

# Towards Inclusive External Human-Machine Interface: Exploring the Effects of Visual and Auditory eHMI for Deaf and Hard-of-Hearing People

Wenge Xu  
Birmingham City University  
Birmingham, United Kingdom  
wenge.xu@bcu.ac.uk

Foroogh Hajiseyedjavadi  
Birmingham City University  
Birmingham, United Kingdom  
Foroogh.Hajiseyedjavadi@bcu.ac.uk

Kurtis Weir  
Birmingham City University  
Birmingham, United Kingdom  
Kurtis.Weir@bcu.ac.uk

Chukwuemeka Eze  
Birmingham City University  
Birmingham, United Kingdom  
chukwuemeka.eze@mail.bcu.ac.uk

Mark Colley  
UCL Interaction Centre  
London, United Kingdom  
m.colley@ucl.ac.uk

## Abstract

External Human-Machine Interfaces (eHMIs) have been proposed to facilitate communication between Automated Vehicles (AVs) and pedestrians. However, no attention was given to Deaf and Hard-of-Hearing (DHH) people. We conducted a formative study through focus groups with 6 DHH people and 6 key stakeholders (including researchers, assistive technologists, and automotive interface designers) to compare proposed eHMIs and extract key design requirements. Subsequently, we investigated the effects of visual and auditory eHMI in a virtual reality user study with 32 participants (16 DHH). Results from our scenario suggesting that (1) DHH participants spent more time looking at the AV; (2) both visual and auditory eHMIs enhanced trust, usefulness, and perceived safety; and (3) only visual eHMIs reduced the time to step into the road, time looking at the AV, gaze time, and percentage looking at active visual eHMI components. Lastly, we provided five practical implications for making eHMI inclusive of DHH people.

## CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; **Accessibility design and evaluation methods**.

## Keywords

Deaf, Hard-of-Hearing, External Human-Machine Interface, Automated Vehicles, Accessibility.

## ACM Reference Format:

Wenge Xu, Foroogh Hajiseyedjavadi, Kurtis Weir, Chukwuemeka Eze, and Mark Colley. 2026. Towards Inclusive External Human-Machine Interface: Exploring the Effects of Visual and Auditory eHMI for Deaf and Hard-of-Hearing People. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*, April 13–17, 2026, Barcelona, Spain. ACM, New York, NY, USA, 21 pages. <https://doi.org/10.1145/3772318.3790738>



This work is licensed under a Creative Commons Attribution 4.0 International License. *CHI '26, Barcelona, Spain*

© 2026 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-2278-3/26/04  
<https://doi.org/10.1145/3772318.3790738>

## 1 Introduction

External human-machine interfaces (eHMI) have been proposed for several years to address the absence of traditional explicit communication cues between pedestrians and human drivers [89] and assist in communication between pedestrians and automated vehicles (AVs) [74]. However, a significant gap in the eHMI literature is that most papers only focus on non-disabled people, excluding 1.3 billion (16%) of the world population who live with significant disability<sup>1</sup>. The disability of pedestrians is a critical contributory factor to fatal or serious collisions with vehicles<sup>2</sup>. The design of eHMIs needs to consider the needs of disabled people; otherwise, it may be inaccessible to them, creating a further divide and inequality in transport. There are initial explorations towards making eHMIs accessible for low vision or blind people [21], intellectually disabled people [51], and wheelchair users [6]. However, there is a lack of investigation on deaf and hard-of-hearing (DHH) people. According to the World Health Organisation, the population of DHH people reached 430 million worldwide and is estimated to reach 700 million in 2050 (i.e., 1 in every 10 people)<sup>3</sup>.

Most existing work on eHMIs focused on visual communication design patterns (e.g., abstract light [30], anthropomorphic features [15], text [62], symbols [16], or projection [13]). Providing only visual eHMIs would not benefit low vision or blind people or those who experience situational impairments (e.g., being distracted by secondary activities, occluded view) [23]. Several works have suggested that using multi-modal eHMI, e.g., combined visual-auditory communication, could make AVs accessible and benefit more pedestrians [21, 51]. However, what visual designs would work for DHH people remains unclear due to the lack of involvement of DHH in the design and evaluation phase. Additionally, how DHH people would perceive the use of auditory communication, such as speech, remains unclear.

We first conducted a focus group study with DHH people and key stakeholders (i.e., human factor researchers, eHMI researchers, accessibility researchers, assistive technologist, and HMI designer) to help us (1) minimise the visual design candidates to be used in

<sup>1</sup>WHO: Disability and Health; accessed 14.04.2025

<sup>2</sup>Reported road casualties in Great Britain: pedestrian factsheet, 2022; accessed 14.08.2024

<sup>3</sup>WHO: Deafness and Hearing Loss; accessed 14.04.2025

our VR study, and (2) come up with initial design requirements. We then conducted a mixed-design VR study with N=32 participants to investigate the effect of Visual (*No Visual*, *Abstract Light*, *Abstract Light + Text*, *Abstract Light + Symbol*) and Auditory (*Without Speech* and *With Speech*) eHMIs among Hearing group (N=16) and DHH group (N=16) with regards of their crossing experience (trust, acceptance, perceived safety, mental load) and behaviour (gaze, step into the road time, early step into the road count). Importantly, we further investigated the extent to which the DHH group valued these differently from the Hearing group.

**Contribution Statement:** Our main contributions: (1) A first focus group study with DHH people (N=6) and relevant stakeholders (N=6) comparing six visual external communication designs from the literature and exploring initial design requirements. (2) A first VR user study that involves DHH participants (16 out of 32 participants) showing that (i) there are behaviour (i.e., eye gaze) differences between participants in DHH group and Hearing group, (ii) providing visual communication improves crossing experience (i.e., trust, usefulness, perceived safety) and behaviour (step-in road time, time spent on the AV, time and percentage of gaze spent on active visual eHMI components), and (iii) providing speech as auditory communication improves crossing experience (i.e., trust, usefulness, perceived safety). (3) 5 practical implications that pave the way for future eHMI design and research.

## 2 Related Work

### 2.1 DHH People and Road Crossing

Hearing and eyesight are the most relevant senses to act adequately in traffic situations [55]. Challenges such as obstructed views due to parked vehicles, low lighting, adverse weather conditions, or dazzling produced by adjacent lights could affect the crossing visually [4, 48]. Hearing helps us understand the direction and distance of sounds, allowing us to judge the location of potential threats or obstacles [63], which could be particularly challenging for DHH people [99]. Previous study with DHH teenagers showed that almost half of DHH teenagers have been involved in traffic accidents, which is a much higher rate than hearing teenagers Gür et al. [50].

Hearing technologies such as hearing aids and hearing implants are the most commonly used technologies by DHH people [95]. Hearing aids are intended to help people with mild to moderate hearing loss<sup>4</sup>. Hearing aids may improve sound localisation, but have failed to improve it consistently and could even impair it [99]. As the perceived benefits not meeting the expectation and the wearing discomfort, only about 20% people who would benefit from hearing aids use one [76]. Hearing implants such as cochlear implants help provide a sense of sound to people with severe to profound deafness<sup>5</sup>. This technology does not restore normal hearing; instead, they restore some sensation of auditory perception to help deaf people understand sounds or speech [83]. However, data from the UK showed that only 1.3% of severely or profoundly deaf people use hearing implants<sup>6</sup>. In short, neither technology is perfect, nor would it help DHH people address the absence of

hearing cues, which may lead to missing critical information when interacting with AVs through eHMIs, especially if they employ auditory feedback.

Limited work has been conducted to understand road crossing among DHH pedestrians. Pecchini and Giuliani [82] conducted a field study in an urban environment with manually driven vehicles and found that DHH people often experienced heightened apprehension when initiating a crossing and exercised greater caution toward approaching vehicles to ensure safety. Lundälv [71] suggest that pedestrians with moderate deafness are at a higher risk of being injured by a vehicle because they have difficulty in identifying the direction the sound is coming from. Overall, the lack of access to auditory information has reduced feelings of safety among DHH people [54]. Over time, this could constitute a negative and fatiguing experience, which may discourage active travel like walking [81], resulting in reduced physical activity levels [14] and broadening inequities in mobility [56].

Regarding crossing strategies, Pecchini and Giuliani [82] observed that DHH pedestrians employed deliberate tactics to communicate with drivers to secure their right-of-way. These included actively seeking direct visual contact with approaching drivers and using hand gestures to capture attention and signal them to slow down or stop. Such cooperative, bidirectional interaction allowed DHH pedestrians to create safer crossing conditions. However, this strategy poses a critical challenge in AVs due to the absence of human drivers. This highlights the need for investigation visual and auditory eHMIs with DHH people.

### 2.2 External Human-Machine Interface

As eyesight (visual) and hearing (auditory) are the most relevant senses (modalities) to act adequately [55], we first discuss visual eHMIs and auditory eHMIs, then move on to discussing eHMIs and disabilities, and finish with multi-modal eHMIs as a potential solution for disabled people when interacting with AVs.

**2.2.1 Visual eHMIs.** Dey et al. [29] found that the visual modality is the most commonly adopted modality among the eHMI literature, and multiple studies [26, 70] have shown that the presence of a visual eHMI could assist in the communication between vulnerable road users and AVs. These proposed visual eHMI concepts can be categorized into the following six visual design patterns, including (1) Road-based Projection [13], (2) Symbols [16], (3) Text [62], (4) Anthropomorphic, including Smiley<sup>7</sup>, eyes [15], and gestures [74], (5) Abstract Light: one-dimensional light segment [30] or two-dimensional light display, and (6) Situational Awareness: a light tracker to show the situational awareness of the AV [29].

Dey et al. [33] compared Situational Awareness, Progress Bar (Abstract Light), Moving Light Bar (Abstract Light), and Road-based Projection and found that Road-based Projection is preferred when the AV has to interact with multiple pedestrians. Text was found preferred by participants in a video-based study Ackermann et al. [2] and a survey study [62]. Despite much research into comparing visual concepts, there is still no explicit agreement on which visual design is the best and how many visual signals are sufficient [31]. Due to the lack of agreement towards the choice of visual design

<sup>4</sup><https://www.nidcd.nih.gov/health/hearing-aids>; accessed 02.08.2024

<sup>5</sup><https://www.nidcd.nih.gov/health/cochlear-implants>

<sup>6</sup>Hearing loss statistics in the UK: <https://www.hearinglink.org/your-hearing/about-hearing/facts-about-deafness-hearing-loss/>

<sup>7</sup>Semcon: The Smiling car; accessed 14.07.2024

and the gap in the research with DHH people, we first conducted a formative study to understand DHH people and the needs of relevant stakeholders and to narrow down visual eHMI designs that we will be using in the VR study. Nevertheless, we locate the visual eHMI on the front of the vehicle grill, which mimics the current expectations and living experience of pedestrians, who generally look towards the location of the driver's head or vehicle movement [34, 40], but also follows the prior works [18, 31].

**2.2.2 Auditory eHMI.** There is relatively less attention to the auditory eHMI design for AVs. Deb et al. [28] suggested that visual eHMIs had a much larger effect on the willingness to initiate crossing when compared to auditory eHMIs (horn, music, and verbal warning saying "safe to cross"). However, auditory eHMI holds several benefits, such as announcing situation awareness and detection [78] and is a reliable source for vision disabled people [21]. The implementations of auditory eHMI are either through speech (verbal messages) Deb et al. [28], Hudson et al. [60], Mahadevan et al. [73] or non-speech auditory signals (e.g., jingles, humming, bells) [10, 31, 46].

Dey et al. [31] compared two non-speech concepts (i.e., bell and droning) in a video-based study, and found that participants had completely different associations and mental models for using these two sounds. Many people perceived the bell as a calm, inviting, and friendly signal, but others perceived it as urgent and rushed and associated it with a warning. Similarly, while several people felt the choice of drone was an intuitive and natural concept for the vehicle's engine sound and speed, others found it unpleasant and burdensome. Prior works have suggested that verbal messages might be more reliable than non-speech auditory signals. Deb et al. [28] investigated multiple sound-related features in a VR study, finding that the verbal message was the most favoured audible feature compared to horn, music, and no sound. Hudson et al. [60] found similar results where verbal messages are preferred over music. In addition, verbal messages may be more inclusive for vision disabled people [21]. Therefore, we used speech in our VR study to explore DHH people's opinions regarding using auditory eHMI.

**2.2.3 eHMI and Disabilities.** Despite extensive research having been conducted on the development of eHMIs, only limited research has been conducted with disabled people. Some works focus on co-designing with disabled people from scratch, whereas others focus on evaluating existing concepts and tailoring them for disabled people. Asha et al. [6] focused on the design of eHMI tailored for wheelchair users; they employed inclusive design practices to work with a wheelchair user to explore the user requirements of wheelchair users when communicating with AVs as pedestrians. On the other hand, Colley et al. [21] evaluated existing auditory concepts with the low vision and blind people through a workshop study and found that the speech auditory eHMI was best received. Their follow-up VR study suggests that having more content in the speech message could reduce the mental load. Haimerl et al. [51] compared eHMI concepts (baseline, visual-only, auditory-only, multi-modal—both visual and audio) through an online video-based survey study with intellectual disabled participants and non-intellectual disabled participants; they found that auditory eHMIs performed worse than visual or multi-modal eHMIs. In general,

Haimerl et al. [51] found that multi-modal eHMIs positively affect quality and inclusion.

The need for multi-modal eHMIs has been a common agreement among the studies worked with disabled people [6, 21, 22, 51]. In addition, it has been suggested by a review paper on eHMI accessibility [42], a core reason is that each modality has specific trade-off Mahadevan et al. [74]. Multi-modal interfaces or feedback designs have been commonly used by other fields to make the interaction accessible for disabled people [5, 24, 92].

**2.2.4 Multi-Modal eHMIs.** The literature suggests that multi-modal eHMI brings several benefits to the pedestrian. Dou et al. [36] conducted a VR study to evaluate 12 eHMI concepts, in combinations of visual (smile/arrow), audio (human voice/warning sound), and vehicle movement style (the approaching speed decreases gradually/remains unchanged), and concluded that multi-modal eHMIs resulted in more satisfactory interaction and improved safety compared to the unimodal eHMI. He et al. [53] conducted a VR study and found that audio-visual modality (symbol and anthropomorphic voice) was more appealing than the eHMI with a unimodal modality (visual or auditory). Results from the Wizard-of-Oz study by [3] suggested that a combination of audio-visual modality is most effective in understanding information.

Despite these potential benefits discussed in the literature, a recent video-based study by Dey et al. [31] found that multi-modal eHMI do not outperform unimodal eHMI regarding effectiveness and user experience. Their qualitative data suggested that multi-modal eHMI may also overwhelm pedestrians, although it also offers assurance of vehicle intention. To our knowledge, no studies have been conducted on exploring eHMIs with DHH people. Therefore, we conducted a VR study to explore the visual eHMI designs, whether auditory eHMIs would be well perceived by DHH people, and how DHH people may perceive eHMIs differently from hearing people.

### 3 Formative Study

As there is a lack of work in eHMI with Deaf and Hard-of-Hearing (HoH) people, we first conducted a formative study with two focus groups (a Deaf-focused group for profoundly deaf people and a HoH-focused group for mild to severe hearing loss people) to gain insights from DHH people and key stakeholders (1) regarding high-level visual eHMI design requirements and (2) about visual eHMI candidates for our VR-based user study.

Focus groups are well-suited to the exploratory nature of formative study and are commonly used in early-stage HCI studies [7, 77, 93]. We intentionally brought together DHH people and key stakeholders to ensure a wide range of perspectives and to facilitate more inclusive discussions about current accessibility challenges in pedestrian–AV interaction [11, 25, 37]. The group format enabled us to capture a breadth of viewpoints and to observe how DHH people and stakeholders collectively interpret, negotiate, and debate these concepts.

#### 3.1 Participants

Twelve participants (4F, 8M, mean age=35.17, SD=9.20) participated in the study (see Table 1). The DHH people were recruited through posters, social media platforms, and word-of-mouth. Stakeholders

like eHMI researchers, accessibility researchers, HMI industry specialists, and people who work in the charity sector were recruited through research papers, charity networks, social media platforms, and prior project databases. The group allocation was done based on participants' lived experience, (2) research interest, or (3) industry expertise. The hearing loss levels were categorised based on recommendations from National Health Services<sup>8</sup>: normal (<20 dB), mild (21 - 40 dB), moderate (41 - 70 dB), severe (71 - 95 dB), and profound (>95 dB).

