for our mobility [for the guide dog or a cane]" [P4]. While the participants would be happy to have tactile feedback with a special pattern integrated into the cane, it is a general theme among VIP that it should not always be the person with the disability having to adjust ([P4] "Why should the affected person now again equip oneself with other technology and thus be restricted?") Another theme of explicit communication was *Authority*. As VIP rely on seeing people and follow their advice/directive, it was mentioned that the vehicle could fill out this role. This deviates from communicating intent and awareness [51] as the vehicle has the authority and decides for the blind pedestrian. The final topic mentioned is *scalability*. The question how to handle situations with many cars arose: [P4] "How are people supposed to know what to do with so many cars out there these days?" Explicit communication (cf. Figure 3) leads to increased awareness as it provides information about the location of the vehicle (implicit) and, depending on the message, information about the intention (e.g., "stopping") and the appropriate action of the pedestrian (e.g., "you can cross"). Additionally, it provides safety as clear evidence is given, which does not have to be derived of e.g., engine noises.

Situation Awareness: While some participants mentioned the need for types of communication that allow for the perception of ambient noise providing additional background information, all participants agreed that the preferred message was **"I'm stopping, you can cross"**. This message should be played by **one** vehicle that has knowledge about the whole traffic situation (e.g., via Vehicle-to-everything(V2X)-technology). If a manually driven vehicle is approaching, this message should be shortened to **"I'm stopping"** since the actions of the human driver cannot be predicted by the AV. The length of the message was emphasized, arguing that "one word is not enough for me because you may not be prepared for it, 2-3 words would be better" [P1]. This choice of preferred message shows that detailed situation awareness is not the main goal of VIP. According to Endsley [29], perception leads to comprehension leading to projection. With this, a decision can be derived. In this case, the decision is "outsourced" to an omniscient narrator, the vehicle/traffic system, which again refers to the *Authority* theme.

Communication Modality: All participants stated that the two-sense principle [63] has to be applied in this case. While the auditory cue is essential, it was mentioned that it should be possible to choose to additionally use a tactile cue. For this, the cane was regarded as feasible. Most VIP use this aid and therefore have only one free hand resulting in the wish to not use the other for a special device or the smartphone ([P4] "Because then you are already restricted again because a hand cannot be used"). It was also highlighted that visual cues are also essential for many VIP: [P1] "Most try to work with their residual vision." [P4] "[...] use a contrast-rich design and diversity." These contrasts [36] are primarily important for the luminance contrast. Turquoise seems to be a promising colour for AVs [84].

Locus of Communication: Regarding the location of the communication ([51]: vehicle, infrastructure, or personal device), the signal was preferred to be emitted directly by the vehicle. There were concerns about vandalism on infrastructure and problems with the connection necessary for V2X-technology. Additionally, [P5] mentioned communication at unsecured crossings where such infrastructure could be missing. [P4] stated that at crossings with traffic light sounds ancillary sounds would be irritating.

Workshop Setup: Participants agreed that the setting was appropriate. They were especially content for being involved in the process. [P1] stated "It's a good thing we're at the table; that it's done with us, not over our heads.". For the participants with residual vision, using a VR-headset was reported to increase immersion. Fully blind participants did not mind wearing a VR-headset.

In summary, the workshop highlights the needs of VIP in traffic situations. Based on their lack of visual stimuli, this population is under constant perceived and experienced risk. VIP and people with other disabilities have found ways to deal with these limitations but need further aid [3]. AVAS already supports the detectability of electric and hybrid vehicles, which are otherwise much later recognized [65]. The status quo, however, is not sufficient for VIP. We therefore strongly argue for explicit communication to aid this population. We also unveiled differences across modalities (e.g., duration) which emphasize the relevance of including the affected, as they mentioned themselves. While duration was explicitly mentioned, we argue that the preferred message actually is *longer* **and** has a *higher information content*. In the study, we therefore vary information content and discuss this discrepancy in section 'Does Duration or Information Content Matter?'

STUDY

To evaluate the concepts based on the workshop results compared to earlier proposed concepts, we designed and conducted a between-subject study with N=8 VIP and N=25 seeing people. This study was guided by the research questions (RQ):

RQ1: What impact do the variables "information content" and "number of communicating vehicles" have on pedestrians in terms of (1) dominance, (2) cognitive load, (3) trust, (4) preference, (5) confidence, and (6) appropriateness?
RQ2: To what extent do the ratings between VIP and seeing participants differ?

Apparatus

We used the VR simulation of the workshop, modified the evaluated sounds/messages, and added waypoints (cf. Figure 1a) that had to be reached. As VIP "can't find their way around when there are no edges" [P6 of the workshop], we fixated quarter rounds (as haptic feedback [92]) on the floor to mark the edge of the curb [86]. Due to space constraints, we added a gain factor in the straight forward and sideways (not height) axis. While a gain of 1.26 was found to be unnoticed [75], we employed, for the seeing participants, a gain of 2.0 ("Seven-League Boots"; meaning 1m

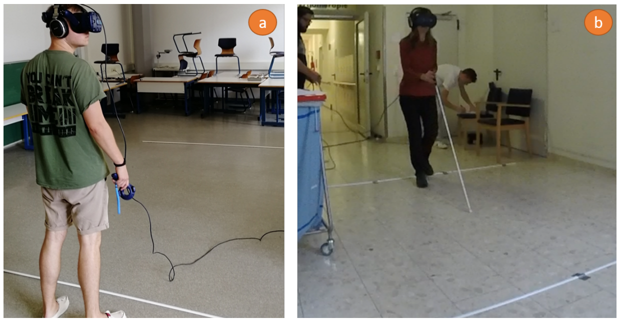

Figure 4: Study setting (a) for seeing participants and (b) for VIP.

in reality equals 2m in VR) due to space and tracking constraints. For VIP, this was 3.0 (as we used their premises). After the trials, we asked participants whether they noticed this gain. Most hadn't noticed it, but said that in retrospective they are aware of it.

