# Should an external human-machine interface flash or just show text? A study with a gaze-contingent setup

Eisma, Yke Bauke; van Gent, Lucas; de Winter, Joost

**DOI**
10.1016/j.trf.2023.07.005

**Publication date**
2023

**Document Version**
Final published version

**Published in**
Transportation Research Part F: Traffic Psychology and Behaviour

Transportation Research Part F: Psychology and Behaviour 97 (2023) 140–154

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

# Should an external human-machine interface flash or just show text? A study with a gaze-contingent setup

Yke Bauke Eisma, Lucas van Gent, Joost de Winter *

*Department of Cognitive Robotics, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Netherlands*

A R T I C L E   I N F O

*Keywords:*
External human–machine interfaces
Peripheral vision
Response times
Acceptance

A B S T R A C T

Automated vehicles need to prioritize pedestrian safety. One way to achieve this is through external human–machine interfaces (eHMIs) that send visual signals to pedestrians. eHMIs can be either text-based or light-based. However, there has been limited research on the effects of these types of eHMI on human information processing and attention allocation. This study aimed to fill this gap by using a gaze-contingent approach, which blurs the view outside a circular aperture, to test the hypothesis that text-based eHMIs, which require focused or foveal attention, result in longer response times compared to light-based eHMIs, which can be understood using peripheral vision. In this study, 23 participants watched animated video clips of traffic situations involving automated vehicles with either no eHMI, a flashing-light eHMI, or a text-based eHMI. Their eye movements were tracked, and they were asked to press the spacebar when they felt it was safe to cross the road. The results showed faster response times when an eHMI was present, with no significant difference between the two types of eHMIs. Further analysis suggested that the flashing-light eHMI captured attention briefly, while the text-based eHMI held attention for a longer period. When no eHMI was present, participants focused on the approaching vehicle for the longest time. The gaze-contingent window resulted in fewer eye movements and slower response times. In conclusion, the study showed that the gaze-contingent window negatively affected response times and eye movements, emphasizing the importance of considering peripheral vision when designing eHMIs for pedestrian safety.

## 1. Introduction

Automated vehicles (AVs) should be designed to prioritize the safety of pedestrians and other vulnerable road users (Berge et al., 2022; GOV.UK, 2022; SWOV, 2022; US DOT, 2018). This can be achieved by implementing technology that allows the AV to detect the presence of pedestrians and to respond accordingly to avoid collisions and unsafe situations (Combs et al., 2019; Park et al., 2017). Additionally, it may be necessary to implement new regulations and laws that address the risks posed by AVs, such as the ability for remote problem-solving and more stringent speed limits for AVs in areas where pedestrians are present (European Union, 2022; NACTO, 2016; NHTSA, 2016). It may also be necessary to educate pedestrians about how to interact with AVs so that they know what to expect (see Faas et al., 2021; Gremmelmaier et al., 2022, for discussions).

In addition to ongoing progress in enhancing AV technology itself, it is equally vital to consider the pedestrian's perspective by designing AVs to interact in a predictable and comprehensible manner. Specifically, external human–machine interfaces (eHMIs) could

* Corresponding author.
*E-mail address:* j.c.f.dewinter@tudelft.nl (J. de Winter).

https://doi.org/10.1016/j.trf.2023.07.005
Received 11 March 2023; Received in revised form 9 July 2023; Accepted 10 July 2023
Available online 17 July 2023
1369-8478/

be implemented to convey the AV's intentions and pedestrian detection status through visual signaling to surrounding road users (Bazilinskyy et al., 2021; Dey et al., 2021; Epke et al., 2021; Mahadevan et al., 2018; Verstegen et al., 2021). This approach may facilitate pedestrians making more appropriate navigation decisions through enhanced situational awareness and projected vehicle behavior.

The human eye uses both foveal and peripheral vision (Granit & Harper, 1930; Strasburger et al., 2011). Foveal vision, the small area within 2.5° of the fixation point (Sakurai, 2015), refers to the detailed sight obtained from the central part of the retina, called the fovea. Foveal vision is important for tasks requiring high visual acuity, such as reading and recognizing faces. Peripheral vision, on the other hand, refers to vision outside the direct line of sight, covering a much wider field, beyond 2.5° (for convenience, parafoveal and peripheral vision are grouped here as 'peripheral vision'). The periphery of the retina provides less detailed information but is important for detecting changes in the environment, such as impending obstacles or dangers (Finlay, 1982; Franchak & Adolph, 2010; Huestegge & Böckler, 2016; Strasburger et al., 2011). When a salient or task-relevant stimulus in the periphery is detected, it typically triggers a saccade toward that stimulus to explore it in further detail (Jacobs et al., 1989; Vater et al., 2016). Pedestrians normally focus their foveal vision (i.e., rotate the eyes) on relevant environmental cues, such as approaching vehicles, the far side of the road, and visual landmarks (De Winter et al., 2021a; Geruschat et al., 2003; Lévêque et al., 2020). Simultaneously, they monitor the environment with peripheral vision to detect potential hazards (e.g., Marigold, 2008; Uttley et al., 2017; Vater et al., 2022).

A common safety issue is that pedestrians are not always visually attentive; for instance, their foveal vision could be engrossed in a smartphone (Basch et al., 2014; Feld & Plummer, 2019; Larue et al., 2019). A related phenomenon is that pedestrians sometimes cross the street without directing their gaze towards the rest of the traffic, implying that peripheral vision plays a significant role in their crossing behavior (Vater et al., 2022). Additional research in pedestrian-AV interaction found that pedestrians may not realize whether a human driver is present or absent, or do not rely on driver-state information to make their crossing decision (Lee et al., 2021; Moore et al., 2019; Rothenbücher et al., 2016). In addition to conditions like glaucoma that impair the visual field and hence affect pedestrian safety (Hassan et al., 2005; Turano et al., 1999), the visual field can also experience a functional impairment. The so-called functional visual field (e.g., Sanders, 1970) or useful field of view (UFOV; Sekuler et al., 2000) describes the area of the visual field from which information can be extracted without moving the eyes or head. Performance on UFOV tests, and related tests such as the detection response task, is not only negatively correlated with age (Edwards et al., 2006), but also tends to narrow under conditions of cognitive load such as during conversing and memorization (Atchley & Dressel, 2004; Conti et al., 2014; Rantanen & Goldberg, 1999), stress (Bursill, 1958; Weltman et al., 1971), and visual load (Mackworth, 1965; Wood et al., 2006; Van Winsum, 2019a). It is noteworthy here that nonvisual cognitive tasks seem to impair visual perception uniformly throughout the visual field, rather than inducing a specific reduction in the peripheral regions (Ringer et al., 2016; Van Winsum, 2019b). In summary, research clearly highlights how heavily pedestrians rely on peripheral vision, acknowledging that peripheral vision is compromised by factors intrinsic to individuals, like age, as well as other factors commonly experienced by pedestrians, such as increased mental and visual loads.

As noted above, the effectiveness of an eHMI may largely depend on its ability to attract peripheral attention. For example, a bright or flashing light may elicit a faster response (possibly before the pedestrian focuses on it) than an eHMI displaying text, which pedestrians must focus on first to read. However, as of present, little research has examined the fundamental question of which modality (text or light) is optimal for presenting an eHMI. Indeed, although a large number of studies have investigated eHMIs that employ textual and light/lamp elements (e.g., Bazilinskyy et al., 2022b; Chen et al., 2023; Eisele & Petzoldt, 2022; Guo et al., 2022b; Kunst et al., 2022; Lau et al., 2022; Schmidt-Wolf & Feil-Seifer, 2022), there is a lack of systematic comparison between these types of interfaces in terms of attentional measurements, such as eye tracking. This has resulted in a dilemma where on the one hand, various researchers have recommended text- and icon-based eHMIs because they can convey semantically rich information (e.g., 'Please proceed to cross', 'Waiting', 'Stopping', or 'After you'; Eisma et al., 2020; Ferenchak & Shafique, 2022; Hudson et al., 2019; Löcken et al., 2019; see Bazilinskyy et al., 2022a for an overview of 44 papers involving text-based eHMIs for AVs), while, on the other hand, authors and guidelines caution against text-based eHMIs because they may not be attention-grabbing and require foveal vision (Cefkin, 2018; Dey et al., 2022; Tabone et al., 2021; and see ISO, 2018; UNECE, 2019, promoting light signal eHMIs). As pointed out above, bright light signal eHMIs have the potential to attract attention, especially when moving or flashing (Hensch et al., 2019), but caution is needed as the meaning of the light signal could be misinterpreted (Lee et al., 2019).

To address the research gap regarding the choice between text- or light-signal eHMIs, the present study employs a gaze-contingent paradigm, a technique to control the extent to which the participant has access to peripheral vision (Duchowski et al., 2004; Reingold et al., 2003). Gaze-contingent methods have previously been used to simulate visual impairments (Biebl et al., 2022; Glen et al., 2016) and to simulate tunnel vision on computer screens or in virtual environments (Mao et al., 2021). In the current human-subject experiment, we used a gaze-contingent window consisting of a circular aperture with blurry vision outside of it. This was done to simulate how pedestrians focus their foveal attention on specific objects, while other areas are less detailed, as they would experience in the real world.

The present study involved human participants watching animated video clips of a road crossing scenario, in which automated cars equipped with different types of eHMIs (a text message or a flashing light) drove by. Both types of eHMIs were presented to the participant with and without a gaze-contingent window, and two sizes of gaze-contingent windows were used to simulate two levels of access to peripheral vision: a normal-vision condition and a condition that may occur under stress and cognitive demands (cf. Rantanen & Goldberg, 1999). In this study, we tested the hypothesis that a text-based eHMI is particularly disadvantageous when presented peripherally, because it might prompt the pedestrian to fixate on it for reading or verification purposes. Additionally, it can be expected that a flash-based eHMI is relatively easy to detect regardless of whether it is presented centrally or peripherally, while a text-based eHMI may be less detectable if it appears peripherally. Consequently, this could result in a slower response time as compared to conditions under clear simulated vision. In summary, we anticipated an interaction between the eHMI format (flash-based vs. text-based) and the size/presence of the gaze-contingent viewport in relation to the speed at which pedestrians arrived at a crossing decision. Along with investigating response times, the current study also examined the effects of the eHMIs and gaze-contingent window manipulations on eye movements.