**Table 1: Pseudonyms and background of focus group participants**

Pseudonym	Age	Relatedness to the project
D1	30s	Profoundly deaf in both ears and uses BSL
D2	mid-20s	Profoundly deaf in both ears and uses BSL
D3	50s	Human factors and eHMI researcher who has worked with disabled people
D4	30s	Human factors and eHMI researcher who works at a national centre for accessible transport
D5	40s	HMI designer who works at JLR and has a special interest in eHMI and AV
D6	30s	Accessibility researcher who has single-ear profound deafness and has research interests in DHH people
H7	20s	Severe deafness in both ears
H8	late-30s	Mild deafness in a single ear
H9	30s	Severe deafness in one ear and moderate deafness in the other, who is also an assistive technologist
H10	late-40s	Human factor and eHMI researcher who has worked with JLR
H11	mid-30s	Accessibility researcher with a special interest in HoH research
H12	mid-30s	Accessibility researcher with a special interest in HoH research who works for a national charity for DHH people

Pseudonyms consist of focus group initial (D: deaf, H: Hard-of-Hearing) and Participant ID (1-12).

### 3.2 Apparatus

Similar to prior focus group studies [25, 59, 79], we used Teams as the online video conferencing tool. We ensured the accessibility of Teams by asking participants about any accessibility adjustments prior to the study. Transcripts were enabled by default for all users. Two British Sign Language (BSL) interpreters were recruited for BSL users in the Deaf group; interpreters were pinned to the screen for BSL users. Overall, participants found the choice of Teams acceptable.

<sup>8</sup>National Health Service: <https://www.esht.nhs.uk/service/audiology/diagnosis-and-testing/>

### 3.3 Evaluated Concepts

We evaluated the following six commonly used eHMI visual concepts [31]; each concept comes with a video demonstration.

- (1) *Abstract Light*: This concept employs a light bar on the bumper to communicate the AV's state and intention [30]. The video used in the study had the following implementation: While the AV is driving, the light bar is glowing statically in cyan. When yielding for a pedestrian, the bar starts to pulsate between on and off at 1 Hz. Once the AV comes to a complete stop, the light bar turns off.
- (2) *Situational Awareness*: This concept follows the design pattern from [29], where a square cyan indicator at the bottom of the windscreen of the AV would appear when it detects the pedestrian and would point toward the pedestrian to show the AV intends to yield for [18]. If the pedestrian steps onto the road and starts crossing, the cyan indicator will follow the pedestrian.
- (3) *Text*: The *Text* concept displays text to the pedestrian on a bumper display [18]. While driving, the bumper display will show "DRIVING". When an AV begins to yield, the text will switch to "YIELDING". When stopped, it displays "STOPPED".
- (4) *Symbol*: This concept follows the symbol design pattern [16, 29] displayed on a bumper display. The display does not show any additional information and would activate a pedestrian symbol to indicate the AV would yield for the pedestrian [18].
- (5) *Road-based Projection*: This concept was based on the prior works in projection design pattern [35, 69]. Once the AV begins to yield, it projects a rectangle with triangles representing arrows up to 10 m in front of the car. The projection moves with the AV until it is 10 m away from its yielding position. At this point, the end of the rectangle represents the line that marks the spot where the AV will stop. The AV will continue to come closer to the yielding line. Once stopped, the stopping line will remain present [18].
- (6) *Anthropomorphic*: We simplified the design patterns introduced by Dey et al. [29] by combining "Eyes", "Smiling", and "Other Anthropomorphic" into *Anthropomorphic*. For the video presentation of this eHMI design, we implement the same Smiling AV concept as described in [18], where the anthropomorphic interface is displayed through the AV's bumper display with a "mouth" on it. If the AV is not yielding, the mouth remains its natural expression (i.e., a horizontal line). If the AV is yielding to a pedestrian, the mouth turns the horizontal line into a smile.

### 3.4 Procedure

Both focus group sessions were conducted on the same day, with each participant attending only their assigned group. The study began with an introduction from the research team, outlining the research brief and key activities. Upon transitioning to the eHMI section, each group's facilitator (part of the author team) provided a high-level overview of eHMI (e.g., what eHMI is and why it is needed) followed by a question and answer session for participants. Then, the group discussion began with six eHMI concepts discussed



in order. For each concept, the facilitator first gave a brief introduction accompanied by images and a short video presentation. Participants were then asked to take turns sharing their views on the benefits, disadvantages, concerns, and potential barriers of the concept, with the reminder that there were no right or wrong answers. Once all participants shared the corresponding eHMI idea, they can freely discuss the concept. Before moving on to the next concept, the facilitator asked if participants had any additional comments.

This activity took around 150 minutes, including breaks. After participants went through all concepts, they filled in a questionnaire (see Section 3.5). After completing the questionnaire, they had another 30 minutes for open discussion (if they had other design ideas, auditory information).

### 3.5 Quantitative Measurements

We employed the following questions to help us decide candidates for the subsequent VR user study.

- (1) Trust: Trust was measured through the Trust subscale and the Understandability subscale from the Trust in Automation questionnaire [64] in a 5-point Likert scale (1=Strongly disagree to 5=Strongly agree).
- (2) Acceptance: We employed the van der Laan acceptance scale with the subscales "usefulness" and "satisfying" [91]. The scale is measured in a 5-point Likert scale during collection and adjusted to -2 (negative rating) to +2 (positive rating) in the data analysis.
- (3) Perceived Safety: We assessed the Perceived Safety using questions from [44] with a 7-point Likert scale, from -3 (negative) to +3 (positive).
- (4) Mental Workload: We used the raw Mental Workload subscale on a 21-point Likert scale from the NASA-TLX Questionnaire [52].
- (5) Necessity and Reasonability [18]: We employed a 7-point Likert scale (1=Totally Disagree to 7=Totally Agree) to measure participants' perception of the Necessity and Reasonability of overall visual eHMI concept usage.
- (6) Rank: Participants ranked the eHMI concepts regarding their preference (the lower the number, the better).

### 3.6 Findings

This formative study helped us identify (1) appropriate visual eHMI candidates for the VR user study through a combination of questionnaire data and ranking, and (2) top-level design requirements through qualitative data.

**3.6.1 Quantitative Survey Results and Discussion.** We used Shapiro-Wilk to check the normality of the data, for normally distributed data, we used one-way repeated measure ANOVA with eHMI Types as the within-subject factor to help us gain initial understanding of the DHH people and relevant stakeholders' perspectives regarding each eHMI design. We used the Friedman test to analyze data if they were not normally distributed. Bonferroni corrections were always used in all post-hoc analyses.

**Trust in Automation.** Regarding Understandability, an ANOVA revealed a significant effect of eHMI Types ( $F(5, 55) = 7.864, p <$

$.001, \eta_p^2 = .417$ ). Post-hoc pairwise comparisons indicated that *Abstract Light* ( $M = 3.10, SD = 0.77$ ) was rated higher than *Anthropomorphic* ( $M = 1.96, SD = 0.96, p = .029$ ) and *Situational Awareness* ( $M = 1.90, SD = 0.76, p = .013$ ). *Text* ( $M = 3.40, SD = 0.90$ ) was rated higher than *Anthropomorphic* ( $M = 1.96, SD = 0.96, p = .034$ ) and *Situational Awareness* ( $M = 1.90, SD = 0.76, p = .010$ ).

As for Trust, ANOVA yielded a significant effect of eHMI Types ( $F(5, 55) = 8.907, p < .001, \eta_p^2 = .447$ ). Post-hoc tests suggested that participants rated a significant lower rating for *Situational Awareness* ( $M = 1.96, SD = 0.72$ ) than *Abstract Light* ( $M = 3.29, SD = 0.81, p = .004$ ), *Symbol* ( $M = 3.30, SD = 0.92, p = .031$ ) and *Text* ( $M = 3.50, SD = 0.74, p = .003$ ). *Anthropomorphic* was also rated lower than *Abstract Light* ( $M = 3.29, SD = 0.81, p = .028$ ) and *Text* ( $M = 3.50, SD = 0.74, p = .032$ ).

**Perceived Safety.** ANOVA revealed a significant effect of eHMI Types  $F(5, 55) = 7.305, p < .001, \eta_p^2 = .399$ . Post-hoc tests showed that *Situational Awareness* ( $M = -1.77, SD = 0.97$ ) was rated significantly lower than *Symbol* ( $M = 0.69, SD = 1.43, p = .007$ ), *Abstract Light* ( $M = 0.42, SD = 1.47, p = .014$ ), *Text* ( $M = 0.71, SD = 1.45, p = .002$ ), *Road-based Projection* ( $M = 0.02, SD = 1.27, p = .022$ ).

**Acceptance.** As for Usefulness, ANOVA showed a significant effect of eHMI Types  $F(5, 55) = 7.453, p < .001, \eta_p^2 = .404$ . Post-hoc tests suggested that *Situational Awareness* ( $M = -0.85, SD = 0.73$ ) as rated significantly lower than *Symbol* ( $M = 0.63, SD = 0.90, p = .020$ ), *Abstract Light* ( $M = 0.50, SD = 0.69, p = .003$ ), and *Text* ( $M = 0.80, SD = 0.72, p = .001$ ).

Regarding Satisfying, ANOVA yielded a significant effect of eHMI Types  $F(5, 55) = 8.728, p < .001, \eta_p^2 = .442$ . Post-hoc tests suggested that *Situational Awareness* ( $M = -1.02, SD = 0.75$ ) was rated significantly lower than *Symbol* ( $M = 0.67, SD = 0.83, p = .006$ ), *Abstract Light* ( $M = 0.33, SD = 0.69, p = .002$ ), and *Text* ( $M = 0.50, SD = 0.75, p = .011$ ).

**Mental Workload.** ANOVA yielded a significant main effect of eHMI Types ( $F(5, 55) = 7.105, p < .001, \eta_p^2 = .392$ ) on Mental Workload. Post-hoc pairwise comparisons tests suggested that *Situational Awareness* ( $M = 15.92, SD = 3.80$ ) led to greater mental workload than *Symbol* ( $M = 7.25, SD = 5.36, p = .002$ ), *Abstract Light* ( $M = 7.00, SD = 4.00, p < .001$ ), and *Text* ( $M = 9.17, SD = 4.93, p < .001$ ). No other significant results were found.

**Necessity, Reasonability, and Preference Ranking.** Participants stated eHMI designs are necessary ( $M = 5.58, SD = 1.24$ ) and reasonable ( $M = 5.50, SD = 0.90$ ). For ranking data, *Abstract Light* ( $M = 2.25, SD = 1.22$ ) was rated the best, followed by *Text* ( $M = 2.50, SD = 1.00$ ) and *Symbol* ( $M = 2.58, SD = 1.51$ ). *Situational Awareness* is the least preferred ( $M = 5.50, SD = 0.52$ ), with *Anthropomorphic* rated the second worst ( $M = 4.75, SD = 1.29$ ). *Road-based Projection* ( $M = 3.42, SD = 1.62$ ) had the third-worst rating. Friedman's Test showed that there was a significant difference in the ranking data among the eHMI concepts ( $\chi^2(5) = 30.762, p < .001$ ). Post hoc analysis with Bonferroni correction showed that *Situational Awareness* was rated significantly worse than *Symbol* ( $p = .045$ ), *Abstract Light* ( $p = .03$ ), and *Text* ( $p = .03$ ). In addition, *Text* is rated significantly better than *Anthropomorphic* ( $p = .03$ ).

In summary, our quantitative data (questionnaire ratings and ranking) suggested *Abstract Light* was the best perceived visual

design, followed by *Text* and *Symbol*. Overall, the results appeared to align with results reported in prior works by [18, 33], except for the *Road-Projection*.

**3.6.2 High-Level eHMI Design Requirements.** For the qualitative data analysis, we conducted a thematic analysis [12] using an inductive coding approach. The interview transcripts (auto-transcribed by Teams and corrected by an author) were independently read by two coders, who familiarised themselves with the data and generated initial codes. The coders then met to organise and consolidate these codes into preliminary themes and sub-themes. Through a recursive review process, the coders refined and evaluated the emerging themes against both the coded excerpts and the full dataset. Once the themes were clearly defined and sufficiently concise, we produced the final report—the four high-level eHMI design requirements derived from this analysis, which was also used to inform the design of the VR user study:

- (1) **Key Vehicle States and State-Transition Must be Clearly Presented.** Both groups mentioned that there is a need to display key vehicle states and a clear state-transition design from slowing down to fully stopped, so that pedestrians can see the intention and state of the AV. For instance, H10 said: "*Let people see that there is a difference between the two states.*"
- (2) **Avoid Crossing Advice, Let Pedestrians Decide.** Our participants raised concerns that relying on cross advice from AVs could create liability and legibility issues. This is inline with previous works by Faas et al. [45], Zhang et al. [98] and Volvo<sup>9</sup>. Instead of giving crossing advice, they suggest that "*The choice has to be down to the pedestrian*" [D3]. Therefore, in line with works around non-disabled people [45, 98] and suggestion given by ISO technical report [1], eHMIs should present the AV's intention/state.
- (3) **Provide Multi-Modal eHMI.** Participants from both groups first acknowledged the effectiveness of visual eHMI designs, then explicitly mentioned that eHMI should be inclusive for them and others. They expressed the need for wider accessibility and inclusiveness, such as (1) ensuring the eHMI designs are accessible under different weather and light conditions and (2) making eHMI feasible for low vision users, colour-blind people, and people with multiple disabilities. One participant explained that "*for visual impaired people, visual might not be the best solution. They would probably want an audible cue as well.*" Therefore, reflecting the recent works on multi-modal eHMIs and their benefits for being inclusive and accessible to a wider audience [6, 21, 22, 31], there seems to be a need to develop multi-modal eHMIs.
- (4) **Provide Combined-Visual Communications.** Not just requesting multi-modal eHMI, participants suggested that the AV should employ multiple types of visual communications so that the pedestrians could pick up the visual eHMI that works for them (e.g., dyslexic people could refer to the *Abstract Light* signal, colour blind people could rely on the *Text/Symbol*) and rely on one visual eHMI if the other were damaged. "*There needs to be some kind of design that don't*

*show colour alone. You'd need some kind of other visual together.*" [H11].

In conclusion, our focus group findings suggested that *Abstract Light*, *Text*, and *Symbol* were the top three visual eHMI candidates and should be further evaluated in our VR user study. As multi-modal eHMI was highly recommended by our focus group participants, we briefly explored the use of auditory eHMI in our VR user study. As this formative study focuses on visual eHMI designs, we conclude the choice of the auditory eHMI to be *Speech* as its usefulness for other disabled people [21], which met the requirement of our participants (i.e., eHMI should be inclusive for DHH people and others).

We adjusted our designs: (1) display a clear state-transition of eHMI when the vehicle is slowing down and a clear state when the vehicle is fully stopped, and (2) avoid using instructive words or symbols. We proposed eHMI candidates that are (1) multi-modal eHMIs via audio-visual eHMI (Visual: *Abstract Light*, *Text*, and *Symbol*; Audio: *Speech*), and (2) single visual eHMIs (*Abstract Light* alone) or combined-visual eHMIs (*Abstract Light + Text* and *Abstract Light + Symbol*) to understand how DHH people perceive them.

## 4 VR User Study

This study was guided by three research questions (RQ):

RQ1: How do the ratings for experience (i.e., trust, acceptance, perceived safety, mental load) and behaviour (i.e., gaze behaviour, step-in road time, early step into the road count) differ between participants in the Hearing group and in the DHH group?

RQ2: What impact do the visual eHMIs have on pedestrians regarding experience and behaviour?

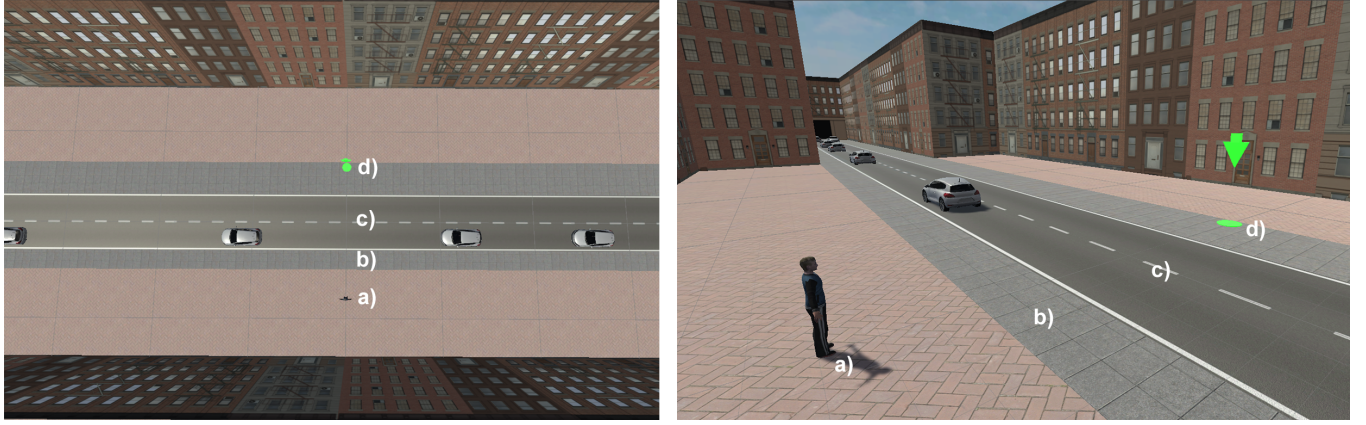
RQ3: What impact does providing auditory speech-based eHMIs have on pedestrians regarding experience and behaviour?