Procedure

Each participant experienced **five** conditions, a *baseline* with no communication and a 2 x 2 design (*number of communicating vehicles* with two levels: one vs. two and message *information content* with two levels: low vs. high; the *independent* variables). The latter resulted in four systems that directly resulted from the uncovered themes of the workshop: *number of communicating vehicles* and *information content*. The messages given were (translated from German): *1. "Cross"* and *2. "I'm stopping, you can cross."* All systems included an AVAS inspired by BMW [24].

Each session started with a brief introduction, signing of the consent form, and a demographic questionnaire. The five conditions were then presented in counterbalanced order. Participants had to cross the street twice (increasing embodiment [18]) in each condition. For the VIP, the situation was described. Afterwards, they answered the questionnaires described below. At the end, participants gave general feedback. On average, a session lasted 45min. Participants were compensated with €8.

Measurements

Objective dependent variables: During each crossing, the system logged position, whether a car was in the center of the visual field (indicating that the participant might look at the vehicle), the lateral angle of the view (0° meaning straight ahead), and the current duration of the crossing in relation to when the auditory crossing signal was given. *Time on street* was calculated as the time from when the auditory signal was given until the waypoint was reached.

Subjective dependent variables: After each condition, we measured:

- affective state: on a 7-point semantic scale using the self-assessment manikin (SAM) [10]
- cognitive load: raw NASA-TLX [37] on a 20-point scale
- subscale *Predictability* and *Trust in Automation* of the *Trust in Automation questionnaire* by Körber [44]

Participants were also asked on individual 7-point Likert scales to what degree they felt safe, whether the crossings seemed risky, the vehicle seemed reliable and trustworthy, and if the communication was comprehensible and unambiguous. A single item misery scale [8] was used to assess simulator sickness.

After all five crossing sessions, participants rated their preferences regarding the systems from greatest (*ranking = 1*) to lowest (*ranking = 5*). Open questions regarding feedback and improvement proposals were also asked. For VIP, the experimenter read the questions and scales aloud.

Participants

ID	Gender	Age	Vision	Since	In %	Aids
P26	F	16	FB	2006	0	
P27	M	78	FB	1944	1	
P28	M	53	FB	1966	0	guide dog, private vehicle
P29	M	23	PB	1996	5	
P30	M	46	FB	2004	2	smartphone with navigation software
P31	F	70	FB	1949	0	— " —
P32	M	66	FB	1953	0	— " —
P33	F	62	FB	1956	0	— " —

Table 4: Demographic information of study participants. All participants used a *walking cane* and *short-range public transportation*. P1-25 are the seeing study participants.

We recruited 25 seeing people (15 male, 10 female) aged 21-56 (*M*=27.30, *SD*=8.62). 8 VIP were recruited (cf. Table 4). These participants were not part of the first workshop. On a 5-point Likert scale (*1=strongly disagree*, *5=strongly agree*), the seeing participants reported a high interest in AVs (*M*=4.00, *SD*=.91) and believed such a system to ease their lives (*M*=4.12, *SD*=.83). The non-seeing participants reported less interest in AVs (*M*=3.13, *SD*=1.13) but also believed such a system to ease their lives (*M*=3.50, *SD*=1.31). Seeing participants believed more in AVs to become reality by 2029 (10 years from today; *M*=3.72, *SD*=1.17) than VIP (*M*=2.25, *SD*=1.17). A Mann-Whitney U-test revealed a significantly higher interest in AVs by seeing participants (*W*=52.5, *Z*=-2.07, *p*=.04, *r*=.46) and that VIP significantly less believed AVs to become reality until 2029 (*W*=39, *Z*=-2.64, *p*=.008, *r*=.59). The *Propensity to Trust* subscale of the *Trust in Automation* questionnaire [44] was administered twice, once before and once after all conditions. *Propensity to Trust* was relatively low (overall: *M*=2.52, *SD*=.64; VIP: *M*=2.33, *SD*=.56; seeing: *M*=2.57, *SD*=.66). No significant differences between groups were found for the first measurement.

RESULTS

In the following, we report descriptive and inferential statistics. We used ANOVAs to compare the five systems. In addition, to investigate interaction effects of *eyesight* (between-group) x *number of communicating vehicles* x *content* (both within-group) and the non-parametric nature of the data [74], we disregard the *baseline* condition and use *nparLD*, which was shown to be a robust method even for small sample sizes [58] and unequal group sizes [56]. We report the ANOVA-type statistics. For post-hoc tests, we always used Bonferroni corrections.

Affective State

Participants' affective state in terms of dominance, introduced as "control over the situation" [12], was overall relatively high (overall: *M*=4.67, *SD*=1.63; VIP: *M*=4.40, *SD*=1.95; seeing: *M*=4.76, *SD*=1.52). The mean values for dominance varied between groups and systems. VIP experienced highest and seeing lowest in the *baseline*. Dominance was lowest for the seeing but highest

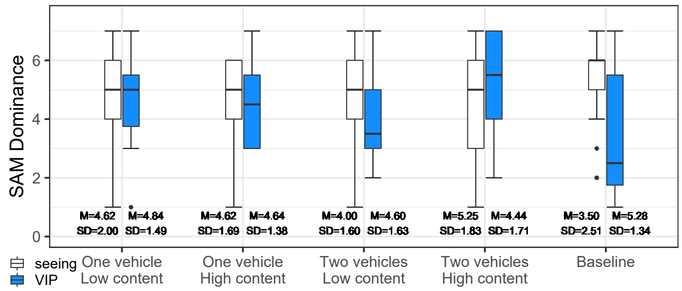

Figure 5: Values for dominance per condition.

for the VIP in the *two vehicles high content* condition. Ratings for dominance were close together for the other conditions (Figure 5).

A Friedman's ANOVA showed a significant difference in the mean ratings in dominance for the VIP (χ^2 (4)=12.4, p=.015). Post-hoc tests showed no significant difference. For dominance, the non-parametric variance analysis (*nparLD*) showed significant interaction effects for *information content* x *number of communicating vehicles* (F=4.25, df=1, p=.04; Figure 6a) and *eyesight* x *information content* (F=5.57, df=1, p=.02; Figure 6b).