## 2. Methods

### 2.1. Participants

The experiment was performed with 23 participants (18 male, 5 female) between 21 and 67 years old ($M = 31.3$, $SD = 13.8$). The participants were acquaintances of the experimenter and students of the Faculty of Mechanical, Maritime and Materials Engineering at the Delft University of Technology, the Netherlands. Our decision to include 23 participants was informed by previous investigations that elucidated the effects of different eHMI conditions on pedestrian crossing behavior (e.g., De Clerq et al., 2019: 28 participants; Mok et al., 2022: 17 pedestrians). All participants provided written informed consent. The research was approved by the TU Delft Human Research Ethics Committee (approval no. 2047).

### 2.2. Apparatus

Participants' eye movements were recorded binocularly at a sampling rate of 2000 Hz using an Eyelink 1000 Plus eye tracker (SR Research, Ottawa, ON, Canada). The stimuli were presented on a 24.5-inch BenQ monitor. Although the monitor had greater pixel capability, we configured its resolution to 1280 × 720 pixels for the experiment and set the refresh rate at 144 Hz. The distance between the participant's eyes and the monitor was approximately 100 cm. The dimensions of the screen were 544 × 303 mm, resulting in a horizontal viewing angle of 30.4 deg and a vertical viewing angle of 17.2 deg. The experimental setup is shown in Fig. 1.

### 2.3. Independent variables

There were two independent variables: (1) eHMI type (Flash, Text, No eHMI), and (2) Gaze-contingent window (Small, Large, None).

Both eHMI signals were presented on a main screen located on the front of the vehicle, measuring 1000 × 400 $mm^2$. In addition, two smaller screens were placed on the left and right fronts of the car, respectively. The Flash eHMI consisted of white rectangular light. Upon activation, it flashed twice as follows: On (170 ms), Off (170 ms), On (170 ms). The screens were visible when the car was approaching head-on as well as when turning into the road. The windshield was not see-through, so no driver could be seen. The Text eHMI displayed the message 'Waiting', together with an icon of a zebra crossing and a walking pedestrian (as in Eisma et al., 2020). Once the text message turned on, it stayed on until the end of the video clip. Both eHMIs were white to maximize contrast and to prevent ambiguity in regard to the meaning of colors that are already used in traffic (Bazilinskyy et al., 2021; Eisma et al., 2020).

The gaze-contingent window consisted of a circular aperture that followed the participant's gaze in real-time. Outside the circle, the view of the scene was blurred. FFmpeg (The FFmpeg Developers, 2021) was used to apply Gaussian blur with $\sigma = 10$ to the original video clip, and the number of steps used for Gaussian approximation was 6. The small gaze-contingent window had a diameter of 150 pixels, corresponding to a visual span of 3.7°, whereas the large window had a diameter of 300 pixels, corresponding to a visual span of 7.3°. Note that the foveal region of the eyes can be defined as the area within a span of 5° around the fixation point (Sakurai, 2015). In the current study, foveal vision was thus left relatively intact, but peripheral vision was artificially degraded.

Fig. 2 presents a selection of illustrative screenshots from the video clips (1280 × 720 pixels). In addition, Fig. 3 showcases cropped screenshots (293 × 164 pixels) capturing the moment when the eHMI was activated and its equivalent instance under the No eHMI

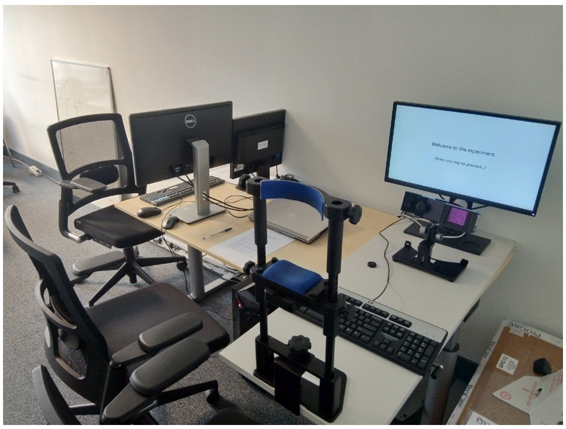

**Fig. 1.** The experimental setup.

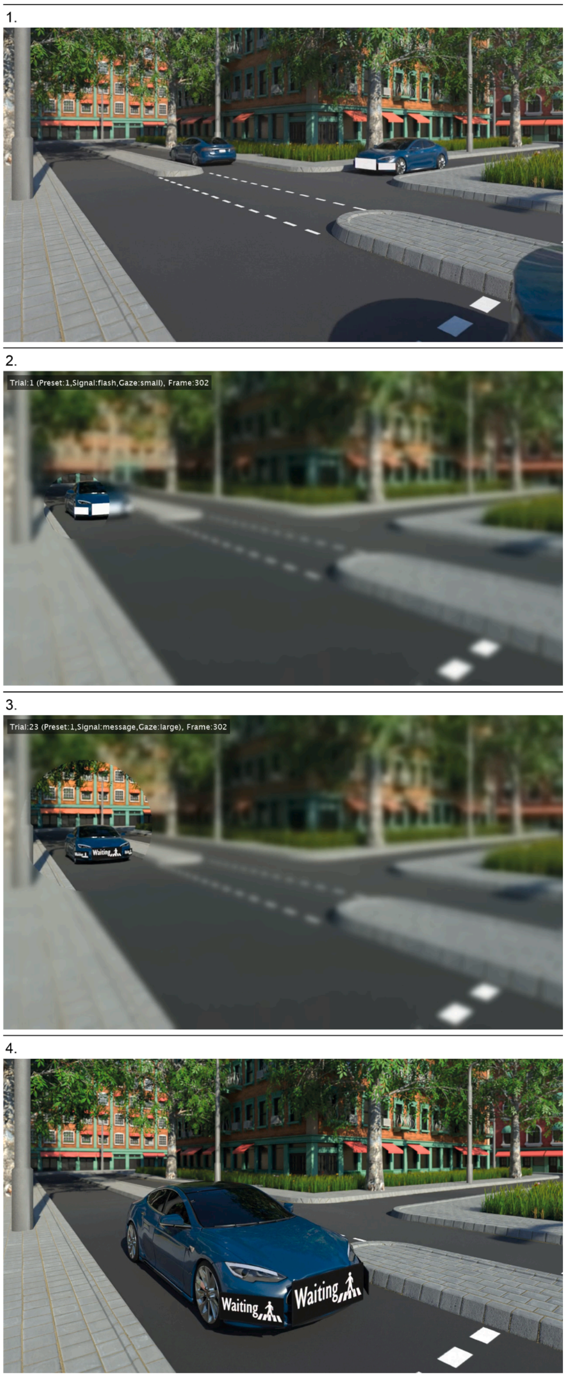


**Fig. 2. 1:** Moment the Flash eHMI turned on and the car started to decelerate ($t$ = 5.02 s; vehicle-path condition 3), **2:** Moment the Flash eHMI turned on and the car started to decelerate ($t$ = 5.02 s; vehicle-path condition 1), with a small gaze-contingent window, **3:** Moment when the Text eHMI turned on and the car started to decelerate ($t$ = 5.02 s; vehicle-path condition 1), with a large gaze-contingent window, **4:** End of the video clip with Text eHMI on ($t$ = 10.00 s; vehicle-path condition 1).

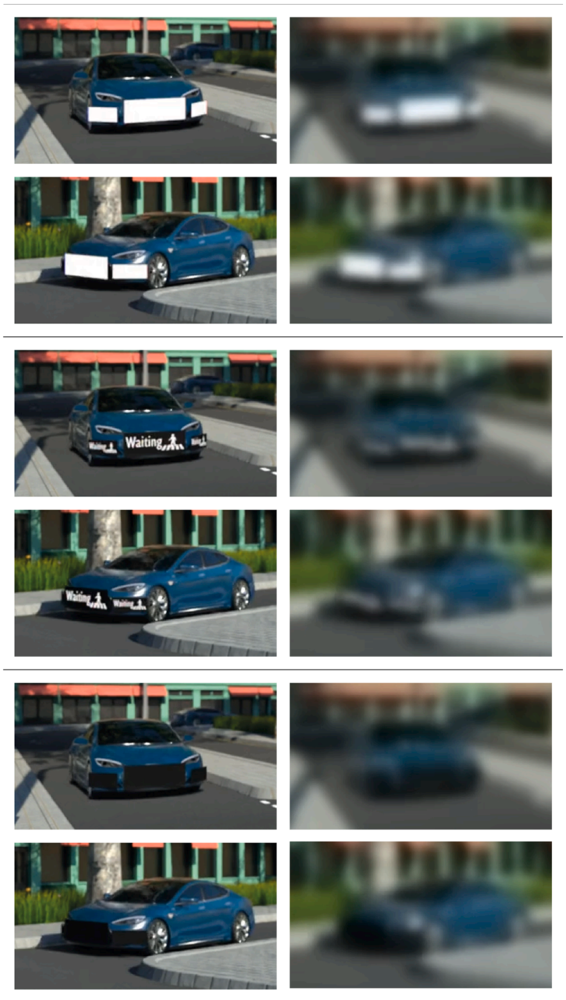


**Fig. 3.** Screenshots of the car at the moment of eHMI onset in clear view (left column) and in blurred view, i.e., when focusing elsewhere (right column). Top two rows: Flash eHMI, middle two rows: Text eHMI, bottom two rows: no eHMI.

condition. It can be seen that while the Text eHMI remains visible when presented in simulated peripheral vision, its salience is less pronounced than that of the Flash eHMI. This observation aligns with our intended design, as presented in the Introduction.

### *2.4. Design of the animated video clips*

Twenty-one non-interactive animated video clips were created: 3 eHMI conditions (Flash, Message, No eHMI) × 7 vehicle-path conditions. When combined with the 3 gaze-contingent overlays, this yielded 63 trials per participant. The video clips were created using the software Blender (The Blender Foundation, 2021). All videos were 10 s long and were played at 60 frames per second.

The road environment consisted of a T-junction and was the same in all video clips. It was derived from previous eHMI research (Eisma et al., 2020) and typical road topology of a real crossing in the Netherlands. The selected scenario was one where it is likely for a pedestrian to cross, while cars are not required to stop but could if they want to. The two-way street had a traffic island, so pedestrians only had to focus on traffic from one direction. The lane width was 3 m. The camera perspective was from the eyes of a pedestrian waiting to cross the road at a crossing with a traffic island.