### 4.1 Virtual Environment

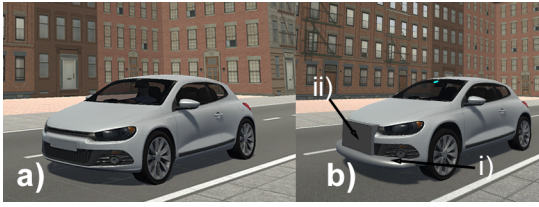
We used Unity V2022.3.44 to develop the virtual environment. Figure 1 illustrates a western-style city environment set up with a straight 2-way 1-lane road used in our study. Vehicles enter the area through a tunnel on one side and exit through another on the other side. Participants needed to start the trial about 4.7 m away from the first lane (see Figure 1a), arrive and wait at the grey pavement (see Figure 1b), and when there was enough gap or vehicles started to stop or stopped for the participant, they would need to cross the road (see Figure 1c) to the other end, until they reach the green waypoint (see Figure 1d), which would proceed them to the subsequent trial or finished the condition if it was the 2nd trial.

For the first lane, the one closer to the participant, there were constantly 10 vehicles, consisting of 1 AV and 9 manual vehicles. For the second lane, we set 2 manual vehicles to drive towards the right-hand side tunnel when launching the environment; these 2 vehicles would disappear and not reappear after entering the tunnel. This setting for the second lane was to raise awareness among the participants of vehicles potentially approaching from the left-hand side. At the same time, we do not want the vehicles in the second lane to affect the crossing decision and behaviour of the participants. The maximum speed of the city environment was set at 50 km/h, while vehicles drive from 40 km/h to 50 km/h. The participant will cross the road from a) to d) twice. Only the AV would stop for

<sup>9</sup>Volvo: Volvo 360c Concept



**Figure 1: Setup of the virtual environment:** The left shows a top-down view, while the right shows a leftward view. a) The starting position of the participant, which is 4.7 m from the road. b) The 2 m wide grey pavement area, which is used to inform the vehicle of the intention of crossing. c) The 6.8 m wide two-lane road that the participant needed to cross. d) The waypoint (green surface with downward arrow indicated), which is 2.5 m away from the road, is used to proceed with the partition to the next repetition. The human model was disabled in the experiment, placed here as a visual illusion.



**Figure 2: a) The Manual vehicle and b) the AV used in the study. i) The Light Strip, ii) The Display.**

the participants; however, we informed the participants that both manual vehicles and AVs may stop for them. The simulation also provided a light ambient background noise.

Figure 2 shows the appearance of the manual and AVs used in the user study. The AV used the same vehicle model as the manual vehicle but with three modifications in line with the prior work [18, 45, 73]: (1) a round cyan light positioned at the top centre of the vehicle's windshield to indicate that the vehicle is driving autonomously, (2) a light strip located at the bottom of the bumper region to display the light design, and (3) a display located at the grill region of the bumper to display the text or symbol. To further differentiate the AV and the manual vehicle, the manual vehicle features a human driver character inside the vehicle, while the AV does not involve any human character inside.

## 4.2 User Study Design

We employed a  $4 \times 2$  within-subjects design with Visual eHMI and Auditory eHMI as the independent variables. The 4 levels involved in the Visual eHMI (see Fig. 3) are:

- (1) *No Visual*. This condition serves as the baseline condition; both the light strip and the display would be presented but not activated; we want to ensure that differences measured

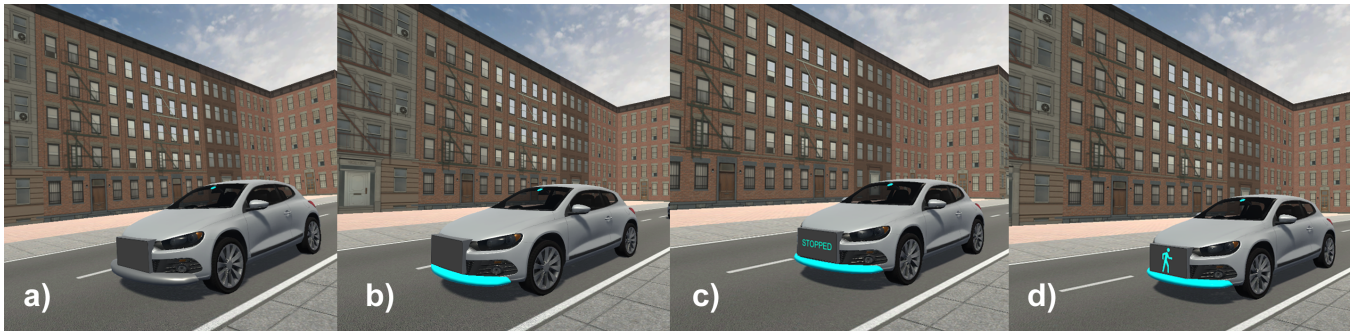
between the eHMI concepts solely rely on the eHMI concept itself.

- (2) *Abstract Light*. While the AV slows down the process, the light strip pulses in Cyan and the frequency goes from fast (pulsates between on and off at a rate of 1 Hz) to slow (pulsates between on and off at a rate of 0.5 Hz) to indicate a speed change. When fully stopped, the light strip would stop pulsing but remain static. Unless the vehicle started to yield for participants, neither the light strip nor the display would show any additional information. This would save energy consumption and make the vehicle more environmentally friendly.
- (3) *Abstract Light + Text*. The implementation of this condition is analogous to *Abstract Light*. The main difference is that the display would show "SLOWING" during the slowdown process and show "STOPPED" when the AV fully stopped. Both words were presented in bold cyan letters.
- (4) *Abstract Light + Symbol*. The implementation of this condition is analogous to *Abstract Light*. The main difference is that the display would show the "Standing Man" symbol during the slowdown process and the "Walking Man" symbol when the AV fully stopped, both symbols presented in colour cyan.

The 2 Auditory levels are:

- (1) *Without Speech*. This condition serves as the baseline condition with no added speech eHMI. We assumed all vehicles would be electric vehicles; therefore, we embedded the Acoustic Vehicle Alerting System (AVAS) sound from BMW into all vehicles. The AVAS had a volume of 65 dB.
- (2) *With Speech*. On top of the above condition, the AV would play a verbal message "I'm decelerating" after 1.5 s when the vehicle started to slow down. The length of the message is 1.3 s. When the vehicle fully stopped, it would immediately play a verbal message "I'm stopped", the length of the





**Figure 3: The visual eHMI conditions used in the user study, showing when the vehicle is fully stopped: a) *No Visual*: No added visual effects; b) *Abstract Light*: the light strip stayed on; c) *Text*: the light strip stayed on while the display shows "STOPPED"; d) *Symbol*: the light strip stayed on while the display shows a symbol of "Walking Man".**

verbal message is 0.8 s. Both messages only played once. We used a free online Text-to-Voice service<sup>10</sup> for generating the message.

The order of Visual eHMI  $\times$  Auditory conditions was counter-balanced in this study.

### 4.3 Apparatus and Setup

We used a Varjo XR-4 focal edition device as our VR headset, which offers a 90 Hz refresh rate with  $3840 \times 3744$  resolution and a  $120^\circ \times 105^\circ$  field of view. We also used the Varjo XR-4 built-in eye tracker, which offers 200 Hz eye tracking and provides gaze visualisation, video recording, and eye measurements such as pupil iris diameter, openness, and interpupillary distance<sup>11</sup>. To provide an immersive experience, we used Cyberith Virtualizer Elite 2 as our walking simulation for locomotion. Both devices were connected to a PC with an i9 CPU, 64 GB RAM, and a GeForce RTX 4090 Graphics card. Sound was provided by the VR device directly and the sound loudness level on the PC setting for Varjo headset was set to the same level across all participants; we did not employ any additional headphones or earphones, as they would clash with hearing aids or implants use. The study was conducted in a well-illuminated indoor laboratory.

### 4.4 Outcome Measurement

We used both subjective and objective measurements:

**4.4.1 Crossing Experience.** We used the same set of questionnaires (Trust in Automation [64], van der Laan acceptance scale [91], Perceived Safety [44]) as we used in the Formative Study, except the NASA-TLX question for Mental Demand. Instead, we employed the Low/High Index of Pupillary Activity (LHIPA) [38] for measuring the workload calculated based on the pupil diameter data provided by the Varjo-XR 4 device.

**4.4.2 Crossing Behaviour.** This was evaluated by Eye Behaviour data and Movement data. For Eye Gaze Behaviour, we recorded the number of fixations among three areas of interest (AOI) on the AV—(1) the Light Strip, (2) the display, (3) the other parts of

the AV. We reported the following 5 data metrics: (1) **Light Strip Duration**: Fixation on the Light Strip measured in second, (2) **Display Duration**: Fixation on the Display measured in second, (3) **Whole Vehicle Duration**: Fixation on the combination of Light Strip, Display, and other parts of Vehicle, measured in second (4) **Active Visual eHMI Duration**: duration on active visual eHMI for conditions that employ Visual eHMI, i.e., the Light Strip for *Abstract Light*, combination of the Light Strip and the Display for both *Abstract Light + Text* and *Abstract Light + Symbol* conditions, and (5) **Active Visual eHMI Percentage**: (Active Visual eHMI Duration / the Whole Vehicle Duration)  $\times 100\%$ .

We counted the time when the AV was 20 m away from the participant's crossing point till the participant stepped into the road. We choose 20 m as the starting point for us to calculate the results, because it provides higher accuracy, and early research suggests that pedestrians started to look more at the vehicle than the road ahead of the vehicle from this range [34].

*Step Into the Road Time* and *Early Step Into the Road Count* were reported as our Movement data. *Step Into the Road Time*: The time taken by the participant to start crossing the road from the moment an AV starts to slow down on the nearest lane. *Early Step Into the Road Count*: The number of times the participant stepped onto the road before the vehicle fully stopped.

**4.4.3 Post-Study Questionnaire. Necessity and Reasonability:** Following [18], we employed a 7-point (1=Totally Disagree to 7=Totally Agree) Likert scale to measure participants' perception of the Necessity and Reasonability of (1) visual and (2) auditory eHMI concepts. *Rank*: We also collected the ranking of the visual eHMI concepts and audio conditions; the lower the number, the better.

**4.4.4 Semi-structured Interview.** We first asked the participants to reflect on their design-making strategy: "Can you reflect on how you decided to cross in front of the approaching AV?" Then the interview moved to each visual eHMI design used in the studies (i.e., *Abstract Light*, *Abstract Light + Text*, and *Abstract Light + Symbol*), we asked three questions "Overall, what do you think about the [condition name]", "What did you like about this condition?", "What did you not like about this condition?". For auditory, we asked "Overall, what do you think about the verbal feedback?", "What did you like about it?", "What did you not like about it?", "Have the hearing

<sup>10</sup>TTSMP3 Text-to-Speech services; setting: US English / Kendra

<sup>11</sup>Varjo Developer Eye Tracking

**Table 2: Demographic information of DHH participants in the VR user study**

ID	Age	Gender	Left Ear HL Level	Right Ear HL Level	Preferred Communi- cation	Use Hear- ing Tech- nology?
P17	77	Male	Severe	Severe	English	Yes
P18	70	Male	Moderate	Moderate	English	Yes
P19	23	Female	Mild	Normal	English	No
P20	66	Female	Mild	Mild	English	Yes
P21	59	Female	Profound	Severe	English	Yes
P22	69	Female	Moderate	Severe	English	Yes
P23	28	Female	Moderate	Normal	English	Yes
P24	28	Female	Mild	Normal	English	No
P25	27	Male	Profound	Severe	BSL	Yes
P26	25	Female	Profound	Profound	BSL	No
P27	21	Male	Profound	Profound	English	Yes
P28	49	Female	Moderate	Mild	English	Yes
P29	33	Female	Profound	Profound	BSL	Yes
P30	41	Male	Profound	Profound	BSL	Yes
P31	61	Male	Moderate	Severe	English	Yes
P32	52	Female	Moderate	Profound	English	Yes

Hearing Loss (HL)

levels impacted your communication with the autonomous vehicle? Were you able to hear it?". In the end, we offer an open question to ask if participants have anything to add.

#### 4.5 Participants

Participants were recruited via physical posters, social media platforms, charities, and informal referrals through word-of-mouth. Thirty-two participants took part in the study. 16 (*Mean age* = 25.63, *SD* = 5.82; 10 women; 6 men) participants self-identified as having average hearing for both ears. The other 16 (*Mean age* = 45.56, *SD* = 19.49; 10 women; 6 men) participants self-identified as DHH people; full details can be found in Table 2. A BSL interpreter was presented in-person to assist with the study for BSL users.

All participants were instructed to cross as if they were in real life. For DHH participants, this meant they had the option to wear or not wear their hearing aids or implants, depending on their daily experience. Thirteen DHH participants who used hearing aids and implant daily decided to use it in this study as this is their everyday experience, 1 Deaf participant (i.e., P26) who used implant but only for indoor environment decided not to use it in this study as this is not usual for him, and 2 HoH participants were not using hearing aids at all and did not have one. DHH participants who used hearing aids or implants in the study were instructed to put it to their usual mode when they walk on the street.

On a 5-point Likert scale (1 = Not at all, 5 = Definitely), Hearing participants showed slight interest in AVs ( $M = 2.69$ ,  $SD = 1.20$ ) and had some knowledge about AV ( $M = 4.13$ ,  $SD = 1.02$ ). Similar results were found among DHH participants, while ratings for the interests in AVs had an average score of 2.63 ( $SD = 1.45$ ), and the ratings for their knowledge in AV were 3.88 ( $SD = 0.96$ ).

#### 4.6 Procedure

Each study started with a brief introduction, the consent form to review and sign, and a demographic questionnaire. Before the formal study, participants would wear the Varjo XR-4 and undergo two trials of road crossing in the same environment but without any vehicles present to familiarise themselves with the Virtualizer Elite 2. Once they were ready, they would proceed to the formal experiment conditions. In each condition, participants had to cross the street twice. Eye calibration was checked through Varjo Base software at the beginning of each condition. After each condition, participants answered the required questionnaires. At the end of the study, they completed a post-study questionnaire and then participated in a semi-structured interview.

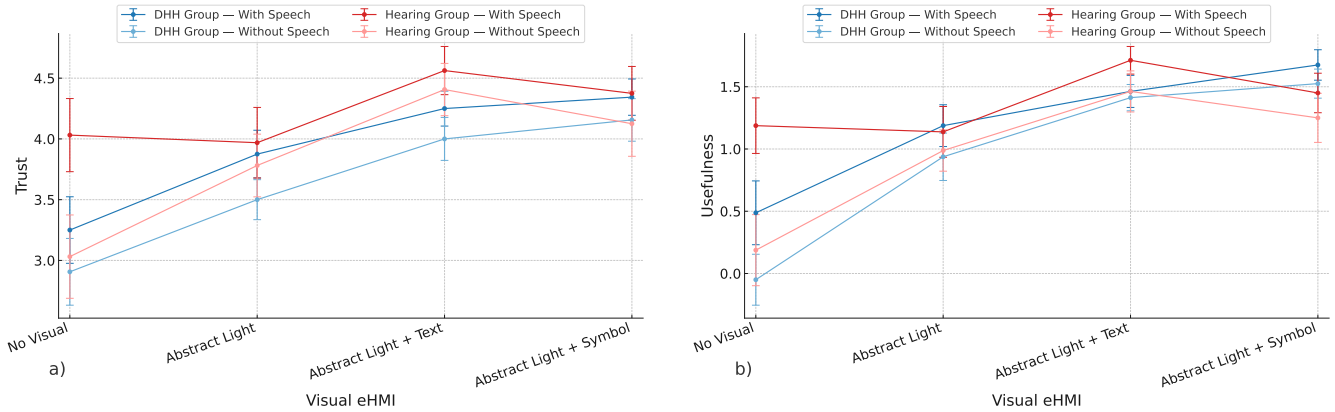
#### 5 Results

Shapiro-Wilk suggested that all data were not normally distributed, so we applied the Aligned Rank Transform (ART) to the data [94]. For questionnaire data, we employed a three-way mixed ANOVA with Visual (*No Visual*, *Abstract Light*, *Abstract Light + Text*, and *Abstract Light + Symbol*) and Audio (*With Speech* and *Without Speech*) as within-subjects variable and Group (1. Hearing and 2. DHH) as the between-subjects variable. For necessity and reasonability of visual and auditory eHMI, we employed a two-way mixed ANOVA after transforming the data with the ART method; the within-subject factor was Modality (Visual and Auditory), while the between-subject factor was the Group (Hearing and DHH). Post-hoc tests were conducted using the ART-C method [41] with Bonferroni corrections.