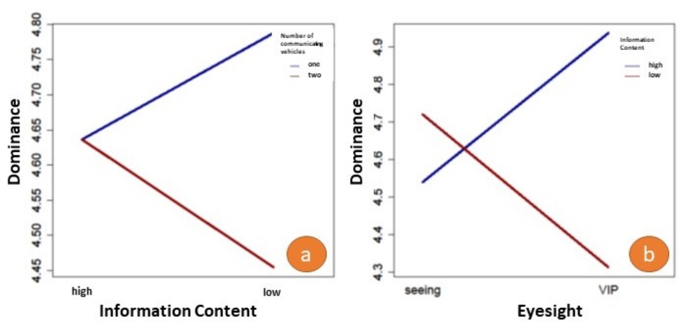

Figure 6: Interaction effects for dominance: (a) equal for high content independent of the number of communicating vehicles. With low content, dominance was higher when one vehicle was communicating. (b) Dominance for the VIP was low for low content and high with high content. Seeing participants reported similar dominance values for low and the high content.

Cognitive Load

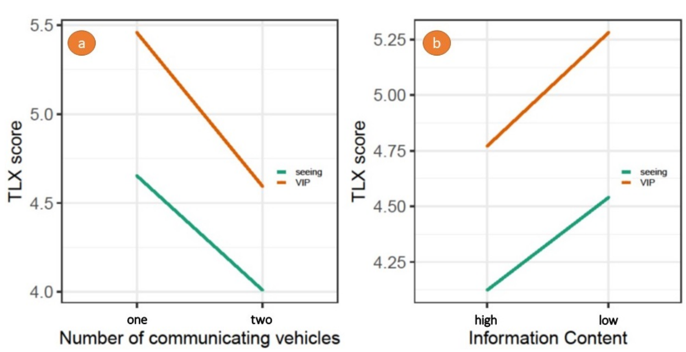

Figure 7: Main effects on TLX score. (a) Two communicating vehicles and (b) high content messages reduce cognitive load.

Cognitive load was measured using the raw NASA-TLX. Overall scores were low for all conditions (range: M=3.89, SD=2.43 *two vehicles high content* to M=6.31, SD=3.77 *baseline*).

A Friedman's ANOVA showed a significant difference in the mean overall scores (χ^2 (4)=36.6, p<.0001). Post-hoc tests showed that the *baseline* received a significantly higher TLX score compared to all other conditions but the *one vehicle low content* system. The same significant effect was also shown for the subscale mental effort (χ^2 (4)=41.4, p<.0001). The non-parametric variance analysis showed a significant main effect on the overall score of *number of communicating vehicles* (F=8.71, df=1, p=.003) and *content* (F=5.51, df=1, p=.02). A significant effect of *number of communicating vehicles* (F=8.42, df=1, p=.004) as well as *content* (F=1.37, df=1, p=.001) on the mental effort subscale could also be shown. Another significant effect of *number of communicating vehicles* on effort was shown (F=5.04, df=1, p=.02). With *high content* messages and multiple vehicles communicating, the overall TLX score and the subscores mental effort and effort are decreased (cf. Figure 7).

Predictability, Propensity to and Trust in Automation

A Wilcoxon signed-rank test revealed that, after the simulation, overall the participants had a significantly higher *Propensity to Trust* (M=2.76, SD=.30; W=104, Z=-2.47, p=.01, r=.55). The non-parametric variance analysis showed a significant main effect on *Trust in Automation* for *number of communicating vehicles* (F=4.44, df=1, p=.04), showing that the level *two* (M=3.98) received higher trust values than *one* (M=3.66). The *number of communicating vehicles* also had a significant main effect on *Predictability* (F=4.15, df=1, p=.04; two: M=3.38, one: M=2.99).

Head and Body Movements

Position and head orientation were logged with 5 Hz. The absolute differences in angles for the head orientation (indicator for head movement) were summed up for each participant and each condition. A Wilcoxon signed-rank test revealed that the seeing participants turned their heads significantly more (W=10841, Z=-8.41, p<.0001, r=1.88). Similar to Ashmead et al. [4], *time on street* was analyzed. Despite the higher gain factor in the VR simulation (cf. 'Apparatus'), time on the street was also non-significantly higher for VIP (M=13.63s, SD=4.87 compared to M=11.89s, SD=2.87) [4]. Assuming that the velocity remains the same, we adjusted the time on street for VIP by multiplying it with *3.0 / 2.0 =* ***1.5***. A Wilcoxon signed-rank test with the adjusted data revealed that the seeing participants crossed the virtual street significantly quicker (W=39, Z=-9.30, p<.0001, r=2.08). We also analyzed the acceleration, defined as change of speed per 0.2s (5 Hz), which was derived by the position values in the forward direction. A Wilcoxon signed-rank test revealed that the seeing participants in absolute values accelerated significantly more (seeing: M=.0104, SD=.0029; VIP: M=.0086, SD=.0034; W=3132, Z=-2.88, p=.004, r=.64). We interpret this as a higher change in speed, indicating multiple stop and go situations. VIP seem to, when they made their decision, cross the street without stopping.

Condition Preferences

The *two vehicles high content* system received rankings indicating the highest preference, i.e., the lowest mean (VIP: M=1.88, SD=1.64; seeing (M=1.88, SD=1.09). There was a clear ranking:

no explicit communication (*baseline*) was the least preferred followed by the single vehicle systems (the *high content* message was preferred). The system with *low* information in the *two vehicle* system was equally liked to the system with the *high content* message by the seeing participants (M=1.88, SD=.83). VIP rated this as their number two (M=2.38, SD=.92).

A Friedman's ANOVA showed a significant difference in the mean rankings both for the seeing (χ^2 (4)=56.9, p<.001) and for the VIP participants (χ^2 (4)=13.5, p=.009). Post-hoc tests showed that for the seeing, the *baseline* system was rated significantly worse compared to all other systems except the *one vehicle low content* system. The *one vehicle low content* system was also rated significantly worse than both *two vehicle* systems. For the VIP, the tests showed that the *baseline* system was rated significantly worse than the *two vehicles high content* system (*difference*=21; critical difference=17.75).