There were a total of seven vehicle-path conditions. In all videos, four cars were present. The cars for each vehicle-path condition had one of three trajectories (Arriving head-on and going straight at the junction, Arriving from the right and turning left at the junction, and Driving on the other side of the road and away from the pedestrian). All cars drove at a speed of 30 km/h. In six of the seven videos, one of the cars had an eHMI.

At 5.02 s into the video, the eHMI activated, and the car started to decelerate at 3 $m/s^2$. At an elapsed time of 7.82 s, the car came to a full stop, 5 m before the pedestrian. In four vehicle-path conditions, the car with eHMI was approaching head-on, while in two vehicle-path conditions, the car with eHMI was making a left turn at the junction (see Fig. 2, for screenshots). In the seventh video, none of the cars stopped, and no eHMI was activated. The seventh video was included to induce a sense of unpredictability and was not analyzed in the current study.

### 2.5. Procedures

First, participants read and signed an informed consent form. Next, they sat behind the computer and adjusted the seat height so that they could comfortably position their head in a head support. Introductory information was presented on the computer screen: "*In the following half hour, you will be seated in front of this screen with an eye tracker. You will be shown a number of videos of a traffic situation. In this situation, you are a pedestrian who wants to cross the road half-way. A number of cars will cross your path, after which some may stop. All the cars in this experiment are equipped with an external display, which can be used to show signals to you, the pedestrian.*".

After that, the eHMIs were introduced, specifying what the eHMIs looked like and the information they conveyed: "*Here is an example of a traffic situation. The initial state of the external display on the front is always as shown" (black front of the car). When a car is stopping, this state may change to any of the examples shown*" (three screenshots were provided, no eHMI, Text eHMI, Flash eHMI). Participants were also provided with the following instruction: "*Whenever you deem it safe to cross the road half way, you press and hold the spacebar key on the keyboard.*".

Next, the eye tracker was calibrated with a standard 9-point calibration. After the calibration, participants completed three 10-s training trials: one trial without a gaze-contingent window and Flash eHMI, one trial with a small gaze-contingent window and Message eHMI, and one trial with a large gaze-contingent window without eHMI (in this order). Following the training trials, the actual trials began. The instruction to hold the spacebar was repeated on the screen: "*Remember to press and hold the spacebar key whenever you deem it safe to cross the road half-way*".

All trials were randomized, with a restriction that prevented subsequent trials from having the same gaze-contingent window size. The 63 trials were presented back-to-back, with no interaction between the participant and the experimenter. After each trial, there was a pause screen ("*Press any key to proceed*"), and a drift correction was performed. The experiment lasted between 30 and 45 min, including the calibration and practice trials.

After finishing all 63 trials, the participants were asked to complete a paper-and-pencil acceptance questionnaire (Van der Laan et al., 1997), one for each eHMI condition. In this questionnaire, respondents were required to evaluate nine bipolar adjectives using a five-point Likert scale. We computed the measure of usefulness by averaging the scores from the following five items: Item 1: useful–useless; Item 3: bad–good; Item 5: effective–superfluous; Item 7: assisting–worthless; and Item 9: raising alertness–sleep-inducing. Satisfaction was assessed based on the scores from these four items: Item 2: pleasant–unpleasant; Item 4: nice–annoying; Item 6: irritating–likeable; and Item 8: undesirable–desirable. We inverted the scores for Items 1, 2, 4, 5, 7, and 9 so that a higher score indicates greater usefulness/satisfaction.

### 2.6. Data processing and dependent variables

After the experiment, the horizontal ($x$) and vertical (y) gaze data as well as the pupil diameter data were filtered for further processing. Periods during which vertical gaze data on the screen were unavailable, such as caused by eye blinks, were labeled as gaps. A margin of 100 ms was added before and after each gap, and the gaps in the data were linearly interpolated. Next, the × and $y$ data, and pupil diameter, were median-filtered using a window length of 100 ms. Saccades were identified based on the speed of the gaze point and filtered using a Savitzky-Golay filter. Saccades were defined based on whether the speed of the gaze point exceeded 30°/s, and the minimum fixation duration was set to 40 ms (Nyström & Holmqvist, 2010).

The following dependent measures were used.

- Response time (ms): The time elapsed between the car starting to decelerate (which coincided with the moment the eHMI turned on) and the first press of the spacebar. Response times faster than 150 ms were not taken into consideration.
- Gaze dispersion (pixels): Next to response times, we explored whether eye movements differed between experimental conditions. A standard approach in the analysis of viewing behavior involves defining Areas of Interest (AOIs) and then calculating the duration or frequency that the participant's gaze falls within each AOI (Holmqvist et al., 2011). However, such an approach cannot be straightforwardly applied to our experimental results due to the dynamic nature of the traffic situation; cars in the video clip change in size and position within the visual field, and there exists a certain arbitrariness in how AOIs could be defined. To be able to analyze where participants focused their attention, the idea was originated to adopt a group-based approach (Eisma et al., 2020). In this approach, it is determined for each video frame whether participants were focusing more on a particular point or were distributing their attention. If participants' attention is centered around one point, it can be concluded that they were attracted to that point and likely used cues from that point at that moment to carry out their task. Conversely, if attention is more distributed, i.e., different participants were looking at different points in the scene, this could imply that there are multiple relevant cues or that participants were engaged in visual search, instead of a single dominant cue to be identified. For this purpose, we used an index called gaze dispersion, which is indicative of the extent to which participants were consistent in their eye behavior (Eisma et al., 2020). More specifically, we defined gaze dispersion as the mean distance from the participants' overall mean gaze point for that particular experimental condition. A dispersion score of, for example, 100 pixels, means that participants' gaze point was, on average, 100 pixels away from the mean fixation gaze point of all participants.
- Number of saccades: Saccades are rapid eye movements that cause the eyes to move from one location to another. The number of saccades was used as an index of eye movement activity (De Winter et al., 2021b). Note that the number of saccades corresponds to one less than the number of fixations, as the fixation filter divides the gaze data into fixations and saccades. In other words, the

number of fixations is redundant with the number of fixations, and since each trial had exactly the same duration (10 s), inversely related to the mean fixation duration.

These measures were computed per trial, and subsequently averaged across the six trials (i.e., vehicle-path conditions) for the corresponding eHMI × gaze-contingent condition. Thus, we obtained a total of nine scores per measure per participant (3 eHMI conditions × 3 gaze-contingent conditions). To simplify the analysis, the six vehicle-path conditions were averaged prior to statistical examination. However, it should be acknowledged that the mean response time for the two vehicle-path conditions in which the obstructing vehicle approached from the right was about 250 ms slower compared to the vehicle-path condition in which the obstructing vehicle approached head-on.

The statistical analyses involved conducting repeated-measures ANOVAs with eHMI condition and gaze-contingent condition as independent variables. In instances where no statistically significant eHMI × gaze-contingent window interaction was observed, paired *t*-tests were performed to compare the results for the three eHMI conditions (with the averages obtained for the three gaze-contingent conditions) and the three gaze-contingent conditions (with the averages obtained for the three eHMI conditions). A Bonferroni correction was applied, which meant that we set the significance level to 0.05/3.

In addition to the above measures, we analyzed pupil diameter as an index of mental workload. The use of pupil diameter as a measure of mental workload has been established in studies since the 1960 s (Hess & Polt, 1964; Kahneman & Beatty, 1966) and has been confirmed in numerous more recent studies (e.g., Klingner et al., 2011; Marquart & De Winter, 2015). Pupil diameter is also gaining popularity in transportation research (Radhakrishnan et al., 2023). Finally, we statistically analyzed the acceptance questionnaire, which comprised Usefulness (mean of 5 items) and Satisfaction (mean of 4 items) components scored from −2 to + 2. The three eHMI conditions were compared using paired-samples *t*-tests.

## 3. Results

### 3.1. Spacebar pressing and response times

Fig. 4 depicts the participants' spacebar pressing behavior as a function of elapsed time relative to the eHMI onset at $t = 0$ s. It can be seen that both eHMIs convinced participants that they could cross the road. The response times for the No-eHMI condition were slower, likely because participants had to rely on implicit cues (i.e., a reduction of speed) rather than an explicit visual signal.

Response times could be computed for a total of 1218 of 1242 (23 participants × 54 trials) trials. In the remaining 24 trials, the participant did not press the spacebar. There were no missing values in the statistical analysis, since, by averaging across the vehicle-path conditions, response times were available for all 23 participants and all nine combinations of eHMI and gaze-contingent window.

Table 1 shows the mean and standard deviations of response times for the nine combinations of eHMI and gaze-contingent conditions. According to a repeated-measures ANOVA of the mean response times, there was a significant effect of eHMI condition (Flash, Text, No eHMI), $F(2, 44) = 114.3$, $p < 0.001$, partial $\eta^2 = 0.84$, a significant effect of gaze-contingent condition (Small, Large, None), *F*

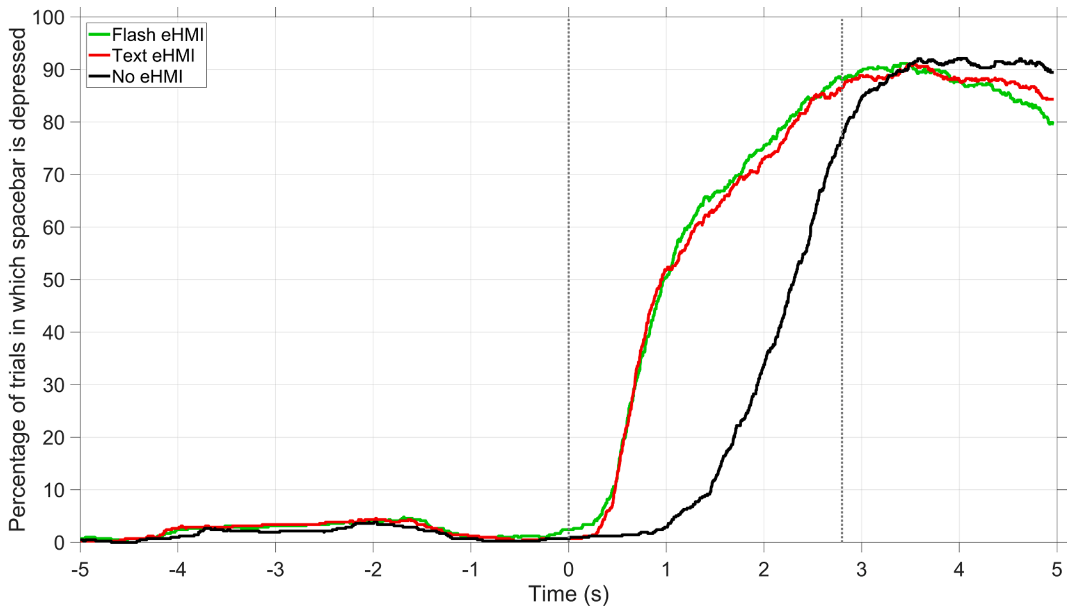


**Fig. 4.** Percentage of trials in which the spacebar was pressed as a function of elapsed time in the trial. The average was taken of all trials (23 participants × 6 vehicle paths × 3 gaze-contingent conditions) per eHMI condition. Only trials in which the participant pressed the spacebar at least once were included. The eHMI activated at 0 s, and the car came to a full stop at 2.8 s, as signified by the two vertical lines.