##### 5.1 Crossing Experience

**5.1.1 Trust in Automation.** Regarding Understandability, we could not observe any significant effect. As for Trust subscale, ANOVA results showed a significant main effect of Visual ( $F(3, 210) = 26.403$ ,  $p < .001$ ,  $\eta_p^2 = .274$ ) and Audio ( $F(1, 210) = 20.253$ ,  $p < .001$ ,  $\eta_p^2 = .088$ ) on Trust. Post-hoc pairwise comparisons revealed that the *Abstract Light + Symbol* ( $M = 4.25$ ,  $SD = 0.82$ ) and the *Abstract Light + Text* ( $M = 4.30$ ,  $SD = 0.75$ ) had significantly higher Trust ratings than the *Abstract Light* ( $M = 3.78$ ,  $SD = 0.93$ ,  $p < .001$ ) and the *No Visual* ( $M = 3.30$ ,  $SD = 1.25$ ,  $p < .001$ ). In addition, the *Abstract Light* ( $M = 3.78$ ,  $SD = 0.93$ ) had a better Trust rating than the *No Visual* ( $M = 3.30$ ,  $SD = 1.25$ ,  $p = .025$ ). Post-hoc pairwise comparisons on the main effect Audio showed that *With Speech* condition ( $M = 4.08$ ,  $SD = 0.97$ ) led to significantly better Trust ratings than *Without Speech* condition ( $M = 3.74$ ,  $SD = 1.07$ ,  $p < .001$ ). We did not find any interaction effect. Mean Trust ratings among each condition can be found in Figure 4a.

**5.1.2 Acceptance.** Usefulness. Mean Usefulness ratings among each condition across participants in Hearing and DHH groups can be found in Figure 4b. ANOVA tests yielded significant main effects of Visual ( $F(3, 210) = 42.088$ ,  $p < .001$ ,  $\eta_p^2 = .375$ ) and Audio ( $F(1, 210) = 27.10108$ ,  $p < .001$ ,  $\eta_p^2 = .114$ ). Post-hoc pairwise comparisons for the main effect Visual showed that both the *Abstract Light + Symbol* ( $M = 1.48$ ,  $SD = 0.61$ ) and the *Abstract Light + Text* ( $M = 1.51$ ,  $SD = 0.52$ ) had significantly higher Usefulness ratings than the *Abstract Light* ( $M = 1.06$ ,  $SD = 0.72$ ,  $p < .001$ ).



**Figure 4: a) Trust and b) Usefulness ratings by Visual × Audio eHMI conditions, separated for DHH and Hearing groups, with and without speech. Error bars show ±SE.**

and the *No Visual* ( $M = 0.45, SD = 1.06, p < .001$ ). In addition, the *Abstract Light* ( $M = 1.06, SD = 0.72$ ) had a better Usefulness rating than the *No Visual* ( $M = 0.45, SD = 1.06, p < .001$ ). Post-hoc pairwise comparisons for the main effect Audio showed that *With Speech* condition ( $M = 1.29, SD = 0.78$ ) resulted in a significantly higher Usefulness rating than *Without Speech* condition ( $M = 0.96, SD = 0.92, p < .001$ ).

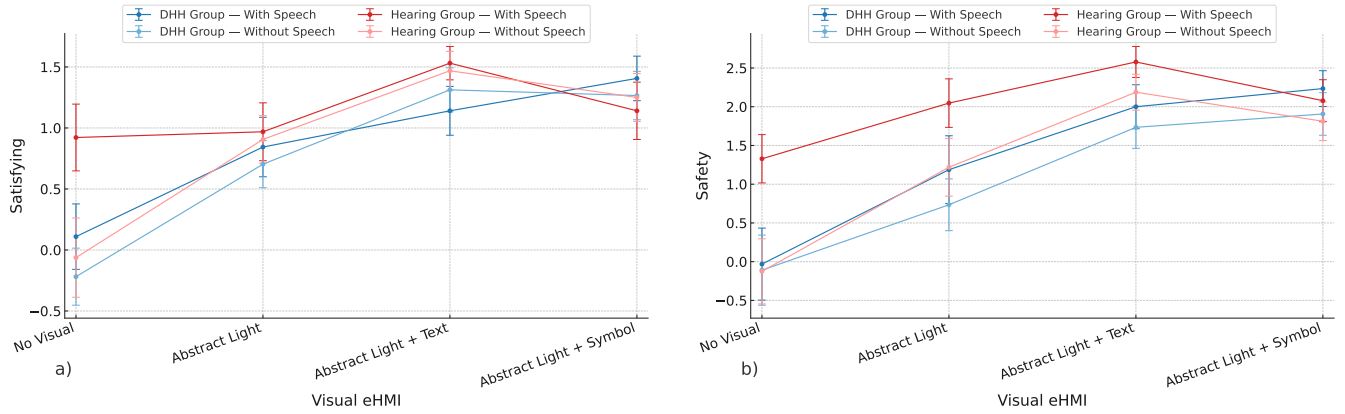
We observed an interaction effect of Visual × Group ( $F(3, 210) = 4.617, p = .004, \eta_p^2 = .062$ ) on Usefulness. Post-hoc pairwise comparisons results among the participants in the Hearing group showed that the *Abstract Light + Symbol* ( $M = 1.35, SD = 0.71$ ) was significantly better than the *No Visual* ( $M = 0.69, SD = 1.13, p < .001$ ), the *Abstract Light + Text* ( $M = 1.59, SD = 0.57$ ) was significantly better than the *Abstract Light* ( $M = 1.06, SD = 0.74, p = .001$ ) and the *No Visual* ( $M = 0.69, SD = 1.13, p < .001$ ). Post-hoc pairwise comparisons results among the DHH participants showed that the *Abstract Light + Symbol* ( $M = 1.60, SD = 0.48$ ) had a significantly higher Usefulness rating than the *Abstract Light* ( $M = 1.06, SD = 0.72, p < .001$ ) and the *No Visual* ( $M = 0.22, SD = 0.95, p < .001$ ). The *Abstract Light + Text* ( $M = 1.44, SD = 0.46$ ) and the *Abstract Light* ( $M = 1.06, SD = 0.72$ ) were both significantly rated higher than the *No Visual* ( $M = 0.22, SD = 0.95$ , both  $p < .001$ ).

We also observed an interaction effect of Visual × Audio ( $F(3, 210) = 5.157, p = .002, \eta_p^2 = .069$ ) on Usefulness. Post-hoc pairwise comparisons results among the *No Visual* condition showed that *With Speech* condition ( $M = 0.84, SD = 1.01$ ) resulted in a higher Usefulness score than *Without Speech* condition ( $M = 0.07, SD = 0.98, p = .003$ ). Post-hoc pairwise comparisons results also showed that when Speech was provided (1) the *Abstract Light + Symbol* ( $M = 1.56, SD = 0.57$ ) was significantly better than the *Abstract Light* ( $M = 1.16, SD = 0.74, p = .042$ ) and the *No Visual* ( $M = 0.84, SD = 1.01, p < .001$ ), (2) the *Abstract Light + Text* ( $M = 1.59, SD = 0.49$ ) was significantly better than the *No Visual* ( $M = 0.84, SD = 1.01, p < .001$ ). Post-hoc pairwise comparisons results showed that when Speech was not provided, (1) the *Abstract Light + Symbol* ( $M = 1.39, SD = 0.65$ ) had a significantly higher score than the *Abstract Light* ( $M = 0.96, SD = 0.70, p = .035$ ) and

the *No Visual* ( $M = 0.07, SD = 0.98, p < .001$ ), (2) the *Abstract Light + Text* ( $M = 1.44, SD = 0.55$ ) had a significantly higher score than the *Abstract Light* ( $M = 0.96, SD = 0.70, p = .007$ ) and the *No Visual* ( $M = 0.07, SD = 0.98, p < .001$ ), and (3) the *Abstract Light* ( $M = 0.96, SD = 0.70$ ) had a significantly higher score than the *No Visual* ( $M = 0.07, SD = 0.98, p = .003$ ).

Satisfying. Figure 5a illustrates mean Satisfying ratings among each condition across participants in the Hearing and DHH groups. ANOVA tests yielded significant main effect of Visual ( $F(3, 210) = 33.807, p < .001, \eta_p^2 = .326$ ) and Audio ( $F(1, 210) = 7.335, p = .007, \eta_p^2 = .034$ ). Post-hoc pairwise comparisons for the main effect Visual showed that both the *Abstract Light + Symbol* ( $M = 1.27, SD = 0.80$ ) and the *Abstract Light + Text* ( $M = 1.36, SD = 0.68$ ) had significantly higher Satisfying ratings than the *Abstract Light* ( $M = 0.86, SD = 0.86, p < .001$ ) and the *No Visual* ( $M = 0.19, SD = 1.17, p < .001$ ), in addition, the *Abstract Light* ( $M = 0.86, SD = 0.86$ ) had a significantly higher Satisfying rating than the *No Visual* ( $M = 0.19, SD = 1.17, p < .001$ ). Post-hoc pairwise comparisons for the main effect Audio showed that *With Speech* condition ( $M = 1.01, SD = 0.97$ ) resulted in a significantly higher Satisfying rating than *Without Speech* condition ( $M = 0.83, SD = 1.03, p = .007$ ).

We also observed an interaction effect of Visual × Audio ( $F(3, 210) = 4.643, p = .004, \eta_p^2 = .062$ ) on Satisfying. Post-hoc pairwise comparisons showed when Speech was provided, the *Abstract Light + Symbol* ( $M = 1.27, SD = 0.84$ ) and the *Abstract Light + Text* ( $M = 1.34, SD = 0.70$ ) were both significantly better than the *No Visual* ( $M = 0.52, SD = 1.14$ , both  $p < .001$ ). Post-hoc pairwise comparisons showed when Speech was not provided, (1) the *Abstract Light + Symbol* ( $M = 1.26, SD = 0.78$ ) was significantly better than the *Abstract Light* ( $M = 0.80, SD = 0.77, p = .026$ ) and the *No Visual* ( $M = -0.14, SD = 1.12, p < .001$ ), (2) the *Abstract Light + Text* ( $M = 1.39, SD = 0.67$ ) was significantly better than the *Abstract Light* ( $M = 0.80, SD = 0.77, p = .002$ ) and the *No Visual* ( $M = -0.14, SD = 1.12, p < .001$ ), and (3) the *Abstract Light* ( $M = 0.80, SD = 0.77$ ) was significantly better than the *No Visual* ( $M = -0.14, SD = 1.12, p = .002$ ).



**Figure 5: a) Satisfying and b) Perceived Safety ratings by Visual  $\times$  Audio eHMI conditions, separated for DHH and Hearing groups, with and without speech. Error bars show  $\pm$ SE.**

**5.1.3 Perceived Safety.** Figure 5b illustrates mean ratings among each condition across participants in the Hearing and DHH groups. ANOVA tests yielded significant main effect of Visual ( $F(3, 210) = 37.671, p < .001, \eta_p^2 = .350$ ) and Audio ( $F(1, 210) = 21.134, p < .001, \eta_p^2 = .091$ ) on Perceived Safety ratings. Post-hoc pairwise comparisons for the main effect Visual showed that the *Abstract Light + Symbol* ( $M = 2.01, SD = 1.02$ ) had a significantly higher Perceived safety rating than the *Abstract Light* ( $M = 1.30, SD = 1.51, p = .001$ ) and the *No Visual* ( $M = 0.27, SD = 1.74, p < .001$ ). Similarly, the *Abstract Light + Text* ( $M = 2.12, SD = 1.02$ ) had a significantly higher Perceived Safety rating than both the *Abstract Light* ( $M = 1.30, SD = 1.51, p < .001$ ) and the *No Visual* ( $M = 0.27, SD = 1.74, p < .001$ ). In addition, the *Abstract Light* ( $M = 1.30, SD = 1.51$ ) also had a significantly higher Perceived Safety rating than the *No Visual* ( $M = 0.27, SD = 1.74, p < .001$ ). Post-hoc pairwise comparisons for the main effect Audio showed that *With Speech* condition ( $M = 1.68, SD = 1.49$ ) resulted in a significantly higher Perceived Safety rating than *Without Speech* condition ( $M = 1.17, SD = 1.56, p < .001$ ).

We also observed an interaction effect of Visual  $\times$  Group ( $F(3, 210) = 3.162, p = .026, \eta_p^2 = .043$ ) on Perceived Safety. Post-hoc pairwise comparisons results among the participants in the Hearing group showed that the *Abstract Light + Symbol* ( $M = 1.95, SD = 1.03$ ) had a significantly higher rating than the *No Visual* ( $M = 0.60, SD = 1.63, p < .001$ ), the *Abstract Light + Text* ( $M = 2.38, SD = 0.88$ ) had a significantly higher rating than the *Abstract Light* ( $M = 1.63, SD = 1.42, p = .017$ ) and the *No Visual* ( $M = 0.60, SD = 1.63, p < .001$ ). The *Abstract Light* ( $M = 1.63, SD = 1.42$ ) had a significantly higher rating than the *No Visual* ( $M = 0.60, SD = 1.63, p = .006$ ). Post-hoc pairwise comparisons results among the participants in the DHH group showed that the *Abstract Light + Symbol* ( $M = 2.07, SD = 1.01$ ) had a significantly higher rating than the *Abstract Light* ( $M = 0.96, SD = 1.55, p < .001$ ) and the *No Visual* ( $M = -0.07, SD = 1.81, p < .001$ ). Similarly, the *Abstract Light + Text* ( $M = 1.87, SD = 1.10$ ) had a significantly higher rating than the *Abstract Light* ( $M = 0.96, SD = 1.55, p = .007$ ) and the *No Visual* ( $M = -0.07, SD = 1.81, p < .001$ ).

**5.1.4 Workload - Ihipa.** We observed an interaction effect of Visual  $\times$  Audio ( $F(3, 210) = 2.710, p = .046, \eta_p^2 = .037$ ). However, post-hoc pairwise comparisons showed no significant differences between conditions. No other main effects and interaction effects were found.

## 5.2 Crossing Behaviour

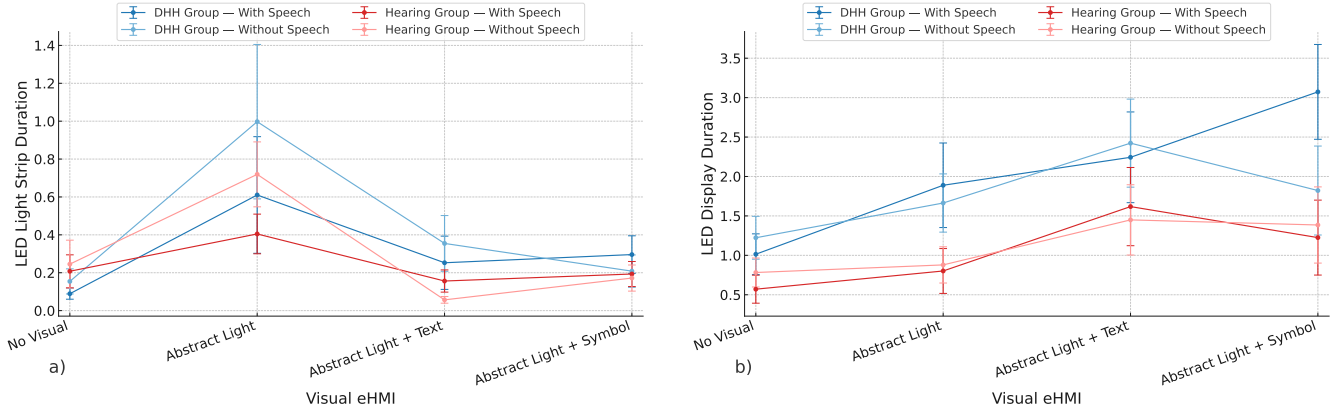
**5.2.1 Eye Behaviour.** All conditions were included for the data analysis of **Light Strip Duration**, **Display Duration**, and **Whole Vehicle Duration**. For **Active Visual eHMI Duration** and **Active Visual eHMI Percentage** analysis, we discarded No Visual condition as it did not involve active eHMIs.

**Light Strip Duration.** Figure 6a shows duration of participants in the Hearing and DHH groups spent on the Light Strip. ANOVA tests yielded significant main effect of Visual ( $F(3, 210) = 13.940, p < .001, \eta_p^2 = .166$ ) and Audio ( $F(1, 210) = 8.341, p = .004, \eta_p^2 = .038$ ). We also observed an interaction effect of Visual  $\times$  Audio ( $F(3, 210) = 4.167, p = .007, \eta_p^2 = .056$ ). We carried out post-hoc pairwise comparisons among the interaction effect, which indicated that when speech was not provided, participants looked significantly longer at the *Abstract Light* ( $M = 0.86, SD = 1.24$ ) than the *No Visual* ( $M = 0.20, SD = 0.40, p < .001$ ), as well as the *Abstract Light + Text* ( $M = 0.21, SD = 0.44, p = .003$ ) and the *Abstract Light + Symbol* ( $M = 0.19, SD = 0.31, p < .001$ ). We did not find other significant results.