Safety, Reliance, Unambiguousness

The non-parametric variance analysis showed that participants felt significantly safer during the crossing when two vehicles communicated (F=11.59, df=1, p=.0007). In this case, participants also trusted the vehicle more (F=10.92, df=1, p=.001). Information content had a significant effect on comprehension (F=5.21, df=1, p=.02) and unambiguousness (F=9.42, df=1, p=.002) of the communication. High content enhanced comprehension and improved unambiguousness. This shows clear preferences for the communication: high content messages of both vehicles are better understood and perceived safety is increased.

Reasonable, Necessary, Appropriate

The mean value for the item rating the explicit communication as reasonable was high for the VIP (M=5.88, SD=1.55) and for the seeing participants (M=6.16, SD=1.11). Necessity was also rated high (VIP: M=6.13, SD=1.81; seeing: M=6.20, SD=1.04). Asked about the appropriateness of the communication, the high content (VIP: M=5.00, SD=1.93; seeing: M=5.12, SD=1.54) and the multiple vehicle communication were rated highest (VIP: M=5.50, SD=1.77; seeing: M=6.32, SD=1.03). Mann-Whitney's U tests found no significant effects in the difference in the responses to our 7-Likert scale questions.

Open Feedback

In the open feedback, most participants stated that explicit verbal communication enhanced their trust (20/29; 4 did not give feedback). While some argued that *low content* messages are "easier to understand" [P24], most agreed that high content messages are superior, as one has "double protection" [P7] and the "state of the system is clear" [P11]. Some problems were mentioned with the communication of two vehicles: [P14] encountered what he felt was an echo and [P6] only heard one message as both vehicles arrived at the exact same time. Improvement suggestions included repeating the communication [P2] and using sounds instead of verbal messages [P8]. Ten seeing participants mentioned the need for a visual signal.

DISCUSSION

Workshop and study participants agree on the need of auditory external communication of AVs as evidenced by their preferences and comments.

Does Duration or Information Content Matter?

The workshop revealed that the *duration* of communication is important. The workshop showed that this seems necessary to ensure a high perceptibility. While Mahadevan et al. [51] shortened their auditory messages as seeing participants found "that the audio messages were too long" [51], we showed that both two vehicles communicating systems were rated equally by the seeing participants. However, the VIP clearly preferred the long message. We conclude that while it is essential for the VIP to have a long message, the seeing participants are also content with this variation. The workshop participants stated that a 2-3 word or a repeating message would be preferable to the one word message, however, in the end they agreed on the longest message: they preferred the message with more information ("I'm stopping **and** you can cross"; representing intent + instruction) over a simple "you can cross" (only instruction) or "safe to cross"(only information) [40] message. The workshop showed that the longer *duration* is especially important for the *Perceptibility* (cf. Figure 3), the added information seems to provide relevant information about the state of the vehicle and therefore allows for a more accurate forecast. As (the seeing) [P7] stated, one has "double protection" with this message. Concluding, it seems that for the introduction of AVs, a longer message with more information content seems reasonable, which might be reduced in the future after some accommodation (helping to avoid disturbing concentration [86]).

Which Vehicles Should Communicate?

Workshop participants suggested to have an "omniscient narrator" that solely communicates with the pedestrian(s). This system was introduced by saying that "one can be sure that the other vehicle will stop after this message is played". We believe there are several reasons why this system was rated lower but was still well received. Firstly, we have encountered several participants that were, despite the introduction, not aware that the other vehicle will stop. Secondly, participants stated their distrust in the vehicles to communicate reliably. For the VIP, we argue that receiving information of multiple vehicles also increases their *Situational Awareness* (cf. Figure 3). Our findings suggest that while future autonomous traffic could be viewed as *one system* as it could be optimized through V2X-technology and algorithms optimizing traffic flow [78], people today think about every vehicle as a single unit. The effects of connectivity were unknown to the participants and distrust was still high, especially for VIP. It seems reasonable to assume that a transition time will be necessary.

Control and Annoyance

There were significant interaction effects for dominance (cf. Figure 6). Seeing participants reported relatively equal levels of dominance for the levels of *information content*, while VIP showed significantly higher levels for the high content message. *Explicit Communication* (cf. Figure 3) with additional relevant information seems to improve their perceived ability to master the situation. Dominance was also influenced by the interaction between *duration* and *number of communicating vehicles*. For high content messages, the level stayed the same. For low content messages, there was higher dominance with one communicating vehicle. This could indicate that the directive "Cross" emitted by two vehicles lowers one's own perceived control.

Need for Explicit Communication

While the seeing population could derive enough information of implicit communication of AVs, i.e., movement, this is not the case for VIP. VIP are already limited in information retrieval [85]. While an AVAS [1] is explicitly mentioned by the workshop participants and the study participants with vision impairment as necessary, additional explicit communication was requested. The seeing participants also reported their preference of explicit auditory communication in addition to visual cues (visual cues are significantly favored by seeing people over audible warnings [70]). As stated in the workshop and in line with Rasouli and Tsotsos [62], this communication should be standardized [42].

The Importance of Visual Cues

Pedestrians today rely on movement patterns of vehicles. While gestures and eye-gaze play a significant role [61], movement seems to be of even more importance [27, 69, 79]. Some seeing participants reported unease at the sight of a vehicle approaching and then breaking late and abruptly in various conditions — even though they were told to be able to rely on the message given by the vehicle in the *one vehicle* conditions. This indicates that visual cues are still important, which is in line with work by Mahadevan et al. [51]. Participants suggested an additional visual cue that indicates stopping from greater distance. In this case, a participant argued that "one vehicle could then talk for both vehicles" [P15]. Some participants also mentioned drawbacks of auditory communication. One participant felt that surrounding noise was modeled as too quiet, therefore not accounting for interference (Fiedler et al. [31] report noise level in a Brazilian city regularly exceeds 65 db(A)). Another participant encountered the situation that both relevant vehicles "spoke" at almost the exact same time, resulting in an echo which he felt irritating. Visual cues can reduce these short-comings, however, this is no solution for VIP.