**Table 1**
Mean and standard deviation of the response times ($n = 23$).

| eHMI condition | Gaze-contingent condition | Mean (ms) | Standard deviation (ms) |
|---|---|---|---|
| Flash eHMI | Small window | 1294 | 630 |
| Flash eHMI | Large window | 1353 | 867 |
| Flash eHMI | No window | 1322 | 847 |
| Text eHMI | Small window | 1362 | 874 |
| Text eHMI | Large window | 1396 | 868 |
| Text eHMI | No window | 1267 | 749 |
| No eHMI | Small window | 2330 | 613 |
| No eHMI | Large window | 2404 | 666 |
| No eHMI | No window | 2198 | 619 |

*Note.* The results of six trials (i.e., vehicle-path conditions) of each participant were averaged.

(2, 44) = 4.06, $p = 0.024$, partial $\eta^2 = 0.16$, and no significant eHMI condition × gaze-contingent condition interaction, $F(4, 88) = 1.43$, $p = 0.232$, partial $\eta^2 = 0.06$.

Paired-samples *t*-tests revealed a significantly shorter response time for the Flash eHMI compared to No eHMI ($p < 0.001$), and for the Text eHMI compared to No eHMI ($p < 0.001$); the difference between Flash and Text was not statistically significant ($p = 0.730$) (three gaze-contingent conditions averaged). Paired-samples *t*-tests also revealed a shorter response time for the large window compared to no window ($p = 0.004$) (three eHMI conditions averaged), while the comparisons between the small window and no window and between the small and large window were not statistically significant ($p = 0.112$ and $p = 0.274$, respectively).

### 3.2. Eye-movement dispersion

Fig. 5 shows the gaze dispersion score for the three eHMI conditions without a gaze-contingent window. One thing that is apparent is that the approaching car (with or without eHMI) attracted attention, as signified by the low dispersion around $t = 0$ s. It can also be seen that the dispersion remained relatively low for the No-eHMI condition up to 2 s after the car started to decelerate. The observed outcome may be due to the participants' efforts to predict whether the car would halt or proceed. In the Flash eHMI condition, on the other hand, the eHMI provided only two brief flashes, which may explain why the participants' attention diversified soon after the eHMI onset, around $t = 1$ s.

Table 2 shows the means and standard deviations of the gaze dispersion times between the moment of eHMI onset and the moment the car came to a full stop, for each of the nine combinations of eHMI and gaze-contingent conditions. According to a repeated-measures ANOVA, there was a significant effect of eHMI condition, $F(2, 44) = 10.45$, $p < 0.001$, partial $\eta^2 = 0.32$, no significant effect of gaze-contingent condition, $F(2, 44) = 2.69$, $p = 0.079$, partial $\eta^2 = 0.11$, and a significant eHMI-condition × gaze-contingent condition interaction, $F(4, 88) = 5.35$, $p < 0.001$, partial $\eta^2 = 0.20$.

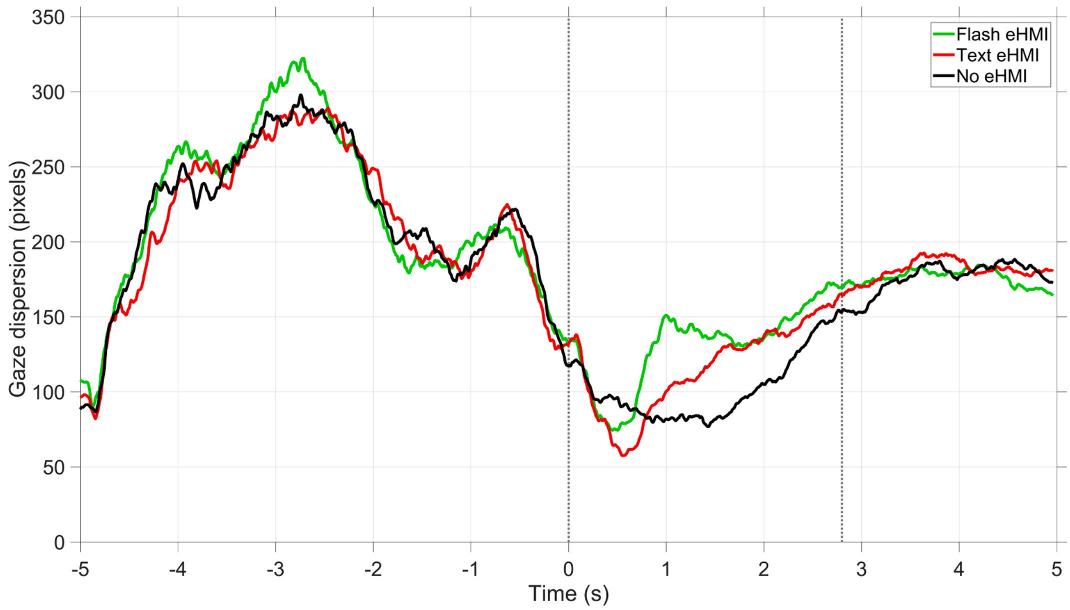


**Fig. 5.** Gaze dispersion for the three eHMI conditions without gaze-contingent window as a function of elapsed time in the trial. The average was taken of all 23 participants and 18 trials per participant (i.e., 6 vehicle-path conditions × 3 gaze-contingent conditions) per eHMI condition. The eHMI activated at 0 s, and the car came to a full stop at 2.8 s, as signified by vertical lines.

**Table 2**
Mean and standard deviation of gaze dispersion between the moment the car started to decelerate ($t = 5.02$ s) and the moment the car came to a full stop ($t = 7.80$ s) ($n = 23$).

| eHMI condition | Gaze-contingent condition | Mean (pixels) | Standard deviation (pixels) |
|---|---|---|---|
| Flash eHMI | Small window | 124 | 60 |
| Flash eHMI | Large window | 127 | 66 |
| Flash eHMI | No window | 143 | 64 |
| Text eHMI | Small window | 112 | 56 |
| Text eHMI | Large window | 115 | 60 |
| Text eHMI | No window | 122 | 53 |
| No eHMI | Small window | 106 | 43 |
| No eHMI | Large window | 105 | 35 |
| No eHMI | No window | 96 | 39 |

*Note.* The results of six trials (i.e., vehicle-path conditions) of each participant were averaged.

The significant interaction effect suggests that the high dispersion for the Flash eHMI was particularly prevalent without a gaze-contingent window (Table 2). Paired-samples *t*-tests for the no-window condition revealed a significantly higher dispersion for the Flash eHMI compared to No eHMI ($p < 0.001$) and for the Flash eHMI compared to the Text eHMI ($p = 0.002$), as well as for the Text eHMI compared to no eHMI ($p = 0.002$).

### 3.3. Number of saccades

In addition to gaze dispersion, it is useful to examine the overall eye-gaze activity. For this purpose, we tabulated the number of saccades during the trials (Table 3). It can be seen that the number of saccades was higher without a gaze-contingent window compared to a large and especially a small gaze-contingent window. According to a repeated-measures ANOVA, there was no significant effect of eHMI condition, $F(2, 44) = 0.91$, $p = 0.408$, partial $\eta^2 = 0.04$, a significant effect of gaze-contingent condition, $F(2, 44) = 52.31$, $p < 0.001$, partial $\eta^2 = 0.70$, and no significant eHMI-condition × gaze-contingent condition interaction, $F(4, 88) = 1.82$, $p = 0.132$, partial $\eta^2 = 0.08$.

Paired-samples *t*-tests revealed no significant differences in the number of saccades between the three eHMI types (Flash vs. Text, Flash vs. No eHMI, and Text vs. No eHMI; three gaze-contingent conditions averaged). However, paired-samples *t*-tests did reveal significant differences between the three gaze-contingent window conditions ($p < 0.001$ for all three comparisons) (three eHMI conditions averaged).

### 3.4. Pupil diameter

Fig. 6 shows the average pupil diameter under the three eHMI conditions. Initial results suggested a significant increase in pupil diameter in trials without eHMIs, possibly due to the uncertainty of the car stopping. The Flash eHMI condition saw a moderate increase, while the text-based eHMI condition resulted in the smallest pupil diameter. However, upon further analysis, we attributed these results partly to luminance differences within the stimulus. The Flash eHMI, which included two bright flashes, caused noticeable dips in pupil diameter at around 0.5 s and 0.9 s. After two flashes, the eHMI remained inactive, appearing identical to the No-eHMI condition. Conversely, the Text eHMI stayed visible till the video's end. Although the stimuli were small and displayed on a screen at a 1 m distance, which could potentially make brightness seem irrelevant, we employed a 'local darkness' metric (De Winter et al., 2021b) for quantitative evidence. This measures the average darkness of a 41 × 41-pixel area surrounding each participant's gaze coordinate in each trial. Fig. 7 depicts average local darkness levels for the three eHMI conditions: the flashes and the text-based eHMI's higher brightnesses were clearly distinguishable, possibly contributing to the smaller pupil diameter for the Text eHMI seen in Fig. 6.