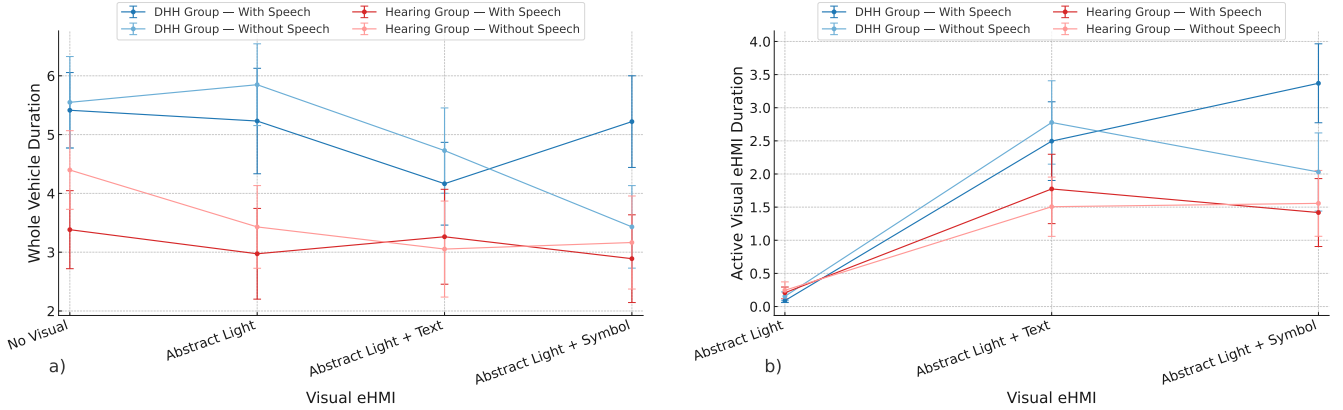
**Display Duration.** ANOVA tests yielded significant main effect of Visual ( $F(3, 210) = 6.163, p < .001, \eta_p^2 = .081$ ). Post-hoc pairwise comparisons for the main effect Visual showed that participants looked more frames on the Display in the *Abstract Light + Symbol* ( $M = 1.88, SD = 2.21, p = .002$ ) and the *Abstract Light + Text* ( $M = 1.93, SD = 2.08, p = .001$ ) conditions than in the *No Visual* condition ( $M = 0.90, SD = 0.92$ ). We also observed an interaction effect of Group  $\times$  Visual  $\times$  Audio ( $F(3, 210) = 2.672, p = .048, \eta_p^2 = .037$ ); however, post-hoc analysis did not yield any significant results. The duration participants spent on the Display can be found in Figure 6b.

**Whole Vehicle Duration.** Figure 7a shows duration of participants in the Hearing and DHH groups spent on the whole vehicle.





**Figure 6: a) Light Strip Duration and b) Display Duration by Visual × Audio eHMI conditions, separated for DHH and Hearing groups, with and without speech. Error bars show ±SE.**



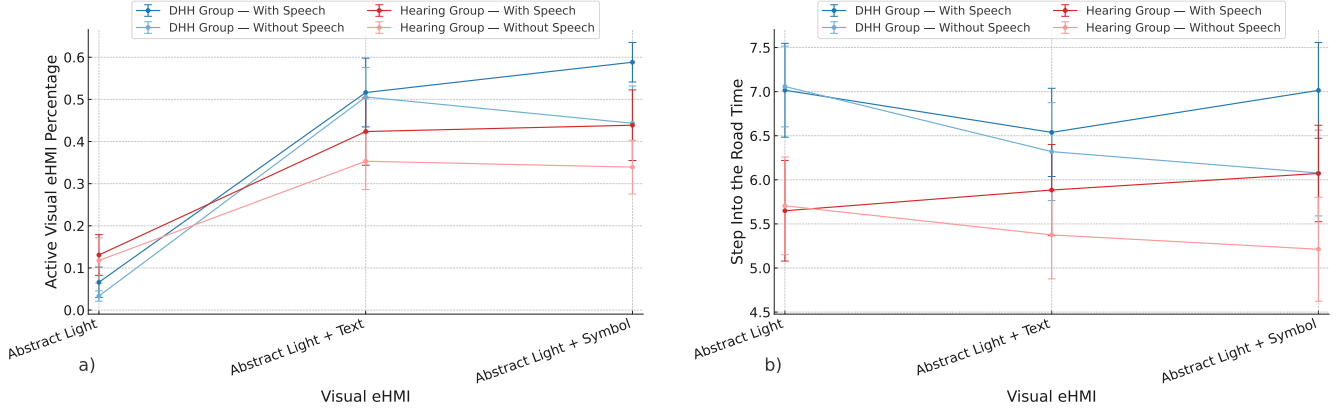
**Figure 7: a) Whole Vehicle Duration and b) Active Visual eHMI Duration by Visual × Audio eHMI conditions, separated for DHH and Hearing groups, with and without speech. Error bars show ±SE.**

ANOVA tests yielded significant main effect of Visual ( $F(3, 210) = 5.330, p = .001, \eta_p^2 = .071$ ) and Group ( $F(1, 30) = 4.271, p = .047, \eta_p^2 = .125$ ). Post-hoc pairwise comparisons for the main effect Visual showed that participants looked significantly more time on the vehicle in the *No Visual* condition ( $M = 4.69, SD = 2.83$ ) than in the *Abstract Light + Text* condition ( $M = 3.80, SD = 3.06, p = .013$ ), the *Abstract Light + Symbol* ( $M = 3.68, SD = 3.09, p = .003$ ). Post-hoc pairwise comparisons for the main effect Group showed that participants in the DHH group ( $M = 4.95, SD = 2.99$ ) looked at the vehicle significantly more time than participants in the Hearing group ( $M = 3.32, SD = 2.94, p = .048$ ).

**Active Visual eHMI Duration.** We conducted further analysis to explore how long participants looked at active eHMI. ANOVA tests yielded significant main effect of Visual ( $F(2, 150) = 74.271, p < .001, \eta_p^2 = .498$ ). Post-hoc pairwise comparison results suggest that participants spent more time on the active visual eHMI in *Abstract Light + Text* and *Abstract Light + Symbol* than *Abstract*

*Light* (both  $p < .001$ ). We also observed interaction effect between Visual × Group ( $F(2, 150) = 8.097, p < .001, \eta_p^2 = .097$ ), Visual × Audio ( $F(2, 150) = 3.305, p = .039, \eta_p^2 = .042$ ), and Visual × Audio × Group ( $F(2, 150) = 4.550, p = .012, \eta_p^2 = .057$ ). The post-hoc results of these three interaction effects confirmed the same post-hoc results of the Visual main effect. In addition, we also observed a significant interaction effect of Audio × Group ( $F(1, 150) = 4.485, p = .025, \eta_p^2 = .029$ ), but no significant post-hoc results were found. The gaze duration spent on active eHMI component can be found in Figure 7b.

**Active Visual eHMI Percentage.** Figure 8a illustrates percentage of gaze duration on active eHMI component among the whole vehicle. ANOVA tests yielded significant main effect of Visual ( $F(2, 150) = 57.038, p < .001, \eta_p^2 = .432$ ) and an interaction effect between Visual × Group ( $F(2, 150) = 6.710, p = .002, \eta_p^2 = .082$ ). Both post-hoc pairwise comparisons showed the same results where



**Figure 8: a) Active Visual eHMI Percentage and b) Step Into the Road Time by Visual  $\times$  Audio eHMI conditions, separated for DHH and Hearing groups, with and without speech. Error bars show  $\pm SE$ .**

participants spent a higher percentage of active visual eHMI components in *Abstract Light + Text* and *Abstract Light + Symbol* than in the *Abstract Light* condition ( $p < .001$ ).

**5.2.2 Movement Data.** We did not encounter a trial of participants crossing in front of the manual vehicles; meanwhile, no crashes were observed. Both data sets below covered 512 trials (32 participants  $\times$  8 conditions  $\times$  2 repetition) of crossings.

**Step Into the Road Time.** ANOVA tests yielded significant main effect of Visual ( $F(3, 210) = 5.684, p < .001, \eta_p^2 = .075$ ). Post-hoc pairwise comparisons for the main effect Visual revealed that participants spent significantly longer time to step into the road in the *No Visual* condition ( $M = 7.10s, SD = 2.50s$ ) than the *Abstract Light* ( $M = 6.36s, SD = 2.18s, p = .045$ ), the *Abstract Light + Text* ( $M = 6.03s, SD = 2.07s, p = .002$ ), and the *Abstract Light + Symbol* ( $M = 6.09s, SD = 2.21s, p = .006$ ) conditions. The value for each condition across participants in the Hearing and DHH groups can be found in Figure 8b.

**Early Step Into the Road Count.** ANOVA tests yielded significant main effect of Visual ( $F(3, 210) = 2.699, p = .047, \eta_p^2 = .037$ ). However, post-hoc pairwise comparisons showed no significant difference between the visual eHMI conditions, suggesting the number of times participants stepped onto the road before the vehicle fully stopped was similar across visual eHMI conditions.

### 5.3 Necessity and Reasonability for Visual and Auditory eHMIs

On average, participants in the Hearing group stated that there is a necessity for visual eHMIs ( $M = 5.88, SD = 1.31$ ) and that their use is reasonable ( $M = 5.69, SD = 1.49$ ). For auditory eHMIs, participants also perceived a moderate necessity ( $M = 5.06, SD = 1.73$ ) and found them reasonable ( $M = 5.31, SD = 1.49$ ). DHH Participants rated the necessity of visual eHMIs slightly higher ( $M = 6.44, SD = 0.51$ ) and found them to be reasonable ( $M = 6.31, SD = 0.70$ ). However, ratings for auditory eHMIs were lower, with perceived necessity ( $M = 4.31, SD = 2.15$ ) and reasonableness ( $M = 4.44, SD = 2.10$ ) showing more variability.

We were interested in understanding whether participants value the Visual and Auditory eHMI differently and whether hearing condition would impact this; therefore, we explored the impact of Modality (Visual vs Auditory) and Group (Hearing and DHH) on the necessity and reasonability ratings. As the data are not normally distributed, we employed ART ANOVA. ANOVA tests yielded significant main effect of Modality ( $F(1, 30) = 17.940, p < .001, \eta_p^2 = .374$ ) on necessity ratings. Post-hoc analysis suggested that the necessity ratings for the visual eHMIs were significantly higher ( $p < .001$ ) than the auditory eHMI. We could not find the main effect of the Group or any interaction effect. Regarding reasonability, ANOVA tests yielded a significant main effect of Modality ( $F(1, 30) = 9.48, p = .004, \eta_p^2 = .240$ ). Post-hoc analysis suggested that the reasonability ratings for the visual eHMIs were significantly higher than the auditory eHMI ( $p = .004$ ). We could not find the main effect of the Group or any interaction effect.

### 5.4 Ranking

The ranking of visual designs shows a preference for the *Abstract Light + Symbol* ( $M = 1.66, SD = 0.87$ ; 19 ranked it first while 8 ranked it second) and the *Abstract Light + Text* ( $M = 1.66, SD = 0.48$ ; 12 ranked it first while 18 ranked it second). They were followed by the *Abstract Light* ( $M = 2.81, SD = 0.74$ ; 1 ranked it first and 3 ranked it second, while dominantly ranked 3rd with 24 votes) and Base ( $M = 3.88, SD = 0.34$ ; 28 ranked it last and never ranked 1st or 2nd). Friedman's Test showed that there was a statistically significant difference in the ranking data among the eHMI concepts ( $\chi^2(5) = 65.51, p < .001$ ). Post-hoc analysis showed that the *Abstract Light + Symbol* was rated significantly better than the *Abstract Light* ( $p = .004$ ) and the *No Visual* ( $p < .001$ ). Similarly, the *Abstract Light + Text* was rated significantly better than the *Abstract Light* ( $p < .001$ ) and Base ( $p < .001$ ). In addition, we found that the *Abstract Light* was rated significantly better than Base ( $p < .001$ ).

Regarding the sound preference. 19 out of 32 participants ranked Speech as their first choice, while 13 ranked Without Speech first. The average rank score for Speech was 1.41 ( $SD = 0.50$ ), compared

to 1.59 ( $SD = 0.50$ ) for Without Speech. The difference between the ratings was not significant.

## 5.5 Qualitative Results

We conducted thematic analysis [12] with inductive coding on the interview transcripts (auto-transcribed by Teams and corrected by an author). In our reporting, we reported the insights related to each of the evaluated eHMIs, auditory eHMIs, and multi-modal eHMIs, with combined-visual eHMIs, along with selected participants' quotes (Hearing: P1-P16; DHH: P17-P32).

**5.5.1 Reflection on Decision Making strategy.** Several participants ( $N=22$ ; Hearing: 11, DHH: 11) mentioned that they either mainly rely on the visual eHMIs "*Relied mainly on visual cues (faster)*" [P4], "*Relied on visual cues; did not use audio cues*" [P17] or prioritized it over other cues "*relied more on visual cues than audio*" [P22] which could be due to their disability "*I don't trust my hearing*" [P18], "*primarily visual eHMI; sound was secondary help if attention was low*" [P12]. 3 participants (Hearing: 2, DHH:1) mentioned that they mainly relied on implicit communication (i.e., "*the car slows down*" [P7, P10, P26]. However, they also acknowledged that certain eHMI design (i.e., "*the walking man symbol helped me and made my decision*" [P26]. 1 participants (DHH:1) mentioned that they did not have a priority among these eHMIs, either eHMI work for them, they wait "*any signal from the vehicle, visual or audio cue*" [P19] and 1 participant (Hearing: 1) mentioned they would rely on both visual and auditory eHMI. 2 people from DHH group mentioned that due to the nature of deafness, they would rather check everything before they made the crossing decision (implicit cues, visual, auditory eHMIs, potential upcoming vehicle from other side of the road) [P32]. 2 participants (Hearing: 1; DHH:1) mentioned regardless of everything, they would wait till the vehicle fully stop.

**5.5.2 Reflection on Visual eHMIs. Abstract Light:** Several participants ( $N=25$ ; Hearing: 9, DHH: 16) expressed positive feeling of the light component and its design, suggesting it "*can see it from a distance*" [P3, P10, P16], "*aesthetically pleasing*", and "*reassuring that I'm seeing you*" [P4, P29, P31]. As for the design pattern, participants expressed they "*liked the blinking effects*" [P2, P8, P10, P14, P23, P30], it felt like the lights "*giving me the indication that it's gonna decelerating and it's gonna stop*" [P4, P14, P26].

However, several participants had raised concerns about the light component and its design, mainly because they believed it could be confusing, particularly when they first perceived the light design alone. A few participants ( $N=8$ ; Hearing: 3, DHH: 5) stated that the light "*was not understandable, not sure what the light was telling me to do*" [P5, P22, P31], and some participants ( $N=4$ ; Hearing: 2, DHH: 2) admitted that they "*had to guess the meaning of the light*" [P6, P12, P21, P25]. Additionally, a few participants ( $N=8$ ; Hearing: 3, DHH: 5) expressed concerns that "*it may not be clear for everybody*" [P16], "*some people may misinterpret the flashing*" [P26], "*could confuse people, is it part (design) of the car? Is it aesthetics?*" [P32], "*might be difficult for people who don't know what the light (is) for*" [P3, P9, P22, P23, P27].

We observed a strong preference for Combined-Visual or Multi-Modal Cues (i.e., combining visual, textual or symbolic and even auditory cues), 8 participants explicitly mentioned it (Hearing: 2,

DHH: 6). P15 said, "*It's more visual to have the light, but if you have something written or the symbol, it's even better*". A primary reason of the claim is due to the fact of confusion (there will be previously discussion on confusion) made by lights own it's own, "*lights on its own ... it's telling me what the car is doing, but it's not communicating to me what it wants me to do*." [P28].

**Abstract Light + Text:** Majority of participants ( $N=30$ ; Hearing: 15, DHH:15) expressed positive feeling of this combination and their design, stating it was "*clear*" [P2, P4, P7, P10, P22, P25, P30], "*easy to understand*" [P2, P18, P30], "*reassuring and confidence-building*" [P4, P10, P20, P23, P31], and "*reliable and helpful*" [P5, P17, P23, P26, P29]. Despite the majority of the people saying this combination was great, 3/32 participants expressed that using the Abstract Light itself is enough (P1, P2, P10) and "*you always see the light first*" [P2], with 1 participant arguing that "*I do not think the light was necessary*". 8/32 participants raised concerns towards the difficulties with the use of Text, as (1) it generates language barrier for "*some Deaf sign language people, English is not a preferred language for them*" [P26], (2) seeing the Text from distance could be challenge, with P31 stated "*not necessarily be able to read it at a distance*", and (3) it could be "*too much information (to see)*" [P1] and "*takes a bit more time (to process), it slows the reaction time*" [P23],—however, they later admitted that "*it is for safety*" to keep the Text.