Inclusive Design for Overall Beneficial Concepts

By including VIP in the evaluation and design process of VPC, we found relevant themes for this population. A novel approach was therefore systematically evaluated. While we found the *high content* messages to be best received, the approach of using one "omniscient narrator" was still well received. We found, in contrast to others who found that seeing people preferred shorter messages [51], that both seeing people and VIP preferred high content messages or rated them equally. We emphasize that this finding, which is beneficial for both groups, resulted from an inclusive design approach, showing the need to include various populations.

Usage of VR Simulation

We found comparable results for the crossing time difference between seeing and non-seeing participants between VR and real-life [4]. A major benefit of the VR simulation is that more data was logged and no dangerous situation could occur. This is especially important when working with a population with special needs. We showed that with careful consideration of the environment (cf. Figure 4b), VR simulations are feasible and well received by the population.

Practical Implications

The use of instruction as a communication message is discouraged by the current standard on external communication of AVs [42]. While this is recommended for a general population, we specifically looked into preferences of VIP. While addressing this population specifically seems unfeasible to avoid stigma [71], [P5] explicitly wanted this: "a signal [when] cane is recognized or my badge". One possible approach is to detect whether a VIP is trying to cross alone and then provide a signal, therefore avoiding stigma and addressing current mechanisms for navigation (cf. [86] "sighted guide" technique). It is, however, clear that an auditory signal is necessary for VIP and should not be too short to avoid missing it. The most important implication therefore is to include auditory cues in external communication concepts for AVs and not too focus solely on brevity [51].

LIMITATIONS

The size of the focus group was relatively small (N=6). However, this does not have to decrease validity [81]. As we included VIP in our study, we had to read the questions aloud for them. This was especially a limitation for the SAM [10], which relies on pictorial representations. Differences in the gain factor were necessary due to space constraints, but somewhat diminish the comparison of the results for *time on street* [4]. Due to the VR simulation, transferability to a real-world scenario is restricted. While the concept "omniscient narrator" addresses aspects of scalability in a pure AV traffic by limiting the number of AV communication participants, this concept has disadvantages especially in mixed and complex traffic, e.g., knowing the "whole traffic situation" seems unfeasible in complex mixed traffic. The aspect of scalability is still an open research question [48] and the proposed concepts need to be further investigated in the light of heavy traffic.

CONCLUSION & FUTURE WORK

Overall, our work aims to highlight the lack of inclusive concepts in the design of external communication for AVs and presents a user-centered design approach including VIP. We argue that populations across abilities all have individual and different concerns with road traffic and have to be considered in the design process to make AVs not only usable for middle aged non-disabled people. We show that the outcome of workshops has to be taken with caution and does not necessarily lead to the best rated systems as there might be unconsidered factors in situ. In a next step, we intend to test the systems in mixed traffic and scenarios with a higher density of vehicles, addressing the aspect of *scalability*. In the future, we want to evaluate these findings in various simulated settings such as multi-lane roads. We also want to evaluate tactile feedback to enable multiple modalities for VIP. While VR has its benefits, it is clear that these concepts must be investigated in a real-world scenario.

ACKNOWLEDGEMENTS

The authors thank all workshop and study participants. This work was conducted within the project 'Interaction between automated vehicles and vulnerable road users' (Intuitiver) funded by the Ministry of Science, Research and the Arts of the State of Baden-Württemberg.

REFERENCES

[1] Regulation No. 540/2014. 2014. Regulation (EU) No 540/2014 of the European Parliament and of the Council of 16 April 2014 on the sound level of motor vehicles and of

replacement silencing systems, and amending Directive 2007/46/EC and repealing Directive 70/157/EEC Text with EEA relevance. (2014).

[2] Claudia Ackermann, Matthias Beggiato, Sarah Schubert, and Josef F Krems. 2019. An experimental study to investigate design and assessment criteria: what is important for communication between pedestrians and automated vehicles? Applied ergonomics 75 (2019), 272–282.

[3] Afnen Arfaoui, Geoffrey Edwards, Ernesto Morales, and Patrick Fougeyrollas. 2019. Understanding risk in daily life of diverse persons with physical and sensory impairments. Journal of Accessibility and Design for All 9, 1 (2019), 66–89.

[4] Daniel H Ashmead, David Guth, Robert S Wall, Richard G Long, and Paul E Ponchillia. 2005. Street crossing by sighted and blind pedestrians at a modern roundabout. Journal of Transportation Engineering 131, 11 (2005), 812–821.

[5] Pavlo Bazilinskyy, Dimitra Dodou, and JCF De Winter. 2019. Survey on eHMI concepts: The effect of text, color, and perspective. Manuscript submitted for publication (2019).

[6] Cynthia L Bennett, Erin Brady, and Stacy M Branham. 2018. Interdependence as a frame for assistive technology research and design. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, 161–173.

[7] Marc-Philipp Böckle, Anna Pernestål Brenden, Maria Klingegård, Azra Habibovic, and Martijn Bout. 2017. SAV2P: Exploring the Impact of an Interface for Shared Automated Vehicles on Pedestrians' Experience. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct. ACM, 136–140.

[8] Jelte E Bos, Scott N MacKinnon, and Anthony Patterson. 2005. Motion sickness symptoms in a ship motion simulator: effects of inside, outside, and no view. Aviation, space, and environmental medicine 76, 12 (2005), 1111–1118.

[9] Rupert RA Bourne, Seth R Flaxman, Tasanee Braithwaite, Maria V Cicinelli, Aditi Das, Jost B Jonas, Jill Keeffe, John H Kempen, Janet Leasher, Hans Limburg, and others. 2017. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. The Lancet Global Health 5, 9 (2017), e888–e897.

[10] Margaret M Bradley and Peter J Lang. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry 25, 1 (1994), 49–59.