**Table 3**
Mean and standard deviation of the number of saccades per trial ($n = 23$).

| eHMI condition | Gaze-contingent condition | Mean | Standard deviation |
|---|---|---|---|
| Flash eHMI | Small window | 12.46 | 2.89 |
| Flash eHMI | Large window | 14.80 | 3.87 |
| Flash eHMI | No window | 15.61 | 4.02 |
| Text eHMI | Small window | 12.51 | 3.49 |
| Text eHMI | Large window | 14.34 | 3.42 |
| Text eHMI | No window | 16.67 | 4.11 |
| No eHMI | Small window | 12.38 | 3.36 |
| No eHMI | Large window | 14.30 | 2.88 |
| No eHMI | No window | 15.62 | 4.24 |

*Note.* The results of six trials (i.e., vehicle-path conditions) of each participant were averaged.

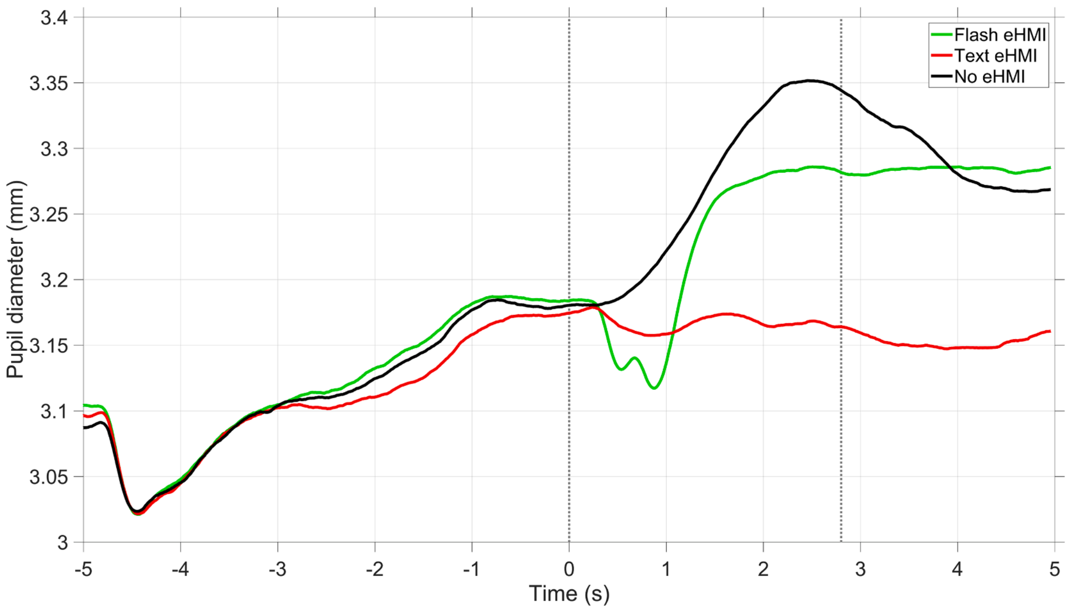


**Fig. 6.** Mean pupil diameter averaged across 18 videos and 23 participants as a function of elapsed time in the trial. The eHMI activated at 0 s, and the car came to a full stop at 2.8 s, as signified by the two vertical lines.

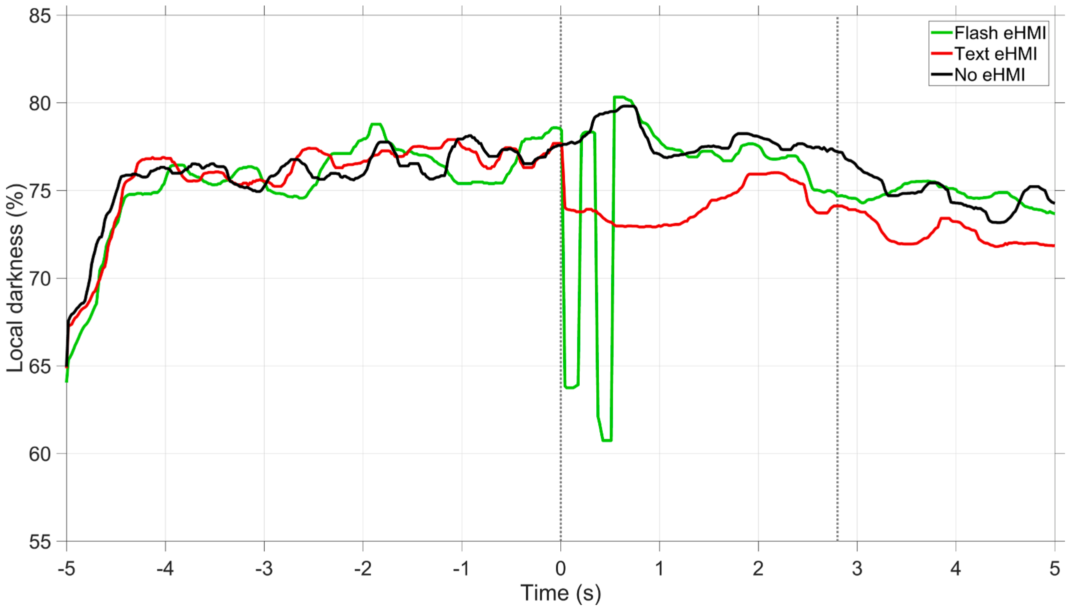


**Fig. 7.** Mean local darkness averaged across 18 videos and 23 participants as a function of elapsed time in the trial. The eHMI activated at 0 s, and the car came to a full stop at 2.8 s, as signified by the two vertical lines. The overall local darkness was median-filtered with a time constant of 200 ms.

### *3.5. Acceptance*

On a scale of −2 (minimum possible score) to + 2 (maximum possible score), the mean (*SD*) usefulness scores were 1.11 (0.77), 1.21 (0.57), and −0.87 (0.62), for the Flash eHMI, Text eHMI, and No-eHMI conditions, respectively. According to a paired-samples *t*-test, the difference between the Flash and Text eHMI was not statistically significant ($t(22) = -0.57, p = 0.577$), while the difference between the Flash eHMI and no eHMI was significant ($t(22) = 8.25, p < 0.001$), and the difference between the Text eHMI and no eHMI was as well ($t(22) = 11.36, p < 0.001$).

The mean (*SD*) satisfaction scores were 0.46 (1.05), 1.05 (0.71), and −0.40 (0.67), for the Flash eHMI, Text eHMI, and No-eHMI conditions, respectively. According to a paired-samples *t*-test, the difference between the Flash and Text eHMI was not statistically significant for our significance level ($t(22) = -2.22$, $p = 0.037$). On the other hand, significant differences were observed between the Flash eHMI and no eHMI ($t(22) = 2.94$, $p = 0.008$), and between the Text eHMI and no eHMI ($t(22) = 6.38$, $p < 0.001$).

## 4. Discussion

This study evaluated the effectiveness of a flash-based external human–machine interface (eHMI) compared to a text-based eHMI and a no eHMI condition in pedestrian-crossing scenarios. A gaze-contingent window was employed to simulate a real-life visual experience of traffic conditions. We hypothesized that the Text eHMI, compared to the Flash eHMI, would result in slower response times, particularly when peripheral vision was less accessible.

The results revealed that response times were significantly faster with the use of an eHMI as compared to the condition in which the eHMI was absent (for a similar finding, see, e.g., Bindschädel et al., 2021; De Clercq et al., 2019; Dietrich et al., 2019; Madigan et al., 2022). This study also found that the Flash eHMI and Text eHMI were perceived as significantly more useful and satisfying than the No-eHMI condition. Even though the effect of the gaze-contingent window on response times was relatively small, it was still statistically significant, with a gaze-contingent window causing slower response times compared to the full-vision condition.

However, contrary to our hypothesis, we did not find a statistically significant interaction between the eHMI and the gaze-contingent conditions. One explanation is that the Text eHMI in our study was still sufficiently easy to detect in the blurred region (for a depiction of the subtly conspicuous yet still discernible text, please refer to Fig. 3) and did not necessarily require reading the text, message was always the same and the car with eHMI would always come to a stop, therefore functioning more or less in an identical manner to the Flash eHMI. To better understand the demands of eHMIs on peripheral vision, future research should add conditions with eHMIs that show a message that the car will *not* stop, thus requiring the participant to explicitly read and understand the text message. In summary, our findings suggest that even with a blurred surrounding environment, participants perceived and understood the eHMIs well, indicating that peripheral vision may not significantly impact eHMI intelligibility under our experimental conditions.

The gaze dispersion results (see Fig. 5) show that homogeneous viewing behavior already occurred in a few tenths of a second before the eHMI was activated. This suggests that the eHMIs in our study did not heavily rely on peripheral vision, as participants were already focusing on the approaching car (for a similar finding, see Eisma et al., 2020). This observation aligns with prior research, which showed that road users tend to focus on objects or agents that have predictive value for their own safety while ignoring other, more irrelevant objects (Garrison & Williams, 2013; Lappi et al., 2017; Palazzi et al., 2019).

Additionally, the effect of eHMI and gaze-contingent windows on eye movements was analyzed. Notably, the gaze-dispersion analysis showed that the Flash eHMI captured participants' attention only briefly. This may be due to the brief flash signal, allowing the participants to press whenever they noticed the flashes. In contrast, the text-based eHMI, and especially the No-eHMI condition, held participants' attention longer. The heightened attention in the Text eHMI scenario might be attributed to the continuous text message in the video, possibly keeping the participants engaged and prompting them to read the text. Conversely, in the absence of eHMI, the increased focus could be a consequence of the necessity for extended vehicle observation to detect any signs of deceleration.

Furthermore, our investigation revealed substantial differences in saccade frequency among the three gaze-contingent conditions, with a smaller window resulting in fewer saccades than a larger window, which in turn yielded fewer saccades than the no-gaze-contingent condition. Previous studies have reported that the use of an aperture-based gaze-contingent window, as employed in our study, leads to longer fixation durations (and thus fewer saccades) compared to full vision (David et al., 2019; Loschky & McConkie, 2002; Nuthmann, 2014; Pomplun et al., 2001). A plausible explanation for this phenomenon is that peripheral vision is important in determining where the observer should direct their next fixation (Ludwig et al., 2014). Many eHMI studies have been conducted in relatively small field-of-view setups (e.g., Dey et al., 2020; Eisma et al., 2020, 2021; Faas et al., 2021; Guo et al., 2022a; Lau et al., 2022), while immersive virtual reality setups employing head-mounted displays or CAVE-based simulators are considerably rarer (for examples, see: Bindschädel et al., 2021; De Clercq et al., 2019; Kaleefathullah et al., 2022; Weber et al., 2019).