**Abstract Light + Symbol:** Twenty-four participants (Hearing: 10, DHH: 14) expressed positive feelings towards this condition and its design. They stated that this was "*Clear*" [P23], "*probably less time to process than text*" [P27], "*reassuring*" [P4, P20]. Many people believe the benefits are due to the signs used looking like "*traffic signs*" ( $N=7$ ). In addition, they believed that using signs could work for people who cannot read [P2, P5, P9, P20], Children [P26], or non-native speakers [P1, P3]. However, 19 participants raised negative comments for Abstract Light + Symbol, while fewer participants ( $N=8$ ) raised negative comments for Abstract Light + Text. For instance, Almost half of the participants ( $N=14$ ) expressed that the Symbol could be unclear or contains abstract meaning (especially the "*standing man*"), they said there was an initial learning curve "*the first time, I wasn't quite sure what it meant*" [P19, P31], people could misinterpretation the Symbol or had wrong assumptions ( $N=5$ ). In addition, a few participants ( $N=7$ ) expressed that Symbol is harder to understand than Text as "*Text is more self-explanatory*" [P10, P11], while Symbol "*you have to actually interpret it; harder at a glance*" [P14].

**5.5.3 Reflection on Auditory eHMIs.** All Hearing group participants ( $N=16$ ) could hear the auditory eHMI from the vehicle. However, this was not the case for DHH participants; only 9/16 stated they could hear the auditory eHMI. 3 HoH participants struggled to hear the auditory eHMI, with [P17] stating that he could hear "Stopped" but not the "Deceleration", where [P18, P32] both admitted that "*I'm struggling*", and they focused more on the visual cues as "*visual is much more effective*" [P32] and are "*used to using visuals*" [P18]. All Deaf (BSL users) ( $N=4$ ) were unable to hear the words, two of them were unable to hear anything (including 1 wearing hearing aids); the other two were able to pick up a sound, but unable to understand the meaning of the sound—"I heard a little bit, but I don't know clearly what they are saying, there is a sound, but not a word and what the word is" [P25] with [P29] sharing

the similar quote *"I didn't understand that, I didn't know what was saying"*. However, both also noted that they were able to interpret the noise in the Abstract Light + Text condition *"I could hear that it said stopped and that helped me go."*

Over half of the participant sample (22/36) mentioned that the auditory eHMI brought benefits to them when they interacted with the AV, *"when audio is provided it felt good that the system tells people what the car is doing"* [P4, P21], *"audio served as a confirmation"* [P14, P21, P27], *"help awareness and decision-making"* [P12, P19], *"reinforced the (visual) message"* [P3, P8, P31], *"increase confidence"* [P13, P22, P31], and they just *"like/love it"* [P5, P10, P11]. However, 1 DHH participant [P17] disliked it *"hindrance rather than a help"* because *"Sound can interrupt or break up other processes, making it harder to concentrate"* and *"I only captured the word stop"* due to my hearing loss. And 1 Hearing participant disliked it as *"It feels uneasy and overloading"* [P14].

In addition to these benefits to themselves, 12 participants mentioned that having auditory eHMI could be helpful for *"people who are blind/cannot see/vision loss"* [P2, P15, P16, P26, P28, P29], *"people who have bad eye sight"* [P1, P7], *"children"* [P5], *"colour blind"* [P17], *"people who don't look up"* [P2], and the distracted people such as *"people wearing headphones"* [P3] or *"everyone if they are not paying attention"* [P9].

**5.5.4 Reflection on Design Requirements. State-Transition:** All participants noticed the changes in the design in different states (i.e., slowing down and fully stopped). Nineteen participants (Hearing: 8; DHH: 11) consistently valued state-transition feedback in eHMI design from the slowdown to fully stopped, noting it improved clarity, trust, and decision-making. Text and light cues that explicitly reflected vehicle state were described as intuitive and reassuring [P3, P6, P12, P16]. The progression helped users anticipate vehicle actions, especially when matched with natural behaviours. Participants also recommended (1) simplifying phrasing (e.g., avoiding technical terms like *"decelerating"*) to ensure quick comprehension, and (2) keeping the message of each state-transition consistent across different modalities.

**Multi-Modal eHMI:** We did not ask questions directly related to multi-modal eHMIs; however, 12 participants (Hearing: 5; DHH: 7) explicitly reported the need for multi-modal eHMIs (i.e., using visual and auditory together), with 1 participant also explicitly suggesting to include haptic feedback for the users through *"a bracelet or an IoT like the Apple Watch that will vibrate"* [P30]. They believe multi-modal eHMI increases clarity of communication, improves reaction times, and ensures inclusivity for people with diverse abilities and needs. For instance, P22 stated that *"everything (seeing and hearing the vehicle slow down) matched could increase confidence and sense of trust"*, while P13 said *"it gave double confirmations, increasing confidence to cross"*. In complex settings or to someone who might not be able to receive a cue from the vehicle, multi-modal eHMI increased accessibility through allowing people to choose their reliable source of information, which was confirmed by our participants (N=6; DHH: 6) with P18 said *"My hearing is quite bad. So I rely on visuals more than anything else"* and P29 explained that *"I relied more on the visual. For some people, maybe rely on the audio, people who are blind and so on, but for me, as a deaf person, no."* Despite there are benefits of having sound stimuli, 8 participants

(Hearing: 1, DHH:7) raised concerns on how effective this could be in the real world, *"in real world with more noise, more information, more people"* [P4], *"busy traffic or noisy environment, the audio might not be audible"* [P20], could be *"hard to detect"* [P25] as *"background noise could block the sound"* [P29, P32], *"making them unhelpful"* [P29] and *"not reliable"* [P32]. In addition, 3 participants argued that having multiple vehicles speak could exacerbate noise pollution and be perceived as *"annoying"* [P1], *"overwhelming"* [P7, P28].

**Combined-Visual eHMI:** Seven participants (Hearing: 1; DHH: 6) explicitly mentioned that there should be multiple visual cues on the visual eHMI design than simply using an Abstract Light eHMI, combining Abstract Light with Text or Symbol displays, to offer clearer, more informative, and more reassuring communication. They help communicate the vehicle's intention, increase user confidence over time, and align with familiar public signage conventions, making them more intuitive for diverse pedestrian groups.

## 6 Discussion

eHMI research still faces a significant lack of focus on evaluating these concepts with disabled people, especially DHH people. Our study is the first VR study that evaluates the effects of visual and auditory eHMI on DHH people's experiences (trust, acceptance, perceived safety, mental load) and behaviour (gaze behaviour, step-in road time, early step into the road count) while comparing it to hearing people. Our RQs were:

**RQ1: To what extent do the ratings for experience and behaviour differ between participants in the Hearing group and in the DHH group?** We found no significant difference between experience-related ratings from the Hearing and DHH group participants. However, crossing behaviour was affected when crossing in front of the AV in the tested scenario. Participants in the DHH group spent significantly more time looking at the AV than the Hearing group. Pecchini and Giuliani [82] had found that DHH pedestrians often had to exercise extreme caution towards manual drivers because of their inability to perceive auditory stimuli. Due to the absence of human drivers in AVs, we suspect DHH spent more time looking at the AV due to caution. This was partly supported by our qualitative data, where 2 participants mentioned they would wait till the AV fully stops, and 2 participants mentioned they would check everything before making the crossing decision, when nobody in the Hearing group mentioned similar comments.

We did notice interesting group-specific patterns. DHH participants rated *Abstract Light* significantly higher in usefulness than the No Visual condition, whereas this difference was not observed in the Hearing group. Conversely, Hearing group participants rated *Abstract Light* significantly higher than No Visual for perceived safety, but this pattern did not hold for DHH participants. These nuanced interaction patterns suggest that while the two groups provide similar overall experience ratings, they differ in how specific visual eHMIs support their crossing experience.

**RQ2: What impact do the visual eHMIs have on pedestrians regarding experience and behaviour?** In line with early works in video-based research [31, 61], VR research [18, 57], and real world Wizard-of-Oz research [45, 75], our results show that providing visual eHMI (i.e., *Abstract Light + Symbol* and *Abstract Light + Text*) on the AV improves subjective crossing assessments (trust,

usefulness, satisfying, and perceived safety). When compared to not having any eHMI, providing *Abstract Light* alone (1) improved usefulness rating among DHH people, (2) improved usefulness and satisfying ratings when speech was not provided, and (3) improved perceived safety among hearing people but not for DHH people.

Our gaze behaviour data further indicate that combined-visual eHMI positively impacted the participants' gaze behaviour; they spent more time on the active visual eHMI and less time on the AV. The crossing behaviour data suggest that participants spent less time stepping on the road, regardless of which visual eHMI was provided. Participants did not encounter any accidents. Therefore, our research clearly supports previous research in that having visual eHMIs assists in crossing decision making [57].

We observed a similar trend where the use of eHMI seems able to rivet and centralise pedestrians' visual attention to the active visual eHMI area, supporting previously reported findings in a video-based study [49]. In addition, we noticed that when *Text* and *Symbol* were active, more visual attention was paid to these visual designs than to the *Abstract Light*. One explanation could be that *Text* and *Symbol* were reported to be clearer and easier to understand when compared to *Abstract Light*, according to the qualitative insights provided by the participants. Participants still paid some attention to the active *Abstract Light* eHMI under *Abstract Light + Text* and *Abstract Light + Symbol* conditions, but they were much less than in the *Abstract Light* condition alone. Under *Abstract Light + Text* and *Abstract Light + Symbol* conditions, participants commented that they check the *Abstract Light* component first when the vehicle is far, which was supported by validating the recorded video with the participants' gaze movements—participants tend to focus on the *Abstract Light* at a far distance and then focus more on *Text* or *Symbol*.

**RQ3: What impact does providing auditory speech-based eHMIs have on pedestrians regarding experience and behaviour?** For our tested scenario, providing Speech-based eHMI generally improves the crossing experiences (trust, usefulness, perceived safety) compared to not providing it. This finding supports the auditory eHMI literature conducted in video-based research [31], simulation research [21], and real world Wizard-of-Oz research [9]. However, we did not observe a significant impact of providing Speech-based eHMI on the pedestrians' crossing behaviour, such as eye gaze behaviour and step into the road decision making, which is not in line with the real world Wizard-of-Oz research [9], where participants made faster crossing decisions when the intention of the vehicle was played.

To answer **RQ3**, we confirm that, for the scenario tested in this study, providing Speech-based auditory eHMI positively impacts crossing experience but does not impact crossing behaviour. The crossing behaviour, especially the speed-related movement measurements, could be improved if an advice/instruction message was used rather than the chosen state/intention message [39]; however, such message design could bring in liability and legibility issues [45, 98] as discussed in the Section 3.6.2.

It is also important to acknowledge that Speech-based auditory eHMIs may not be effective for all pedestrians, particularly sign language users whose first language is not English (i.e., the chosen language for this study) or those who cannot hear anything. Among the four sign language users we recruited, two were able to perceive

the speech sounds. However, they could not interpret the message correctly in most conditions, except in the *Abstract Light + Text* condition, where they could match the sound with the text, even though the spoken stopping message differed from the visual cue. For the other two participants who reported being unable to hear the speech, we conducted an additional check by replaying the sound through a headset after the experiment. Interestingly, one of these participants was able to hear the message when not engaged in the crossing task. We suspect this may be due to inattentive deafness [72], a phenomenon where people fail to notice auditory stimuli (i.e., speech message) when concentrating on a demanding visual task (i.e., crossing).

Nevertheless, three out of four sign-language users believed that Speech-based auditory eHMI is still helpful and would benefit more people. The other sign language user said that he did not care too much, as he could not hear the sound at all. Future work could explore other forms of multi-modal eHMI (e.g., haptics).

## 6.1 Decision Strategy

Twenty-four participants reported relying primarily on explicit communication cues (Visual: 22; Either Visual or Auditory: 1; Both: 1) when making crossing decisions in front of the AV, whereas only 3 participants stated that they depended solely on implicit cues such as vehicle kinematics. This finding supports prior works [9, 68, 90] that the presence of eHMI improves subjective feelings and plays an important role in decision-making. This contrasts with research on understanding pedestrian-manual driven vehicles communication, where Dey and Terken [32] found that explicit communication is rare to non-existent and does not play a significant role, implicit cues—particularly vehicle kinematics—played a more dominant role in making crossing decisions [67, 89]. We argue that even though pedestrians can cross only relying on implicit communication such as the AV's motions [80], the presence of active eHMIs is needed as it could attract pedestrians' attention and support better decision-making and a better crossing experience, especially during the slowdown process.

Deb et al. [28] suggested that visual eHMIs had more impact on the willingness to initiate crossing than auditory eHMIs. This is supported by our results, 22 out of 24 participants explicitly mentioned they relied on eHMI, indicating that they relied more on the visual than the auditory eHMI, while only 1 indicated the opposite. In addition, results of the necessity and reasonableness ratings also indicated that participants believed visual eHMIs were much more necessary and reasonable than the auditory counterpart.

## 6.2 Ecological Validity of Study Approach

Our design choices reflect a deliberate balance between experimental control and ecological relevance, which is consistent with prior eHMI research through VR (e.g., [17, 27, 57]) and CAVE simulators [65].

As lack of realism (appearance, content, task, and setting) in virtual world could negatively impact the validity [66], we ensured our study had a good realism by (1) using visually realistic vehicles and urban surroundings, (2) employing a familiar task (street crossing), and (3) situating the interaction in a typical urban road layout. In addition, to increase internal validity, we intentionally constrained

elements that could confound participants' responses: all vehicles had identical size/colour, yielding behaviours that were consistent, and no additional occlusion-inducing objects were present in the scene (e.g., parked vehicles, trees, bins). We also implemented two vehicles in the second lane at the beginning of the experiment to remind participants to check the left (most participants did check the left while crossing). We told the participants that both manual and automated vehicles may stop for them, although only the AV would stop for them. These implementations enable a clear interpretation of our study; they also come with the drawback of reduced ecological validity. Real world settings can often be more hazardous due to occlusions, variable approach vectors, multiple AVs, and mixed compliance [4, 48]; the crossing logic may be more dynamic, depending on the weather and time of day. Therefore, although our research provides meaningful early-stage research with DHH people, further testing is needed to explore crossing in more complex road setting.

Despite these simplifications, our findings remain meaningful and transferable. As noted by Recarte et al. [84], time-to-arrival estimates are consistent between video simulations and real-life situations. Early browser-based experiments reported that walking behaviour aligned closely with real world data [58]. Shen et al. [87] created a video-based tool to assess the safety of young pedestrians' street-crossing behaviour and found these video-based assessments to be both valid and reliable. Fuest et al. [47] compared video-based, VR-based, and Wizard-of-Oz vehicles' approach and observed only minor descriptive differences between them.

### 6.3 Practical Implications

We revisit the high-level design requirements identified in our formative study, reflect on these in light of our observations, and propose the following practical implications:

**6.3.1 Include Various Populations in the Design and Evaluation.** Despite most quantitative results being rated similarly (no significant difference) between participants in the Hearing and DHH groups, we found that gaze behaviour, like time spent on the vehicles, differed. In addition, many more DHH people asked for the need to have combined-visual eHMIs than participants in the hearing group. In line with Colley et al. [21], we call for future works to include various populations in the design and evaluation of eHMIs to make an inclusive eHMI for all people.

**6.3.2 Display Key State and State-Transition of the Vehicle.** Our qualitative findings suggest that conveying a vehicle's transition between states can enhance DHH pedestrians' clarity, trust, and decision-making when anticipating vehicle actions. We therefore recommend that eHMIs explicitly indicate the state transitions (i.e., slowing down) and the stop in scenarios where vehicles yield to pedestrians. Consideration should be taken when applying (1) *Text* for the visual design, technical terms such as decelerating or deaccelerating should be avoided in favour of plain, widely understood terms like slow or slow down; (2) *Speech* together with the *Text* as the visual eHMI, the terminology should remain consistent across the visual and auditory channels.

**6.3.3 Avoid Traffic Style Symbols.** Traffic signs convey meaning by a combination of symbols, shapes (e.g., circles–give orders,

triangles–warning, rectangles–provide information), and colours (e.g., Red–prohibition or danger, Blue–instructions or routes for specific traffic, Green–direction signs on primary routes). By avoiding input shapes and colours, we assumed the "Standing Man" symbol and the "Walking Man" symbol from typical traffic signs would not confuse participants that the symbols they see are not traffic signs. However, 7 out of 32 participants explicitly mentioned that they treated this as traffic-like signs, as we expected, but took "advice" of the "Walking Man" symbol. Although these symbols are easier for people, this may against our intention of not providing crossing advice, which could lead to liability and legibility issues [1, 45, 98]. Therefore, we encourage future work and design efforts to take this consideration (i.e., avoid using traffic style symbols) into account unless (1) broad public awareness or training campaign are in place to familiarise people with eHMIs [90], or (2) regulations evolve to include AVs as part of future traffic infrastructure [19].