[11] Stacy M Branham and Shaun K Kane. 2015. The invisible work of accessibility: how blind employees manage accessibility in mixed-ability workplaces. In Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility. ACM, 163–171.

[12] Robin N Brewer and Vaishnav Kameswaran. 2018. Understanding the power of control in autonomous vehicles for people with vision impairment. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, 185–197.

[13] Christopher G Burns, Luis Oliveira, Peter Thomas, Sumeet Iyer, and Stewart Birrell. 2019. Pedestrian Decision-Making Responses to External Human-Machine Interface Designs for Autonomous Vehicles. In 2019 IEEE Intelligent Vehicles Symposium (IV). IEEE, 70–75.

[14] Chia-Ming Chang, Koki Toda, Takeo Igarashi, Masahiro Miyata, and Yasuhiro Kobayashi. 2018. A Video-based Study Comparing Communication Modalities between an Autonomous Car and a Pedestrian. In Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, 104–109.

[15] Chia-Ming Chang, Koki Toda, Daisuke Sakamoto, and Takeo Igarashi. 2017. Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, 65–73.

[16] Vicky Charisi, Azra Habibovic, Jonas Andersson, Jamy Li, and Vanessa Evers. 2017. Children's views on identification and intention communication of self-driving vehicles. In Proceedings of the 2017 Conference on Interaction Design and Children. ACM, 399–404.

[17] Michael Clamann, Miles Aubert, and Mary L. Cummings. 2017. Evaluation of Vehicle-to-Pedestrian Communication Displays for Autonomous Vehicles. Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles (2017).

[18] Mark Colley, Marcel Walch, and Enrico Rukzio. 2019. For a Better (Simulated) World: Considerations for VR in External Communication Research. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings (AutomotiveUI '19). Association for Computing Machinery, New York, NY, USA, 442–449. DOI: http://dx.doi.org/10.1145/3349263.3351523

[19] BR Connell, M Jones, R Mace, J Mueller, A Mullick, E Ostroff, J Sanford, E Steinfeld, M Story, and G Vanderheiden. 1997. The principles of universal design, Version 2.0, Raleigh, NC. Center for Universal Design, North Carolina State University (1997).

[20] Koen de Clercq, Andre Dietrich, Juan Pablo Nunez Velasco, Joost de Winter, and Riender Happee. 2019. External Human-Machine Interfaces on Automated Vehicles: Effects on Pedestrian Crossing Decisions. Human factors (2019), 0018720819836343.

[21] Shuchisnigdha Deb, Daniel W Carruth, Muztaba Fuad, Laura M Stanley, and Darren Frey. 2019. Comparison of Child and Adult Pedestrian Perspectives of External Features on Autonomous Vehicles Using Virtual Reality Experiment. In International Conference on Applied Human Factors and Ergonomics. Springer, 145–156.

[22] Shuchisnigdha Deb, Lesley J. Strawderman, and Daniel W. Carruth. 2018. Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment. Transportation research part F: traffic psychology and behaviour 59 (2018), 135–149.

[23] Shuchisnigdha Deb, B. Warner, S. Poudel, and S. Bhandari. 2016. Identification of external design preferences in autonomous vehicles. In Proc. IIE Res. Conf. 69–44.

[24] BMW Deutschland. 2016. Der BMW Vision M NEXT. (Sound by Hans Zimmer). `https://www.youtube.com/watch?v=vFS2HvlHTp4`. (2016). [Online; Accessed: 18-JULY-2019].

[25] Debargha Dey, Marieke Martens, Berry Eggen, and Jacques Terken. 2017. The impact of vehicle appearance and vehicle behavior on pedestrian interaction with autonomous vehicles. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct. ACM, 158–162.

[26] Debargha Dey, Marieke Martens, Chao Wang, Felix Ros, and Jacques MB Terken. 2018. Interface Concepts for Intent Communication from Autonomous Vehicles to Vulnerable Road Users.. In AutomotiveUI (adjunct). 82–86.

[27] Debargha Dey and Jacques Terken. 2017. Pedestrian interaction with vehicles: roles of explicit and implicit communication. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, 109–113.

[28] Addison Duvall. Accessed September 11, 2019. 5 Problems With 'Universal' Design. `https://www.hongkiat.com/blog/universal-design-problems/`.

[29] Mica R Endsley. 2017. Toward a theory of situation awareness in dynamic systems. In Situational Awareness. Routledge, 9–42.

[30] Daniel J Fagnant and Kara Kockelman. 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transportation Research Part A: Policy and Practice 77 (2015), 167–181.

[31] Paulo Eduardo Kirrian Fiedler and Paulo Henrique Trombetta Zannin. 2015. Evaluation of noise pollution in urban traffic hubs—Noise maps and measurements. Environmental Impact Assessment Review 51 (2015), 1–9.

[32] Evelyn Florentine, Mark Adam Ang, Scott Drew Pendleton, Hans Andersen, and Marcelo H. Ang Jr. 2016. Pedestrian notification methods in autonomous vehicles for multi-class mobility-on-demand service. In Proceedings of the Fourth International Conference on Human Agent Interaction. ACM, 387–392.

[33] Lex Fridman, Bruce Mehler, Lei Xia, Yangyang Yang, Laura Yvonne Facusse, and Bryan Reimer. 2017. To walk or not to walk: Crowdsourced assessment of external vehicle-to-pedestrian displays. arXiv preprint arXiv:1707.02698 (2017).

[34] Dominic Furniss, Ann Blandford, and Paul Curzon. 2011. Confessions from a grounded theory PhD: experiences and lessons learnt. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 113–122.

[35] Pedram Gharani and Hassan A Karimi. 2017. Context-aware obstacle detection for navigation by visually impaired. Image and Vision Computing 64 (2017), 103–115.

[36] Hilke Groenewold. 2009. DIN 32975 - Kontraste. `https://nullbarriere.de/din32975.htm`. (2009). [Online; Accessed: 13-JULY-2019].

[37] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Advances in psychology. Vol. 52. Elsevier, 139–183.