Initially, pupil diameter appeared to provide valuable insights into mental workload and the clarity of different eHMI conditions. We reasoned that the Flash eHMI yielded an elevated workload due to the ambiguous ways in which light-based eHMIs may be interpreted, such as signaling lights that could signify a gesture of right of way or indicate a threat (e.g., Lee et al., 2019). Furthermore, the low pupil diameter for the Text eHMI aligned with previous research suggesting that text-based eHMIs are perceived as clear and unambiguous (Bazilinskyy et al., 2019; Ferenchak & Shafique, 2022; Guo et al., 2022b). However, upon more detailed analysis, we found that the differences in pupil diameter could be explained, at least in part, by variations in light intensity caused by the eHMIs. This finding is consistent with recent research that showed changes in pupil diameter previously attributed to psychological phenomena, such as 'pupil mimicry', may actually be explained by changes in light intensity (De Winter et al., 2021b; Derksen et al., 2018). It is important to recognize these limitations, especially given the growing prevalence of eye-tracking technology and the increasing use of pupil diameter in transportation research, such as in Radhakrishnan et al. (2023). In fact, some researchers have suggested that pupil diameter should not be used in combination with visual stimuli at all (Goldwater, 1972; Janisse, 1974). We hope that this comment serves as a useful resource for researchers and practitioners in avoiding potential confounding factors when interpreting changes in pupil diameter.

A limitation of our study was that it was conducted in a safe lab environment. The impacts of eHMIs on the safety of pedestrians in real-life scenarios, and the potential for eHMIs to reduce the number of road traffic accidents is still an understudied topic that requires more research (see Cefkin et al., 2019; Forke et al., 2021; Merat et al., 2018; Monzel et al., 2021; for studies on the effect of eHMIs in naturalistic environments). It must also be acknowledged that the installation of eHMIs is associated with additional costs, a factor that cannot be overlooked in the broader scope of this research. The ethics surrounding the use of eHMIs in pedestrian-crossing scenarios, including issues related to trust and overreliance, also require further elaboration. Though effective at capturing attention and reducing response times, the eHMIs used in this study could promote overreliance if pedestrians come to expect vehicles will always stop. A study by Kaleefathullah et al. (2022) found that pedestrians who had become reliant on a flash-based eHMI were likely to have accidents when the eHMI turned on, yet the AV unexpectedly continued driving. Therefore, engineers developing eHMIs for pedestrians should create systems that promote pedestrian situation awareness rather than systems that elicit stimulus–response behaviors.

## 5. Conclusions

This study evaluated the effectiveness of different types of eHMIs on pedestrian crossing behavior, employing differently-sized gaze-contingent windows. The findings indicate that response times were substantially reduced with the use of an eHMI as compared to the conditions with no eHMI. The flash-based eHMI and text-based eHMI yielded equivalent response times, but there were subtle differences, where the flash-based eHMI required less attention allocation time (thus allowing participants to focus elsewhere). Additionally, this study suggests that the gaze-contingent window, which had the goal of simulating real-life conditions, had a negative effect on response times and on the number of saccades, implying that the participants allocated their attention less. In conclusion, although participants could perform the task effectively even when much of the scene was blurred, researchers and designers of eHMIs are advised to consider the role of peripheral vision in designing eHMIs that aid pedestrians and other vulnerable road users in making the right decisions.

## CRediT authorship contribution statement

**Yke Bauke Eisma:** Conceptualization, Methodology, Writing – review & editing, Supervision, Software, Data curation, Project administration, Validation. **Lucas van Gent:** Conceptualization, Methodology, Resources, Formal analysis, Visualization, Investigation, Writing – original draft, Project administration. **Joost de Winter:** Conceptualization, Methodology, Validation, Formal analysis, Data curation, Visualization, Validation, Supervision, Project administration, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### *Acknowledgments*

This research is funded by grant 016.Vidi.178.047 (“How should automated vehicles communicate with other road users?”), which is provided by the Netherlands Organization for Scientific Research (NWO). More details, such as the rendering settings of the videos, are provided in Van Gent (2022).

### *Data availability*

MATLAB scripts and raw data are available at https://doi.org/10.4121/6dfe7bd5-16f7-4e8d-9ab7-f8203b902310.

## References

Atchley, P., & Dressel, J. (2004). Conversation limits the functional field of view. *Human Factors, 46*, 664–673. https://doi.org/10.1518/hfes.46.4.664.56808

Basch, C. H., Ethan, D., Rajan, S., & Basch, C. E. (2014). Technology-related distracted walking behaviours in Manhattan's most dangerous intersections. *Injury Prevention, 20*, 343–346. https://doi.org/10.1136/injuryprev-2013-041063

Bazilinskyy, P., Dodou, D., & De Winter, J. (2019). Survey on eHMI concepts: The effect of text, color, and perspective. *Transportation Research Part F: Traffic Psychology and Behaviour, 67*, 175–194. https://doi.org/10.1016/j.trf.2019.10.013

Bazilinskyy, P., Dodou, D., & De Winter, J. C. F. (2022). Crowdsourced assessment of 227 text-based eHMIs for a crossing scenario. *Advances in Transportation, AHFE International Conference, 60*, 141–155. https://doi.org/10.54941/ahfe1002444

Bazilinskyy, P., Kooijman, L., Dodou, D., & De Winter, J. C. F. (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*, Article 103450. https://doi.org/10.1016/j.apergo.2021.103450

Bazilinskyy, P., Kooijman, L., Mallant, K. P. T., Roosens, V. E. R., Middelweerd, M. D. L. M., Overbeek, L. D., Dodou, D., & De Winter, J. C. F. (2022b). Get out of the way! Examining eHMIs in critical driver-pedestrian encounters in a coupled simulator. *Proceedings of the 14th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Seoul, Republic of Korea. https://doi.org/10.1145/3543174.3546849.

Berge, S. H., Hagenzieker, M., Farah, H., & De Winter, J. (2022). Do cyclists need HMIs in future automated traffic? An interview study. *Transportation Research Part F: Traffic Psychology and Behaviour, 84*, 33–52. https://doi.org/10.1016/j.trf.2021.11.013

Biebl, B., Arcidiacono, E., Kacianka, S., Rieger, J. W., & Bengler, K. (2022). Opportunities and limitations of a gaze-contingent display to simulate visual field loss in driving simulator studies. *Frontiers in Neuroergonomics, 3*, Article 916169. https://doi.org/10.3389/fnrgo.2022.916169

Bindschädel, J., Krems, I., & Kiesel, A. (2021). Interaction between pedestrians and automated vehicles: Exploring a motion-based approach for virtual reality experiments. *Transportation Research Part F: Traffic Psychology and Behaviour, 82*, 316–332. https://doi.org/10.1016/j.trf.2021.08.018

Bursill, A. E. (1958). The restriction of peripheral vision during exposure to hot and humid conditions. *The Quarterly Journal of Experimental Psychology, 10*, 114–129. https://doi.org/10.1080/17470215808416265

Cefkin, M. (2018). Towards socially acceptable autonomous driving [Presentation]. Nissan Research Center Silicon Valley.

Cefkin, M., Zhang, J., Stayton, E., & Vinkhuyzen, E. (2019). Multi-methods research to examine external HMI for highly automated vehicles. In H. Krömker (Ed.), *HCI in Mobility, Transport, and Automotive Systems* (pp. 46–64). Cham: Springer. https://doi.org/10.1007/978-3-030-22666-4_4.

Chen, Z., Xu, J., Liu, Z., Wei, Z., & Gong, M. (2023). Participatory design of external HMI based on explicit human-vehicle interaction features. In L. C. Jain, V. E. Balas, Q. Wu, & F. Shi (Eds.), *Design studies and intelligence engineering* (pp. 74–83). IOS Press. https://doi.org/10.3233/faia220702.

Combs, T. S., Sandt, L. S., Clamann, M. P., & McDonald, N. C. (2019). Automated vehicles and pedestrian safety: Exploring the promise and limits of pedestrian detection. *American Journal of Preventive Medicine, 56*, 1–7. https://doi.org/10.1016/j.amepre.2018.06.024

Conti, A. S., Dlugosch, C., … Bengler, K. (2014). In *The effect of task set instruction on detection response task performance* (pp. 107–117). Santa Monica, CA: Human Factors and Ergonomics Society.

David, E. J., Lebranchu, P., Perreira Da Silva, M., & Le Callet, P. (2019). Predicting artificial visual field losses: A gaze-based inference study. *Journal of Vision, 19*, 22. https://doi.org/10.1167/19.14.22

De Clercq, K., Dietrich, A., Núñez Velasco, J. P., De Winter, J., & Happee, R. (2019). External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. *Human Factors, 61*, 1353–1370. https://doi.org/10.1177/0018720819836343

De Winter, J., Bazilinskyy, P., Wesdorp, D., De Vlam, V., Hopmans, B., Visscher, J., & Dodou, D. (2021a). How do pedestrians distribute their visual attention when walking through a parking garage? An eye-tracking study. *Ergonomics, 64*, 793–805. https://doi.org/10.1080/00140139.2020.1862310

De Winter, J. C. F., Petermeijer, S. M., Kooijman, L., & Dodou, D. (2021b). Replicating five pupillometry studies of Eckhard Hess. *International Journal of Psychophysiology, 165*, 145–205. https://doi.org/10.1016/j.ijpsycho.2021.03.003

Derksen, M., Van Alphen, J., Schaap, S., Mathot, S., & Naber, M. (2018). Pupil mimicry is the result of brightness perception of the iris and pupil. *Journal of Cognition, 1*, 32. https://doi.org/10.5334/joc.34

Dey, D., Ackermans, S., Martens, M., Pfleging, B., & Terken, J. (2022). Interactions of automated vehicles with road users. In A. Riener, M. Jeon, & I. Alvarez (Eds.), *User experience design in the era of automated driving* (pp. 533–581). Cham: Springer. https://doi.org/10.1007/978-3-030-77726-5_20.

Dey, D., Habibovic, A., Pfleging, B., Martens, M., & Terken, J. (2020). Color and animation preferences for a light band eHMI in interactions between automated vehicles and pedestrians. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu HI. https://doi.org/10.1145/3313831.3376325.

Dey, D., Matviienko, A., Berger, M., Pfleging, B., Martens, M., & Terken, J. (2021). Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior. *it-Information Technology, 63*, 123–141. https://doi.org/10.1515/itit-2020-0025

Dietrich, A., Tondera, M., & Bengler, K. (2019). *Automated vehicles in urban traffic: The effects of kinematics and eHMI on pedestrian crossing behaviour*. University of Iowa.