**6.3.4 Enabling Multi-Modal eHMI.** Data from crossing experience, crossing behaviour, interview, and ranking showed a consistent preference for having a visual eHMI, suggesting all multi-modal systems should rely on visual as the base. Although the ranking data does not show a clear support for the claim that participants prefer to have Speech-based auditory eHMI over not having it, there is a general trend that employing Speech-based auditory eHMI enhances their crossing experiences (i.e., trust, acceptance-fulness, perceived safety). Therefore, our data support the notion that eHMI should be multi-modal and are suitable for people with DHH.

The benefits of having multi-modal eHMI were positively discussed by our participants and related works [21, 31]. However, transport noise ranks among Europe's top three environmental health threats; over 20% of Europeans are exposed to harmful transport noise levels [43]. Considerations are needed when employing auditory eHMI as part of the multi-modal eHMI; designers should ensure auditory eHMI comply with WHO's environmental noise guidelines [96] and are noticeable under different background noise levels [20], as there could be an impact of background noise on auditory eHMI [97].

**6.3.5 Enabling Combined-Visual eHMIs.** Literature and our qualitative feedback suggest that singular visual communication designs have limitations. Using *Abstract Light* alone to express intention may be unintuitive [86]. At the same time, *Text* appeared to be more easily understood than light [39], it may require more visual attention [8, 86] and is less noticeable than *Abstract Light* in far. *Symbols* have fewer language barriers and are easier to observe than *Text* [85], but are less noticeable at a distance than *Abstract Light*. Most critically, pedestrians may treat them as traffic signs.

Based on our findings with DHH people, we suggest future eHMI research/design to enable combined-visual communications as they lead to combined benefits, compensate for the limitation of singular visual communication, and outperform singular visual communication designs. Considering Symbols used in our study may be treated as traffic signs when participants are not educated with the meaning of eHMI, leading to liability issues Zhang et al. [98]. We recommend implementing *Abstract Light + Text* as a starting point. However, this recommendations need to be further tested with people with other disabilities. Additionally, we did not test



other visual methods, such as *Situational Awareness* and *Road-based Projection*, through VR simulation; further work should explore the combined-visual eHMI ideas with these designs.

## 6.4 Limitations and Future Work

This research has some limitations, which also highlight promising directions for future work. Given the limited number of prior research involving DHH participants, we followed previous study [28] and adopted a single controlled scenario featuring a non-signalised crossing. This setup effectively minimised external distractions, allowing participants to focus on the eHMI features being studied. However, such a controlled environment may constrain ecological validity (see Section 6.2). Future research should investigate more complex contexts, such as (1) mixed-traffic with more AVs in the scene, (2) crossings shared with other pedestrians [18], and (3) crossing in intersections or zebra crossings, to examine how these factors influence pedestrians' crossing experience and behaviour.

We only used the *Abstract Light* as the singular visual eHMI concept—the best singular eHMI candidate found through our formative study. We did not test *Text* or *Symbol* independently as singular eHMI. In addition, the lower ranked visual concepts from formative study (i.e., *Situational Awareness*, *Road-based Projection*, and *Anthropomorphic*) were not evaluated in VR as having them would significantly increase the duration of our VR study, hence not feasible to our participants, we encourage future work to further explore them in VR and field studies. For auditory eHMI, we only evaluated *Speech*, as it was considered more inclusive for low vision or blind people [21] and was preferred over other audible cues in prior research [28, 60]. However, *Speech* may not be an ideal medium for all DHH people, especially those whose primary language is sign language. Future work could benefit from deeper collaboration with DHH advocacy groups and from recruiting a more diverse pool of DHH participants—representing different linguistic and cultural backgrounds—to ensure the findings resonate globally. Moreover, exploring other modalities such as non-verbal auditory cues (e.g., music or bells [31, 97]) or haptic feedback could further enhance inclusivity and accessibility.

Our study involved only a single experimental session and was only tested in the UK. Future research could adopt a longitudinal design, extend participation to other groups of disabled pedestrians [6, 21, 88], and test the implications in different countries and cultures to determine if they would apply in the long term, across various countries and cultures, and for other types of disabilities.

## 7 Conclusion

In this research, we first conduct a formative study via focus groups with DHH people and relevant stakeholders to (1) inform the selection of visual eHMI candidates and (2) identify foundational design needs. Then, we investigate the effect of Visual eHMI (*No Visual*, *Abstract Light*, *Abstract Light + Text*, *Abstract Light + Symbol*) and Speech-based Auditory eHMI (*With Speech*, *Without Speech*) for AV-pedestrian communications in a VR study, with additional interest to explore differences regarding crossing experience and behaviours among participants in the hearing and DHH groups. Our results suggest three conclusions for the crossing scenario we evaluated: (1) Considerations should be paid to DHH people regarding eHMI

design, as they show different crossing behaviour (i.e., gaze duration on the AV). (2) Visual eHMIs, singular or combined-visual, improve both crossing experiences and crossing behaviour. (3) Auditory eHMI, like Speech-based auditory cues used in our experiment, might not help with crossing behaviour, but improve the overall crossing experiences. We also proposed five valuable implications that can enhance the inclusivity of AV-pedestrian communications accessible to DHH people.

## Acknowledgments

The authors thank all participants for their time. This work was funded by the Royal Society (RG/R1\241114).

## References

- [1] 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.
- [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. doi:10.1016/j.apergo.2018.11.002
- [3] Seonggeun Ahn, Dokshin Lim, and Byungwoo Kim. 2021. Comparative Study on Differences in User Reaction by Visual and Auditory Signals for Multimodal eHMI Design. In *HCI International 2021 - Posters*, Constantine Stephanidis, Margherita Antona, and Stavroula Ntoa (Eds.). Springer International Publishing, Cham, 217–223.
- [4] Javier Albusac, D. Vallejo, J.J. Castro-Schez, and C. Gzlez-Morcillo. 2018. An expert fuzzy system for improving safety on pedestrian crossings by means of visual feedback. *Control Engineering Practice* 75 (2018), 38–54. doi:10.1016/j.conengprac.2018.03.008
- [5] Savvas Argyropoulos, Konstantinos Moustakas, Alexey A. Karpov, Oya Aran, Dimitrios Tzovaras, Thanos Tsakiris, Giovanna Varni, and Byungjun Kwon. 2008. Multimodal user interface for the communication of the disabled. *Journal on Multimodal User Interfaces* 2, 2 (01 Sep 2008), 105–116. doi:10.1007/s12193-008-0012-2
- [6] Ashratuz Zavin Asha, Christopher Smith, Georgina Freeman, Sean Crump, Sowmya Somanath, Lora Oehlberg, and Ehud Sharlin. 2021. Co-Designing Interactions between Pedestrians in Wheelchairs and Autonomous Vehicles. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference (Virtual Event, USA) (DIS '21)*. Association for Computing Machinery, New York, NY, USA, 339–351. doi:10.1145/3461778.3462068
- [7] Julian Bäumler, Helen Bader, Marc-André Kaufhold, and Christian Reuter. 2025. Towards Youth-Sensitive Hateful Content Reporting: An Inclusive Focus Group Study in Germany. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 219, 22 pages. doi:10.1145/3706598.3713542
- [8] P. Bazilinsky, L. Kooijman, D. Dodou, and J.C.F. de Winter. 2021. How should external human-machine interfaces behave? Examining the effects of colour, position, message, activation distance, vehicle yielding, and visual distraction among 1,434 participants. *Applied Ergonomics* 95 (2021), 103450. doi:10.1016/j.apergo.2021.103450
- [9] Janina Bindschädel, Pauline Weimann, and Andrea Kiesel. 2023. Using eHMI, acoustic signal, and pitch motion to communicate the intention of automated vehicles to pedestrians: A Wizard of Oz study. *Transportation Research Part F: Traffic Psychology and Behaviour* 97 (2023), 59–72. doi:10.1016/j.trf.2023.06.017
- [10] 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 (Oldenburg, Germany) (AutomotiveUI '17)*. Association for Computing Machinery, New York, NY, USA, 136–140. doi:10.1145/3131726.3131765
- [11] Sjoukje E. Bouma, Juliette F. E. van Beek, Manna A. Alma, Ron L. Dierckx, Lucas H. V. van der Woude, Inge van den Akker-Scheek, and Martin Stevens. 2022. What affects the implementation of lifestyle interventions in patients with osteoarthritis? A multidisciplinary focus group study among healthcare professionals. *Disability and Rehabilitation* 44, 26 (2022), 8283–8293. doi:10.1080/09638288.2021.2011438
- [12] Virginia Braun and Victoria Clarke and. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. doi:10.1191/1478088706qp0630a
- [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 Press, New York, NY, USA, 70–75. doi:10.1109/IVS.2019.8814030
- [14] Angela Carlin, Marie H. Murphy, and Alison M. Gallagher. 2016. Do Interventions to Increase Walking Work? A Systematic Review of Interventions in Children and Adolescents. *Sports Medicine* 46, 4 (2016), 515–530. doi:10.1007/s40279-015-0432-6
- [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* (Oldenburg, Germany) (*AutomotiveUI '17*). Association for Computing Machinery, New York, NY, USA, 65–73. doi:10.1145/3122986.3122989
- [16] Michael Clamann, Miles Aubert, and Missy Cummings. 2017. Evaluation of Vehicle-to-Pedestrian Communication Displays for Autonomous Vehicles. In *Transportation Research Board 96th Annual Meeting*. The National Academies of Sciences, Engineering, and Medicine, Washington DC, United States, 13.
- [17] Mark Colley, Elvedin Bajrovic, and Enrico Rukzio. 2022. Effects of Pedestrian Behavior, Time Pressure, and Repeated Exposure on Crossing Decisions in Front of Automated Vehicles Equipped with External Communication. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (*CHI '22*). Association for Computing Machinery, New York, NY, USA, Article 367, 11 pages. doi:10.1145/3491102.3517571
- [18] Mark Colley, Julian Britten, and Enrico Rukzio. 2023. Scalability in External Communication of Automated Vehicles: Evaluation and Recommendations. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 2, Article 51 (jun 2023), 26 pages. doi:10.1145/3596248
- [19] Mark Colley, Surong Li, and Enrico Rukzio. 2021. Increasing Pedestrian Safety Using External Communication of Autonomous Vehicles for Signalling Hazards. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction* (Toulouse & Virtual, France) (*MobileHCI '21*). Association for Computing Machinery, New York, NY, USA, Article 20, 10 pages. doi:10.1145/3447526.3472024
- [20] Mark Colley and Enrico Rukzio. 2020. A Design Space for External Communication of Autonomous Vehicles. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) (*AutomotiveUI '20*). Association for Computing Machinery, New York, NY, USA, 212–222. doi:10.1145/3409120.3410646
- [21] Mark Colley, Marcel Walch, Jan Gugenheimer, Ali Askari, and Enrico Rukzio. 2020. Towards Inclusive External Communication of Autonomous Vehicles for Pedestrians with Vision Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3313831.3376472
- [22] Mark Colley, Marcel Walch, Jan Gugenheimer, and Enrico Rukzio. 2019. Including people with impairments from the start: external communication of autonomous vehicles. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings* (Utrecht, Netherlands) (*AutomotiveUI '19*). Association for Computing Machinery, New York, NY, USA, 307–314. doi:10.1145/3349263.3351521
- [23] 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* (Utrecht, Netherlands) (*AutomotiveUI '19*). Association for Computing Machinery, New York, NY, USA, 442–449. doi:10.1145/3349263.3351523
- [24] Mario Covarrubias, Elia Gatti, Monica Bordegoni, Umberto Cugini, and Alessandro Mansutti. 2014. Improving manual skills in persons with disabilities (PWD) through a multimodal assistance system. *Disability and Rehabilitation: Assistive Technology* 9, 4 (2014), 335–343. doi:10.3109/17483107.2013.799238 PMID: 23692410.
- [25] Chris Creed, Maadh Al-Kalbani, Arthur Theil, Sayan Sarcar, and Ian Williams. 2024. Inclusive augmented and virtual reality: A research agenda. *International Journal of Human-Computer Interaction* 40, 20 (2024), 6200–6219. doi:10.1080/10447318.2023.2247614
- [26] Koen de Clercq, Andre Dietrich, Juan Pablo Núñez Velasco, Joost de Winter, and Riender Happee. 2019. External Human-Machine Interfaces on Automated Vehicles: Effects on Pedestrian Crossing Decisions. *Human Factors* 61, 8 (2019), 1353–1370. doi:10.1177/0018720819836343 PMID: 30912985.
- [27] Shuchisnigdha Deb, Daniel W. Carruth, Muztaba Fuad, Laura M. Stanley, and Darren Frey. 2020. Comparison of Child and Adult Pedestrian Perspectives of External Features on Autonomous Vehicles Using Virtual Reality Experiment. In *Advances in Human Factors of Transportation*, Neville Stanton (Ed.). Springer International Publishing, Cham, 145–156.
- [28] 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. doi:10.1016/j.trf.2018.08.016
- [29] Debargha Dey, Azra Habibovic, Andreas Löcken, Philipp Wintersberger, Bastian Pfleging, Andreas Riener, Marieke Martens, and Jacques Terken. 2020. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives* 7 (2020), 100174. doi:10.1016/j.trip.2020.100174
- [30] Debargha Dey, Kai Holländer, Melanie Berger, Berry Eggen, Marieke Martens, Bastian Pfleging, and Jacques Terken. 2020. Distance-Dependent eHMIs for the Interaction Between Automated Vehicles and Pedestrians. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) (*AutomotiveUI '20*). Association for Computing Machinery, New York, NY, USA, 192–204. doi:10.1145/3409120.3410642
- [31] Debargha Dey, Toros Ufuk Senan, Bart Hengeveld, Mark Colley, Azra Habibovic, and Wendy Ju. 2024. Multi-Modal eHMIs: The Relative Impact of Light and Sound in AV-Pedestrian Interaction. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 91, 16 pages. doi:10.1145/3613904.3642031
- [32] 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* (Oldenburg, Germany) (*AutomotiveUI '17*). Association for Computing Machinery, New York, NY, USA, 109–113. doi:10.1145/3122986.3123009
- [33] Debargha Dey, Arjen van Vastenhoven, Raymond H. Cuijpers, Marieke Martens, and Bastian Pfleging. 2021. Towards Scalable eHMIs: Designing for AV-VRU Communication Beyond One Pedestrian. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) (*AutomotiveUI '21*). Association for Computing Machinery, New York, NY, USA, 274–286. doi:10.1145/3409118.3475129
- [34] Debargha Dey, Francesco Walker, Marieke Martens, and Jacques Terken. 2019. Gaze Patterns in Pedestrian Interaction with Vehicles: Towards Effective Design of External Human-Machine Interfaces for Automated Vehicles. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Utrecht, Netherlands) (*AutomotiveUI '19*). Association for Computing Machinery, New York, NY, USA, 369–378. doi:10.1145/3342197.3344523
- [35] André Dietrich, Jan-Henning Willrodt, Karolin Wagner, and Klaus Bengler. 2018. Projection-Based External Human Machine Interfaces – Enabling Interaction between Automated Vehicles and Pedestrians. In *Proceedings of the Driving Simulation Conference 2018 Europe VR* (2018-09-05), Andras Kemeny, Florent Colombet, Frédéric Merienne, and Stéphane Espié (Eds.). Driving Simulation Association, Driving Simulation Association, Antibes, France, 43–50.
- [36] Jinzhen Dou, Shuang Chen, Zhi Tang, Chang Xu, and Chengqi Xue. 2021. Evaluation of Multimodal External Human-Machine Interface for Driverless Vehicles in Virtual Reality. *Symmetry* 13, 4 (2021), 17. doi:10.3390/sym13040687
- [37] S. Duarte, L. Cunha, L. Moreira, S. Ferreira, and A. Lobo. 2023. Requirements and Expectations for Truck Platooning – A Multidisciplinary Perspective. In *Safety Management and Human Factors (AHFE Open Access, Vol. 105)*, Pedro Arezes and Anne Garcia (Eds.). AHFE International, USA, 9. doi:10.54941/ahfe1003072
- [38] Andrew T. Duchowski, Krzysztof Krejtz, Nina A. Gehrer, Tanya Bafna, and Per Bækgaard. 2020. The Low/High Index of Pupillary Activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3313831.3376394
- [39] Y.B. Eisma, A. Reiff, L. Kooijman, D. Dodou, and J.C.F. de Winter. 2021. External human-machine interfaces: Effects of message perspective. *Transportation Research Part F: Traffic Psychology and Behaviour* 78 (2021), 30–41. doi:10.1016/j.trf.2021.01.013
- [40] Y. B. Eisma, S. van Bergen, S. M. ter Brake, M. T. T. Hensen, W. J. Tempelaar, and J. C. F. de Winter. 2020. External Human-Machine Interfaces: The Effect of Display Location on Crossing Intentions and Eye Movements. *Information* 11, 1 (2020), 18. doi:10.3390/info11010013
- [41] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) (*UIST '21*). Association for Computing Machinery, New York, NY, USA, 754–768. doi:10.1145/3472749.3474784
- [42] Md Atik Enam, Ananta Bastola, and Julian Brinkley. 2024. Are the External Human-Machine Interfaces (eHMI) Accessible for People with Disabilities? A Systematic Review. In *2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS)*. IEEE Press, Toronto, ON, Canada, 1–6. doi:10.1109/ICHMS59971.2024.10555703
- [43] European Environment Agency. 2025. *Environmental Noise in Europe 2025*. EEA Report 05/2025. European Environment Agency, Luxembourg: Publications Office of the European Union. doi:10.2800/1181642 Report published 24 June 2025; based on Environmental Noise Directive data from the 2022 reporting round.
- [44] Stefanie M. Faas, Andrea C. Kao, and Martin Baumann. 2020. A Longitudinal Video Study on Communicating Status and Intent for Self-Driving Vehicle Pedestrian Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3313831.3376484