[38] Shirin E Hassan. 2012. Are normally sighted, visually impaired, and blind pedestrians accurate and reliable at making street crossing decisions? Investigative ophthalmology & visual science 53, 6 (2012), 2593–2600.

[39] Abdelsalam Helal, Steven Edwin Moore, and Balaji Ramachandran. 2001. Drishti: An integrated navigation system for visually impaired and disabled. In Proceedings fifth international symposium on wearable computers. IEEE, 149–156.

[40] Christopher R. Hudson, Shuchisnigdha Deb, Daniel W. Carruth, John McGinley, and Darren Frey. 2018. Pedestrian Perception of Autonomous Vehicles with External Interacting Features. In International Conference on Applied Human Factors and Ergonomics. Springer, 33–39.

[41] Ahmed Hussein, Fernando Garcia, Jose Maria Armingol, and Cristina Olaverri-Monreal. 2016. P2V and V2P Communication for Pedestrian Warning on the basis of Autonomous Vehicles. In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2034–2039.

[42] ISO/TR 23049:2018 2018. Road Vehicles: Ergonomic Aspects of External Visual Communication from Automated Vehicles to Other Road Users. Standard. International Organization for Standardization.

[43] Volodymyr Ivanchenko, James Coughlan, and Huiying Shen. 2008. Crosswatch: a camera phone system for orienting visually impaired pedestrians at traffic intersections. In International Conference on Computers for Handicapped Persons. Springer, 1122–1128.

[44] Moritz Körber. 2018. Theoretical considerations and development of a questionnaire to measure trust in automation. In Congress of the International Ergonomics Association. Springer, 13–30.

[45] Athanasios Koutsoklenis and Konstantinos Papadopoulos. 2011. Auditory cues used for wayfinding in urban environments by individuals with visual impairments. Journal of Visual Impairment & Blindness 105, 10 (2011), 703–714.

[46] Tobias Lagstrom and Victor Malmsten Lundgren. 2015. AVIP-Autonomous vehicles interaction with pedestrians. Master of Science Thesis, Chalmers University of Technology (2015).

[47] Yeti Li, Murat Dikmen, Thana G. Hussein, Yahui Wang, and Catherine Burns. 2018. To cross or not to cross: Urgency-based external warning displays on autonomous vehicles to improve pedestrian crossing safety. In Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, 188–197.

[48] Andreas Löcken, Carmen Golling, and Andreas Riener. 2019. How Should Automated Vehicles Interact with Pedestrians?: A Comparative Analysis of Interaction Concepts in Virtual Reality. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, 262–274.

[49] Victor Malmsten Lundgren, Azra Habibovic, Jonas Andersson, Tobias Lagström, Maria Nilsson, Anna Sirkka, Johan Fagerlönn, Rikard Fredriksson, Claes Edgren, and Stas Krupenia. 2017. Will there be new communication needs when introducing automated vehicles to the urban context? Springer, 485–497.

[50] Karthik Mahadevan, Elaheh Sanoubari, Sowmya Somanath, James E Young, and Ehud Sharlin. 2019. AV-Pedestrian Interaction Design Using a Pedestrian Mixed Traffic Simulator. In Proceedings of the 2019 on Designing Interactive Systems Conference. ACM, 475–486.

[51] Karthik Mahadevan, Sowmya Somanath, and Ehud Sharlin. 2018. Communicating awareness and intent in autonomous vehicle-pedestrian interaction. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 429.

[52] Milecia Matthews, Girish Chowdhary, and Emily Kieson. 2017. Intent communication between autonomous vehicles and pedestrians. arXiv preprint arXiv:1708.07123 (2017).

[53] Natasha Merat, Tyron Louw, Ruth Madigan, Marc Wilbrink, and Anna Schieben. 2018. What externally presented information do VRUs require when interacting with fully Automated Road Transport Systems in shared space? Accident Analysis & Prevention 118 (2018), 244–252.

[54] Nicole Mirnig, Nicole Perterer, Gerald Stollnberger, and Manfred Tscheligi. 2017. Three strategies for autonomous car-to-pedestrian communication: A survival guide. In Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. ACM, 209–210.

[55] Ingunn Moser. 2005. On becoming disabled and articulating alternatives: The multiple modes of ordering disability and their interferences. Cultural Studies 19, 6 (2005), 667–700.

[56] Taylor A Nelson, Andrew A Abeyta, and Clay Routledge. 2019. What makes life meaningful for theists and atheists? Psychology of Religion and Spirituality (2019).

[57] Alan F Newell and Peter Gregor. 2000. "User sensitive inclusive design"—in search of a new paradigm. In Proceedings on the 2000 conference on Universal Usability. ACM, 39–44.

[58] Kimihiro Noguchi, Yulia R Gel, Edgar Brunner, and Frank Konietschke. 2012. nparLD: an R software package for the nonparametric analysis of longitudinal data in factorial experiments. Journal of Statistical Software 50, 12 (2012).

[59] World Health Organization. 2018. Blindness and vision impairment. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment. (2018). [Online; accessed: 07-JUNE-2019].

[60] World Health Organization. 2019. Deafness and hearing loss. https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss. (2019). [Online; accessed: 07-SEPTEMBER-2019].

[61] Amir Rasouli, Iuliia Kotseruba, and John K Tsotsos. 2017. Understanding pedestrian behavior in complex traffic scenes. IEEE Transactions on Intelligent Vehicles 3, 1 (2017), 61–70.

[62] Amir Rasouli and John K Tsotsos. 2019. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. IEEE Transactions on Intelligent Transportation Systems (2019).

[63] Ulrike Rau. 2019. Zwei-Sinne-Prinzip. https://nullbarriere.de/rau-barrierefrei-bauen-zwei-sinne-prinzip.htm. (2019). [Online; Accessed: 18-JULY-2019].

[64] Michael Rettenmaier, Moritz Pietsch, Jonas Schmidtler, and Klaus Bengler. 2019. Passing through the Bottleneck-The Potential of External Human-Machine Interfaces. In 2019 IEEE Intelligent Vehicles Symposium (IV). IEEE, 1687–1692.