Duchowski, A. T., Cournia, N., & Murphy, H. (2004). Gaze-contingent displays: A review. *Cyberpsychology & Behavior, 7*, 621–634. https://doi.org/10.1089/cpb.2004.7.621

Edwards, J. D., Ross, L. A., Wadley, V. G., Clay, O. J., Crowe, M., Roenker, D. L., & Ball, K. K. (2006). The useful field of view test: Normative data for older adults. *Archives of Clinical Neuropsychology, 21*, 275–286. https://doi.org/10.1016/j.acn.2006.03.001

Eisele, D., & Petzoldt, T. (2022). Effects of traffic context on eHMI icon comprehension. *Transportation Research Part F: Traffic Psychology and Behaviour, 85*, 1–12. https://doi.org/10.1016/j.trf.2021.12.014

Eisma, Y. B., Reiff, A., Kooijman, L., Dodou, D., & De Winter, J. C. F. (2021). External human-machine interfaces: Effects of message perspective. *Transportation Research Part F: Traffic Psychology and Behaviour, 78*, 30–41. https://doi.org/10.1016/j.trf.2021.01.013

Eisma, Y. B., Van Bergen, S., Ter Brake, S. M., Hensen, M. T. T., Tempelaar, W. J., & De Winter, J. C. F. (2020). External human–machine interfaces: The effect of display location on crossing intentions and eye movements. *Information, 11*, 13. https://doi.org/10.3390/info11010013

Epke, M. R., Kooijman, L., & De Winter, J. C. F. (2021). I see your gesture: A VR-based study of bidirectional communication between pedestrians and automated vehicles. *Journal of Advanced Transportation, 2021*, 5573560. https://doi.org/10.1155/2021/5573560

European Union. (2022). Commission Implementing Regulation (EU) 2022/1426 of 5 August 2022 laying down rules for the application of Regulation (EU) 2019/2144 of the European Parliament and of the Council as regards uniform procedures and technical specifications for the type-approval of the automated driving system (ADS) of fully automated vehicles. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32022R1426.

Faas, S. M., Kraus, J., Schoenhals, A., & Baumann, M. (2021). Calibrating pedestrians' trust in automated vehicles: Does an intent display in an external HMI support trust calibration and safe crossing behavior?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411764.3445738

Feld, J. A., & Plummer, P. (2019). Visual scanning behavior during distracted walking in healthy young adults. *Gait & Posture, 67*, 219–223. https://doi.org/10.1016/j.gaitpost.2018.10.017

Ferenchak, N. N., & Shafique, S. (2022). Pedestrians' perceptions of autonomous vehicle external human-machine interfaces. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part B: Mechanical Engineering, 8*, Article 034501. https://doi.org/10.1115/1.4051778

Finlay, D. (1982). Motion perception in the peripheral visual field. *Perception, 11*, 457–462. https://doi.org/10.1068/p110457

Forke, J., Fröhlich, P., Suette, S., Gafert, M., Puthenkalam, J., Diamond, L., … Tscheligi, M. (2021). Understanding the headless rider: Display-based awareness and intent-communication in automated vehicle-pedestrian interaction in mixed traffic. *Multimodal Technologies and Interaction, 5*, 51. https://doi.org/10.3390/mti5090051

Franchak, J. M., & Adolph, K. E. (2010). Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. *Vision Research, 50*, 2766–2774. https://doi.org/10.1016/j.visres.2010.09.024

Garrison, T. M., & Williams, C. C. (2013). Impact of relevance and distraction on driving performance and visual attention in a simulated driving environment. *Applied Cognitive Psychology, 27*, 396–405. https://doi.org/10.1002/acp.2917

Geruschat, D. R., Hassan, S. E., & Turano, K. A. (2003). Gaze behavior while crossing complex intersections. *Optometry and Vision Science, 80*, 515–528. https://doi.org/10.1097/00006324-200307000-00013

Glen, F. C., Smith, N. D., Jones, L., & Crabb, D. P. (2016). 'I didn't see that coming': Simulated visual fields and driving hazard perception test performance. *Clinical and Experimental Optometry, 99*, 469–475. https://doi.org/10.1111/cxo.12435

Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin, 77*, 340–355. https://doi.org/10.1037/h0032456

GOV.UK. (2022). Responsible innovation in self-driving vehicles. https://www.gov.uk/government/publications/responsible-innovation-in-self-driving-vehicles/responsible-innovation-in-self-driving-vehicles.

Granit, R., & Harper, P. (1930). Comparative studies on the peripheral and central retina: II. Synaptic reactions in the eye. *American Journal of Physiology-Legacy Content, 95*, 211–228. https://doi.org/10.1152/ajplegacy.1930.95.1.211

Gremmelmaier, H., Töttel, L., Pavlitskaya, S., Orf, S., Zofka, M., & Zöllner, J. M. (2022). Getting ready for all types of users: A virtual reality training center for automated vehicles. In *Proceedings of the 2022 IEEE Symposium Series on Computational Intelligence*. https://doi.org/10.1109/SSCI51031.2022.10022144

Guo, F., Lyu, W., Ren, Z., Li, M., & Liu, Z. (2022). A video-based, eye-tracking study to investigate the effect of eHMI modalities and locations on pedestrian–automated vehicle interaction. *Sustainability, 14*, 5633. https://doi.org/10.3390/su14095633

Guo, J., Yuan, Q., Yu, J., Chen, X., Yu, W., Cheng, Q., … Jiang, X. (2022). External human–machine interfaces for autonomous vehicles from pedestrians' perspective: A survey study. *Sensors, 22*, 3339. https://doi.org/10.3390/s22093339

Hassan, S. E., Geruschat, D. R., & Turano, K. A. (2005). Head movements while crossing streets: Effect of vision impairment. *Optometry and Vision Science, 82*, 18–26.

Hensch, A.-C., Neumann, I., Beggiato, M., Halama, J., & Krems, J. F. (2019). Effects of a light-based communication approach as an external HMI for Automated Vehicles — a Wizard-of-Oz study. *Transactions on Transport Sciences, 10*, 18–32. https://doi.org/10.5507/tots.2019.012

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143*, 1190–1192. https://doi.org/10.1126/science.143.3611.1190

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Areas of interest. In K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst,H. Jarodzka, & J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures* (pp. 187–230). OUP Oxford.

Hudson, C. R., Deb, S., Carruth, D. W., McGinley, J., & Frey, D. (2019). Pedestrian perception of autonomous vehicles with external interacting features. In I. Nunes (Ed.), *Advances in Human Factors and Systems Interaction. AHFE 2018* (pp. 33–39). Cham: Springer. https://doi.org/10.1007/978-3-319-94334-3_5.

Huestegge, L., & Böckler, A. (2016). Out of the corner of the driver's eye: Peripheral processing of hazards in static traffic scenes. *Journal of Vision, 16*, 11. https://doi.org/10.1167/16.2.11

ISO. (2018). Road vehicles: Ergonomic aspects of external visual communication from automated vehicles to other road users (ISO/TR 23049:2018).

Jacobs, A. M., Nazir, T. A., & Heller, O. (1989). Perception of lowercase letters in peripheral vision: A discrimination matrix based on saccade latencies. *Perception & Psychophysics, 46*, 95–102. https://doi.org/10.3758/BF03208079

Janisse, M. P. (1974). Pupillometry: Some advances, problems and solutions. In M. P. Janisse (Ed.), *Pupillary dynamics and behavior* (pp. 209–248). Boston, MA: Springer. https://doi.org/10.1007/978-1-4757-1642-9_1.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*, 1583–1585. https://doi.org/10.1126/science.154.3756.1583

Kaleefathullah, A. A., Merat, N., Lee, Y. M., Eisma, Y. B., Madigan, R., Garcia, J., & De Winter, J. (2022). External human–machine interfaces can be misleading: An examination of trust development and misuse in a CAVE-based pedestrian simulation environment. *Human Factors, 64*, 1070–1085. https://doi.org/10.1177/0018720820970751

Klingner, J., Tversky, B., & Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology, 48*, 323–332. https://doi.org/10.1111/j.1469-8986.2010.01069.x

Kunst, K., Scheuchenpflug, J., Kraft, J., & Flachhuber, M. (2022). Investigating the perception of pedestrians in Car 2 Human communication: A case study using different symbols and dynamics to communicate via an angular restricted eHMI and road projections. *SAE Technical Paper*, 2022-01-0800. https://doi.org/10.4271/2022-01-0800.

Lappi, O., Rinkkala, P., & Pekkanen, J. (2017). Systematic observation of an expert driver's gaze strategy—an on-road case study. *Frontiers in Psychology, 8*, 620. https://doi.org/10.3389/fpsyg.2017.00620

Larue, G., Watling, C., Black, A., Wood, J., & Khakzar, M. (2019). Pedestrian distraction at railway level crossings: Can illuminated in-ground LEDs attract their attention back? *Proceedings of the 12th World Congress on Railway Research (WCRR) 2019.* Rail Safety and Standards Board Ltd, UK.

Lau, M., Jipp, M., & Oehl, M. (2022). One solution fits all? Evaluating different communication strategies of a light-based external human-machine interface for differently sized automated vehicles from a pedestrian's perspective. *Accident Analysis & Prevention, 171*, Article 106641. https://doi.org/10.1016/j.aap.2022.106641

Lee, Y. M., Madigan, R., Garcia, J., Tomlinson, A., Solernou, A., Romano, R., Markkula, G., Merat, N., & Uttley, J. (2019). Understanding the messages conveyed by automated vehicles. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 134–143, Utrecht, The Netherlands. https://doi.org/10.1145/3342197.3344546.

Lee, Y. M., Madigan, R., Giles, O., Garach-Morcillo, L., Markkula, G., Fox, C., … Merat, N. (2021). Road users rarely use explicit communication when interacting in today's traffic: Implications for automated vehicles. *Cognition, Technology & Work, 23*, 367–380. https://doi.org/10.1007/s10111-020-00635-y

Lévêque, L., Ranchet, M., Deniel, J., Bornard, J.-C., & Bellet, T. (2020). Where do pedestrians look when crossing? A state of the art of the eye-tracking studies. *IEEE Access, 8*, 164833–164843. https://doi.org/10.1109/ACCESS.2020.3021208

Löcken, A., Golling, C., & Riener, A. (2019). How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 262–274, Utrecht, The Netherlands. https://doi.org/10.1145/3342197.3344544.

Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal of Experimental Psychology: Applied, 8*, 99–117. https://doi.org/10.1037/1076-898X.8.2.99

Ludwig, C. J. H., Davies, J. R., & Eckstein, M. P. (2014). Foveal analysis and peripheral selection during active visual sampling. *Proceedings of the National Academy of Sciences, 111*, E291–E299. https://doi.org/10.1073/pnas.1313553111

Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic Science, 3*, 67–68. https://doi.org/10.3758/BF03343023

Madigan, R., Lee, Y. M., Lyu, W., Horn, S., Garcia de Pedro, J., & Merat, N. (2022). *Pedestrian interactions with Automated Vehicles: Does the presence of a zebra crossing affect how eHMIs and movement patterns are interpreted?* PsyArXiv. https://doi.org/10.31234/osf.io/7xjvf.

Mahadevan, K., Somanath, S., & Sharlin, E. (2018). Communicating awareness and intent in autonomous vehicle-pedestrian interaction. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montréal, Canada. https://doi.org/10.1145/3173574.3174003.

Mao, C., Go, K., Kinoshita, Y., Kashiwagi, K., Toyoura, M., Fujishiro, I., … Mao, X. (2021). Different eye movement behaviors related to artificial visual field defects–A pilot study of video-based perimetry. *IEEE Access, 9*, 77649–77660. https://doi.org/10.1109/ACCESS.2021.3080687

Marigold, D. S. (2008). Role of peripheral visual cues in online visual guidance of locomotion. *Exercise and Sport Sciences Reviews, 36*, 145–151. https://doi.org/10.1097/JES.0b013e31817bff72

Marquart, G., & De Winter, J. (2015). Workload assessment for mental arithmetic tasks using the task-evoked pupillary response. *PeerJ Computer Science, 1*, e16.

Merat, N., Louw, T., Madigan, R., Wilbrink, M., & Schieben, A. (2018). What externally presented information do VRUs require when interacting with fully Automated Road Transport Systems in shared space? *Accident Analysis & Prevention, 118*, 244–252. https://doi.org/10.1016/j.aap.2018.03.018

Mok, C. S., Bazilinskyy, P., & De Winter, J. C. F. (2022). Stopping by looking: A driver-pedestrian interaction study in a coupled simulator using head-mounted displays with eye-tracking. *Applied Ergonomics, 105*, Article 103825. https://doi.org/10.1016/j.apergo.2022.103825

Monzel, M., Keidel, K., Schubert, W., & Banse, R. (2021). A field study investigating road safety effects of a front brake light. *IET Intelligent Transport Systems, 15*, 1043–1052. https://doi.org/10.1049/itr2.12080

Moore, D., Currano, R., Strack, G. E., & Sirkin, D. (2019). The case for implicit external human-machine interfaces for autonomous vehicles. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 295–307, Utrecht, Netherlands. https://doi.org/10.1145/3342197.3345320.

NACTO. (2016). Policy statement on automated vehicles. https://nacto.org/wp-content/uploads/2016/06/NACTO-Policy-Automated-Vehicles-201606.pdf.

NHTSA. (2016). Federal automated vehicles policy. https://www.nhtsa.gov/document/automated-driving-systems-20-voluntary-guidance.

Nuthmann, A. (2014). How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 342–360. https://doi.org/10.1037/a0033854

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods, 42*, 188–204. https://doi.org/10.3758/BRM.42.1.188

Palazzi, A., Abati, D., Calderara, S., Solera, F., & Cucchiara, R. (2019). Predicting the driver's focus of attention: The DR(eye)VE project. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 41*, 1720–1733. https://doi.org/10.1109/TPAMI.2018.2845370

Park, M.-K., Lee, S.-Y., Kwon, C.-K., & Kim, S.-W. (2017). Design of pedestrian target selection with funnel map for pedestrian AEB system. *IEEE Transactions on Vehicular Technology, 66*, 3597–3609. https://doi.org/10.1109/TVT.2016.2604420

Pomplun, M., Reingold, E. M., & Shen, J. (2001). Investigating the visual span in comparative search: The effects of task difficulty and divided attention. *Cognition, 81*, B57–B67. https://doi.org/10.1016/S0010-0277(01)00123-8

Radhakrishnan, V., Louw, T., Cirino Gonçalves, R., Torrao, G., Lenné, M. G., & Merat, N. (2023). Using pupillometry and gaze-based metrics for understanding drivers' mental workload during automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 94*, 254–267. https://doi.org/10.1016/j.trf.2023.02.015

Rantanen, E. M., & Goldberg, J. H. (1999). The effect of mental workload on the visual field size and shape. *Ergonomics, 42*, 816–834. https://doi.org/10.1080/001401399185315

Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multiresolutional displays: An integrative review. *Human Factors, 45*, 307–328. https://doi.org/10.1518/hfes.45.2.307.27235

Ringer, R. V., Throneburg, Z., Johnson, A. P., Kramer, A. F., & Loschky, L. C. (2016). Impairing the useful field of view in natural scenes: Tunnel vision versus general interference. *Journal of Vision, 16*, 7. https://doi.org/10.1167/16.2.7

Rothenbücher, D., Li, J., Sirkin, D., Mok, B., & Ju, W. (2016). Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication*, 795–802, New York, NY. https://doi.org/10.1109/ROMAN.2016.7745210.

Sakurai, M. (2015). Parafovea. In R. Luo (Ed.), *Encyclopedia of color science and technology*. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-27851-8_215-1.

Sanders, A. F. (1970). Some aspects of the selective process in the functional visual field. *Ergonomics, 13*, 101–117. https://doi.org/10.1080/00140137008931124

Schmidt-Wolf, M., & Feil-Seifer, D. (2022). Vehicle-to-pedestrian communication feedback module: A study on increasing legibility, public acceptance and trust. *Proceedings of Social Robotics: 14th International Conference*, 14–23, Florence, Italy. https://doi.org/10.1007/978-3-031-24667-8_2.

Sekuler, A. B., Bennett, P. J., & Mamelak, M. (2000). Effects of aging on the useful field of view. *Experimental Aging Research, 26*, 103–120. https://doi.org/10.1080/036107300243588

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*, 13. https://doi.org/10.1167/11.5.13

SWOV. (2022). Zelfrijdende voertuigen [Self-driving vehicles] (Factsheet). https://swov.nl/sites/default/files/bestanden/downloads/FS%20Zelfrijdende%20voertuigen_0.pdf.

Tabone, W., De Winter, J. C. F., Ackermann, C., Bärgman, J., Baumann, M., Deb, S., … Stanton, N. A. (2021). Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. *Transportation Research Interdisciplinary Perspectives, 9*, Article 100293. https://doi.org/10.1016/j.trip.2020.100293

The Blender Foundation. (2021). Blender. https://www.blender.org.

The FFmpeg Developers. (2021). FFmpeg. https://ffmpeg.org.

Turano, K. A., Rubin, G. S., & Quigley, H. A. (1999). Mobility performance in glaucoma. *Investigative Ophthalmology & Visual Science, 40*, 2803–2809.

US DOT. (2018). Preparing for the future of transportation. Automated Vehicles 3.0. https://www.transportation.gov/sites/dot.gov/files/docs/policy-initiatives/automated-vehicles/320711/preparing-future-transportation-automated-vehicle-30.pdf.

UNECE. (2019). GRE taskforce on autonomous vehicle signalling requirements, 4th meeting (Draft Report No. AVSR-04-05). https://wiki.unece.org/download/attachments/78742442/AVSR-04-05e.docx?api=v2.

Uttley, J., Fotios, S., & Cheal, C. (2017). Effect of illuminance and spectrum on peripheral obstacle detection by pedestrians. *Lighting Research & Technology, 49*, 211–227. https://doi.org/10.1177/1477153515602954

Van der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies, 5*, 1–10. https://doi.org/10.1016/S0968-090X(96)00025-3

Van Gent, L. (2022). *The effect of cognitive load on the effect of external Human-Machine Interfaces: An eye-tracker experiment*. Delft University of Technology. MSc thesis).

Van Winsum, W. (2019a). Optic flow and tunnel vision in the detection response task. *Human Factors, 61*, 992–1003. https://doi.org/10.1177/0018720818825008

Van Winsum, W. (2019b). A threshold model for stimulus detection in the peripheral detection task. *Transportation Research Part F: Traffic Psychology and Behaviour, 65*, 485–502. https://doi.org/10.1016/j.trf.2019.08.014

Vater, C., Kredel, R., & Hossner, E.-J. (2016). Detecting single-target changes in multiple object tracking: The case of peripheral vision. *Attention, Perception, & Psychophysics, 78*, 1004–1019. https://doi.org/10.3758/s13414-016-1078-7

Vater, C., Wolfe, B., & Rosenholtz, R. (2022). Peripheral vision in real-world tasks: A systematic review. *Psychonomic Bulletin & Review, 29*, 1531–1557. https://doi.org/10.3758/s13423-022-02117-w

Verstegen, R., Dey, D., & Pfleging, B. (2021). CommDisk: A holistic 360° eHMI concept to facilitate scalable, unambiguous interactions between automated vehicles and other road users. *Proceedings of the 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 132–136, Leeds, UK. https://doi.org/10.1145/3473682.3480280.

Weber, F., Chadowitz, R., Schmidt, K., Messerschmidt, J., & Fuest, T. (2019). Crossing the street across the globe: A study on the effects of eHMI on pedestrians in the US, Germany and China. In H. Krömker (Ed.), *HCI in Mobility, Transport, and Automotive Systems. HCII 2019* (pp. 515–530). Cham: Springer. https://doi.org/10.1007/978-3-030-22666-4_37.

Weltman, G., Smith, J. E., & Egstrom, G. H. (1971). Perceptual narrowing during simulated pressure-chamber exposure. *Human Factors, 13*, 99–107. https://doi.org/10.1177/001872087101300202

Wood, J., Chaparro, A., Hickson, L., Thyer, N., Carter, P., Hancock, J., … Ybarzabal, F. (2006). The effect of auditory and visual distracters on the useful field of view: Implications for the driving task. *Investigative Ophthalmology & Visual Science, 47*, 4646–4650. https://doi.org/10.1167/iovs.06-0306