- [45] Stefanie M. Faas, Lesley-Ann Mathis, and Martin Baumann. 2020. External HMI for self-driving vehicles: Which information shall be displayed? *Transportation Research Part F: Traffic Psychology and Behaviour* 68 (2020), 171–186. doi:10.1016/j.trf.2019.12.009
- [46] Evelyn Florentine, Mark Adam Ang, Scott Drew Pendleton, Hans Andersen, and Marcelo H. Ang. 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* (Biopolis, Singapore) (HAI '16). Association for Computing Machinery, New York, NY, USA, 387–392. doi:10.1145/2974804.2974833
- [47] Tanja Fuest, Elisabeth Schmidt, and Klaus Bengler. 2020. Comparison of Methods to Evaluate the Influence of an Automated Vehicle's Driving Behavior on Pedestrians: Wizard of Oz, Virtual Reality, and Video. *Information* 11, 6 (2020), 24. doi:10.3390/info11060291
- [48] Thomas Golob and Will Recker. 2001. Relationships Among Urban Freeway Accidents, Traffic Flow, Weather and Lighting Conditions. *Institute of Transportation Studies, UC Berkeley, Institute of Transportation Studies, Research Reports, Working Papers, Proceedings* 129 (01 2001), 24. doi:10.1061/(ASCE)0733-947X(2003)129:4(342)
- [49] Fu Guo, Wei Lyu, Zenggen Ren, Mingming Li, and Ziming Liu. 2022. A Video-Based, Eye-Tracking Study to Investigate the Effect of eHMI Modalities and Locations on Pedestrian-Automated Vehicle Interaction. *Sustainability* 14, 9 (2022), 21. doi:10.3390/su14095633
- [50] Kamer Gür, Eda Kılınc, Enes Günaydin, and Hilal Gülsün. 2021. The epidemiology of injury among adolescents with hearing loss, health beliefs regarding injury and associated factors. *Disability and Health Journal* 14, 2 (2021), 100994. doi:10.1016/j.dhjo.2020.100994
- [51] Mathias Haimerl, Mark Colley, and Andreas Riener. 2022. Evaluation of Common External Communication Concepts of Automated Vehicles for People With Intellectual Disabilities. *Proc. ACM Hum.-Comput. Interact.* 6, MHCI, Article 182 (Sept. 2022), 19 pages. doi:10.1145/3546717
- [52] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, Amsterdam, Netherlands, 139–183. doi:10.1016/S0166-4115(08)62386-9
- [53] Zhifan He, Zhengyu Tan, Ruifo Zhang, Yanyan Li, and Bin Liu. 2021. How Pedestrian-AV Interaction Is Affected by the eHMI: A Virtual Reality Experiment. In *Advances in Usability, User Experience, Wearable and Assistive Technology*, Tareq Z. Ahram and Christianne S. Falcão (Eds.). Springer International Publishing, Cham, 707–714.
- [54] Marion Hersh, James Ohene-Djan, and Saduf Naqvi. 2010. Investigating Road Safety Issues and Deaf People in the United Kingdom: An Empirical Study and Recommendations for Good Practice. *Journal of prevention & intervention in the community* 38 (10 2010), 290–305. doi:10.1080/10852352.2010.509021
- [55] Marion. A. Hersh. 2016. Improving Deafblind Travelers' Experiences: An International Survey. *Journal of Travel Research* 55, 3 (2016), 380–394. doi:10.1177/0047287514546225
- [56] Isti Hidayati, Wendy Tan, and Claudia Yamu. 2021. Conceptualizing Mobility Inequality: Mobility and Accessibility for the Marginalized. *Journal of Planning Literature* 36, 4 (2021), 492–507. doi:10.1177/08854122211012898
- [57] Kai Holländer, Ashley Colley, Christian Mai, Jonna Häkikilä, Florian Alt, and Bastian Pfleger. 2019. Investigating the Influence of External Car Displays on Pedestrians' Crossing Behavior in Virtual Reality. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (Taipei, Taiwan) (MobileHCI '19). Association for Computing Machinery, New York, NY, USA, Article 27, 11 pages. doi:10.1145/3338286.3340138
- [58] Kai Holländer, Luca Schellenberg, Changkun Ou, and Andreas Butz. 2020. All Fun and Games: Obtaining Critical Pedestrian Behavior Data from an Online Simulation. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–9. doi:10.1145/3334480.3382797
- [59] Elahi Hossain, Anna L. Cox, Anna Dowthwaite, and Yvonne Rogers. 2024. Adaptive, Sociable and Ready for Anything: Undergraduate Students Are Resilient When Faced with Technological Change. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 198 (April 2024), 32 pages. doi:10.1145/3653689
- [60] Christopher R. Hudson, Shuchisnigdha Deb, Daniel W. Carruth, John McGinley, and Darren Frey. 2019. Pedestrian Perception of Autonomous Vehicles with External Interacting Features. In *Advances in Human Factors and Systems Interaction*, Isabel L. Nunes (Ed.). Springer International Publishing, Cham, 33–39.
- [61] Philip Joisten, Ziyu Liu, Nina Theobald, Andreas Webler, and Bettina Abendroth. 2021. Communication of Automated Vehicles and Pedestrian Groups: An Intercultural Study on Pedestrians' Street Crossing Decisions. In *Proceedings of Mensch Und Computer 2021* (Ingolstadt, Germany) (MuC '21). Association for Computing Machinery, New York, NY, USA, 49–53. doi:10.1145/3473856.3474004
- [62] Young Woo Kim, Jae Hyun Han, Yong Gu Ji, and Seul Chan Lee. 2020. Exploring the Effectiveness of External Human-Machine Interfaces on Pedestrians and Drivers. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) (AutomotiveUI '20). Association for Computing Machinery, New York, NY, USA, 65–68. doi:10.1145/3409251.3411725
- [63] Andrew J. Kolarik, Brian C. J. Moore, Pavel Zahorik, Silvia Cirstea, and Shahina Pardhan. 2016. Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, & Psychophysics* 78 (2016), 373–395. doi:10.3758/s13414-015-1015-1 Published online 2015 Nov 20.
- [64] Moritz Körber. 2019. Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*, Sebastiano Bagnara, Riccardo Tartaglia, Sara Albolino, Thomas Alexander, and Yushi Fujita (Eds.). Springer International Publishing, Cham, 13–30.
- [65] Yee Mun Lee, Ruth Madigan, Chinebuli Uzundu, Jorge Garcia, Richard Romano, Gustav Markkula, and Natasha Merat. 2022. Learning to interpret novel eHMI: The effect of vehicle kinematics and eHMI familiarity on pedestrian' crossing behavior. *Journal of Safety Research* 80 (2022), 270–280. doi:10.1016/j.jsr.2021.12.010
- [66] Letitia Lew, Truc Nguyen, Solomon Messing, and Sean Westwood. 2011. Of course I wouldn't do that in real life: advancing the arguments for increasing realism in HCI experiments. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada) (CHI EA '11). Association for Computing Machinery, New York, NY, USA, 419–428. doi:10.1145/1979742.1979621
- [67] 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* (Toronto, ON, Canada) (AutomotiveUI '18). Association for Computing Machinery, New York, NY, USA, 188–197. doi:10.1145/3239060.3239082
- [68] Hailong Liu, Takatsugu Hirayama, and Masaya Watanabe. 2021. Importance of Instruction for Pedestrian-Automated Driving Vehicle Interaction with an External Human Machine Interface: Effects on Pedestrians' Situation Awareness, Trust, Perceived Risks and Decision Making. In *2021 IEEE Intelligent Vehicles Symposium (IV)*. IEEE Press, Nagoya, Japan, 748–754. doi:10.1109/IV48863.2021.9575246
- [69] 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* (Utrecht, Netherlands) (AutomotiveUI '19). Association for Computing Machinery, New York, NY, USA, 262–274. doi:10.1145/3342197.3344544
- [70] Victor Malmsten Lundgren, Azra Habibovic, Jonas Andersson, Tobias Lagström, Maria Nilsson, Anna Sirkka, Johan Fagerlön, Rikard Fredriksson, Claes Edgren, Stas Krupenia, and Dennis Saluäär. 2017. Will There Be New Communication Needs When Introducing Automated Vehicles to the Urban Context?. In *Advances in Human Aspects of Transportation*, Neville A. Stanton, Steven Landry, Giuseppe Di Bucchianico, and Andrea Vallicelli (Eds.). Springer International Publishing, Cham, 485–497.
- [71] Jörgen Lundälv. 2004. Self-reported experiences of incidents and injury events in traffic among hearing impaired people as pedestrians and cyclists. A follow-up study of mobility and use of hearing equipment. *International Journal of Rehabilitation Research* 27 (03 2004), 79–80. doi:10.1097/01.mrr.0000119280.12233.9d
- [72] James S. P. Macdonald and Nilli Lavie. 2011. Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics* 73, 6 (2011), 1780–1789. doi:10.3758/s13414-011-0144-4
- [73] 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* (San Diego, CA, USA) (DIS '19). Association for Computing Machinery, New York, NY, USA, 475–486. doi:10.1145/3322276.3322328
- [74] 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* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3173574.3174003
- [75] Milecia Matthews, Girish Chowdhary, and Emily Kieson. 2017. Intent Communication between Autonomous Vehicles and Pedestrians. arXiv:1708.07123 [cs.HC] <https://arxiv.org/abs/1708.07123>
- [76] Abby McCormack and Heather Fortnum. 2013. Why do people fitted with hearing aids not wear them? *International Journal of Audiology* 52, 5 (2013), 360–368. doi:10.3109/14992027.2013.769066 PMID: 23473329.
- [77] Denisa Qori McDonald, Shruti Mahajan, Richard Vallett, Genevieve Dion, Ali Shokoufandeh, and Erin Solovey. 2022. Interaction with Touch-Sensitive Knitted

- Fabrics: User Perceptions and Everyday Use Experiments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 426, 20 pages. doi:10.1145/3491102.3502077
- [78] 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. doi:10.1016/j.aap.2018.03.018
- [79] Sharifah Roziah Binti Mohd Kassim, Shujun Li, and Budi Arief. 2023. Understanding How National CSIRTs Evaluate Cyber Incident Response Tools and Data: Findings from Focus Group Discussions. *Digital Threats* 4, 3, Article 45 (Oct. 2023), 24 pages. doi:10.1145/3609230
- [80] Dylan Moore, Rebecca Currano, G. Ella Strack, and David Sirkin. 2019. The Case for Implicit External Human-Machine Interfaces for Autonomous Vehicles. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Utrecht, Netherlands) (AutomotiveUI '19). Association for Computing Machinery, New York, NY, USA, 295–307. doi:10.1145/3342197.3345320
- [81] Jenna Rachel Panter and Andy Jones. 2010. Attitudes and the Environment as Determinants of Active Travel in Adults: What Do and Don't We Know? *Journal of Physical Activity and Health* 7, 4 (2010), 551 – 561. doi:10.1123/jpah.7.4.551
- [82] Dario Pecchini and Felice Giuliani. 2015. Street-Crossing Behavior of People with Disabilities. *Journal of Transportation Engineering* 141, 10 (06 2015), 04015022. doi:10.1061/(ASCE)TE.1943-5436.0000782
- [83] Nathaniel R. Peterson, David B. Pisoni, and Richard T. Miyamoto. 2010. Cochlear implants and spoken language processing abilities: Review and assessment of the literature. *Restorative Neurology and Neuroscience* 28, 2 (2010), 237–250. doi:10.3233/RNN-2010-0535 PMID: 20404411
- [84] Miguel Ángel Recarte, Ángela Conchillo, and Luis Miguel Nunes. 2005. Estimation of arrival time in vehicle and video. *Psicothema* 17, 1 (2005), 112–117.
- [85] Michael Rettenmaier, Jonas Schulze, and Klaus Bengler. 2020. How Much Space Is Required? Effect of Distance, Content, and Color on External Human-Machine Interface Size. *Information* 11, 7 (2020), 15. doi:10.3390/info11070346
- [86] Alexandros Rouchitsas and Håkan Alm. 2022. Ghost on the Windshield: Employing a Virtual Human Character to Communicate Pedestrian Acknowledgement and Vehicle Intention. *Information* 13, 9 (2022), 35. doi:10.3390/info13090420
- [87] Zhuo Shen, Jinfei Ma, and Ning Wang. 2023. Development and validation of a video-based assessment tool for children's street-crossing safety. *Journal of Transport & Health* 33 (2023), 101716. doi:10.1016/j.jth.2023.101716
- [88] Pj Standen, F Rees, and David Brown. 2009. Effect of playing computer games on decision making in people with intellectual disabilities. *Journal of Assistive Technologies* 3 (07 2009), 8. doi:10.1108/17549450200900011
- [89] Matus Sucha, Daniel Dostal, and Ralf Rissler. 2017. Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention* 102 (2017), 41–50. doi:10.1016/j.aap.2017.02.018
- [90] Tram Thi Minh Tran, Callum Parker, Marius Hoggenmüller, Yiyuan Wang, and Martin Tomitsch. 2024. Exploring the Impact of Interconnected External Interfaces in Autonomous Vehicles on Pedestrian Safety and Experience. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 89, 17 pages. doi:10.1145/3613904.3642118
- [91] Jinke D. Van Der Laan, Adriaan Heino, and Dick De Waard. 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies* 5, 1 (1997), 1–10. doi:10.1016/S0968-090X(96)00025-3
- [92] Holly S. Vitense, Julie A. Jacko, and V. Kathlene Emery. 2002. Multimodal feedback: establishing a performance baseline for improved access by individuals with visual impairments. In *Proceedings of the Fifth International ACM Conference on Assistive Technologies* (Edinburgh, Scotland) (Assets '02). Association for Computing Machinery, New York, NY, USA, 49–56. doi:10.1145/638249.638260
- [93] Yingna Wang, Qingqin Liu, Xiaoying Wei, and Mingming Fan. 2025. Facilitating Daily Practice in Intangible Cultural Heritage through Virtual Reality: A Case Study of Traditional Chinese Flower Arrangement. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 1126, 17 pages. doi:10.1145/3706598.3713409
- [94] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) (CHI '11). Association for Computing Machinery, New York, NY, USA, 143–146. doi:10.1145/1978942.1978963
- [95] World Health Organization. 2021. *World Report on Hearing*. World Health Organization, Geneva. 252 pages. <https://www.who.int/publications/i/item/9789240020481>
- [96] World Health Organization Regional Office for Europe. 2018. *Environmental Noise Guidelines for the European Region*. Technical Report. World Health Organization, Copenhagen, Denmark. <https://www.who.int/publications/i/item/9789289053563> WHO Regional Office for Europe guideline.
- [97] Wenge Xu, Foroogh Hajiseyedjavadi, Debargha Dey, Tram Thi Minh Tran, and Mark Colley. 2026. Exploring the Impacts of Background Noise on Auditory Stimuli of Audio-Visual eHMs for Hearing, Deaf, and Hard-of-Hearing People. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems* (CHI '26), April 13–17, 2026, Barcelona, Spain (Barcelona, Spain) (CHI '26). Association for Computing Machinery, Barcelona, Spain, 16 pages. doi:10.1145/3772318.3791557
- [98] Jingyi Zhang, Erik Vinkhuyzen, and Melissa Cefkin. 2018. Evaluation of an Autonomous Vehicle External Communication System Concept: A Survey Study. In *Advances in Human Aspects of Transportation*. Springer International Publishing, Cham, 650–661. doi:10.1007/978-3-319-60441-1\_63
- [99] Yunfang Zheng, Jacob Swanson, Janet Koehnke, and Jianwei Guan. 2022. Sound Localization of Listeners With Normal Hearing, Impaired Hearing, Hearing Aids, Bone-Anchored Hearing Instruments, and Cochlear Implants: A Review. *American Journal of Audiology* 31, 3 (2022), 819–834. doi:10.1044/2022\_AJSLP-22-00006