[65] Ryan L Robart and Lawrence D Rosenblum. 2009. Are hybrid cars too quiet? The Journal of the Acoustical Society of America 125, 4 (2009), 2744–2744.

[66] Manaswi Saha, Alexander J Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. 2019. Closing the Gap: Designing for the Last-Few-Meters Wayfinding Problem for People with Visual Impairments. (2019).

[67] Avilash Sahoo, Santosha K. Dwivedy, and P.S. Robi. 2019. Advancements in the field of autonomous underwater vehicle. Ocean Engineering 181 (2019), 145 – 160. DOI:http://dx.doi.org/https://doi.org/10.1016/j.oceaneng.2019.04.011

[68] Johnny Saldaña. 2015. The coding manual for qualitative researchers. Sage.

[69] Sarah Schmidt and B Färber. 2009. Pedestrians at the kerb–Recognising the action intentions of humans. Transportation research part F: traffic psychology and behaviour 12, 4 (2009), 300–310.

[70] JJ Scott and Robert Gray. 2008. A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. Human factors 50, 2 (2008), 264–275.

[71] Kristen Shinohara and Jacob O Wobbrock. 2011. In the shadow of misperception: assistive technology use and social interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 705–714.

[72] Andreas Sieß, Kathleen Hübel, Daniel Hepperle, Andreas Dronov, Christian Hufnagel, Julia Aktun, and Matthias Wölfel. 2015. Hybrid City Lighting-Improving Pedestrians' Safety through Proactive Street Lighting. In 2015 International Conference on Cyberworlds (CW). IEEE, 46–49.

[73] Ye Eun Song, Christian Lehsing, Tanja Fuest, and Klaus Bengler. 2018. External HMIs and their effect on the interaction between pedestrians and automated vehicles. In International Conference on Intelligent Human Systems Integration. Springer, 13–18.

[74] Laerd statistics. 2016. Mixed ANOVA using SPSS Statistics. https://statistics.laerd.com/spss-tutorials/mixed-anova-using-spss-statistics.php. (2016). [Online; Accessed: 12-SEPTEMBER-2019].

[75] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. 2009. Estimation of detection thresholds for redirected walking techniques. IEEE transactions on visualization and computer graphics 16, 1 (2009), 17–27.

[76] Klaas-Jan Stol, Paul Ralph, and Brian Fitzgerald. 2016. Grounded theory in software engineering research: a critical review and guidelines. In 2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE). IEEE, 120–131.

[77] Helena Strömberg, Ingrid Pettersson, Jonas Andersson, Annie Rydström, Debargha Dey, Maria Klingegård, and Jodi Forlizzi. 2018. Designing for social experiences with and within autonomous vehicles–exploring methodological directions. Design Science 4 (2018).

[78] Alireza Talebpour and Hani S Mahmassani. 2016. Influence of connected and autonomous vehicles on traffic flow stability and throughput. Transportation Research Part C: Emerging Technologies 71 (2016), 143–163.

[79] Arenda F te Velde, John van der Kamp, José A Barela, and Geert JP Savelsbergh. 2005. Visual timing and adaptive behavior in a road-crossing simulation study. Accident Analysis & Prevention 37, 3 (2005), 399–406.

[80] Anja Thieme, Cynthia L Bennett, Cecily Morrison, Edward Cutrell, and Alex S Taylor. 2018. I can do everything but see!–How People with Vision Impairments Negotiate their Abilities in Social Contexts. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 203.

[81] Jean Toner. 2009. Small is not too small: Reflections concerning the validity of very small focus groups (VSFGs). Qualitative Social Work 8, 2 (2009), 179–192.

[82] Unity Technologies. 2019. Unity. (2019). https://unity.com/

[83] Himanshu Verma, Guillaume Pythoud, Grace Eden, Denis Lalanne, and Florian Evéquoz. 2019. Pedestrians and Visual Signs of Intent: Towards Expressive Autonomous Passenger Shuttles. 3 (09 2019), 107:1. DOI:http://dx.doi.org/10.1145/3351265

[84] Annette Werner. 2019. New colours for autonomous driving: An evaluation of chromaticities for the external lighting equipment of autonomous vehicles. Colour Turn 1 (2019).

[85] WR Wiener and G Lawson. 1997. The use of traffic sounds to make street crossings by persons who are visually impaired. Journal of Visual Impairment & Blindness 91, 5 (1997).

[86] Michele A Williams, Caroline Galbraith, Shaun K Kane, and Amy Hurst. 2014. Just let the cane hit it: how the blind and sighted see navigation differently. In Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility. ACM, 217–224.

[87] Michele A Williams, Amy Hurst, and Shaun K Kane. 2013. Pray before you step out: describing personal and situational blind navigation behaviors. In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, 28.

[88] Scooter Willis and Sumi Helal. 2005. RFID information grid for blind navigation and wayfinding. In Ninth IEEE International Symposium on Wearable Computers (ISWC'05). IEEE, 34–37.

[89] Jacob O Wobbrock, Shaun K Kane, Krzysztof Z Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-based design: Concept, principles and examples. ACM Transactions on Accessible Computing (TACCESS) 3, 3 (2011), 9.

[90] Su Yang. 2017. Driver behavior impact on pedestrians' crossing experience in the conditionally autonomous driving context. (2017).

[91] Jingyi Zhang, Erik Vinkhuyzen, and Melissa Cefkin. 2017. Evaluation of an autonomous vehicle external communication system concept: a survey study. In International Conference on Applied Human Factors and Ergonomics. Springer, 650–661.

[92] Yuhang Zhao, Cynthia L Bennett, Hrvoje Benko, Edward Cutrell, Christian Holz, Meredith Ringel Morris, and Mike Sinclair. 2018. Enabling People with Visual Impairments to Navigate Virtual Reality with a Haptic and Auditory Cane Simulation. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 116.

[93] Raphael Zimmermann and Reto Wettach. 2017. First step into visceral interaction with autonomous vehicles. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, 58–64.