Head-locked, world-locked, or conformal diminished-reality? An examination of different AR solutions for pedestrian safety in occluded scenarios

Joris Peereboom*, Wilbert Tabone*, Dimitra Dodou, Joost de Winter

Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology

* Joint first authors

Abstract

Introduction: Many collisions between pedestrians and cars are caused by poor visibility, such as occlusion by a parked vehicle. Augmented reality (AR) could help to prevent this problem, but it is unknown to what extent the augmented information needs to be embedded into the world. **Methods:** In this virtual reality experiment with a head-mounted display (HMD), 28 participants were exposed to AR designs, in a scenario where a vehicle approached from behind a parked vehicle. The experimental conditions included a live video feed of the occluded region presented directly on the HMD (*VideoHead*), a video feed displayed across the street (*VideoStreet*), and two conformal diminished reality designs: a see-through display on the occluding vehicle (*VideoSeeThrough*) and a solution where the occluding vehicle has been made semi-transparent (*TransparentVehicle*). A *Baseline* condition without augmented information served as a reference. Moreover, the *VideoHead* and *VideoStreet* conditions were each tested with and without the addition of an on-screen guiding arrow indicating the location of the approaching vehicle. Participants performed 42 trials, 6 per condition, during which they had to hold a key when they felt safe to cross. **Results:** The keypress rates and additional questionnaires showed that the *TransparentVehicle* and *VideoSeeThrough* designs came out most favourably, while the *VideoHead* solution caused some discomfort and dissatisfaction. An analysis of head yaw angle showed that *VideoHead* and *VideoStreet* caused divided attention between the screen and the approaching vehicle. The use of guiding arrows did not contribute demonstrable added value. **Conclusion:** AR designs that display a high level of local embeddedness are beneficial for addressing occlusion problems when crossing. However, the simpler head-up-display solutions should not be immediately dismissed because, according to the literature, such solutions can serve tasks where a salient warning or instruction is beneficial, and because they are easier to implement in practice.

Keywords: Augmented reality, Pedestrian safety, Virtual reality experiment, Local presence, Assisted reality, Diminished reality

Introduction

Pedestrian safety represents a global issue, with approximately 310,000 pedestrian fatalities being recorded annually across the globe, a figure that accounts for 23% of total road traffic fatalities (World Health Organization, 2023). Pedestrian collisions frequently involve a motor vehicle as the impacting entity, contributing to 63% of such incidents in a developed country like the Netherlands (SWOV, 2020). Investigative analysis of accident casualties within the United States and Europe illustrates that pedestrian behaviour is implicated in approximately two-thirds of pedestrian-vehicle collisions, and that 10 to 15% of collisions stem from obstructed views (Bálint et al., 2021; European Road Safety Observatory, 2018; Hunter et al., 1996; Yue et al., 2020).

Improvements in pedestrian safety have been made through vehicle technology like radar and camera systems, which are used to detect approaching pedestrians, alert the driver, and autonomously initiate braking or steering. Enhancements to these systems involve pedestrian path prediction, to identify pedestrians entering the road as early as possible (Rudenko et al., 2020), thereby increasing the time budget for the vehicle to respond. However, recognising a pedestrian or knowing a pedestrian's path, may not always be possible, especially when stepping on the road from between parked vehicles (Palffy et al., 2023).

Additionally, pedestrian warning systems have been developed centred around mobile devices, such as smartphones and smartwatches, to send alerts to pedestrians ahead of potential collisions (Bastani Zadeh et al., 2018; Liu et al., 2015; Won et al., 2020). In the current study, we extend this concept by investigating whether augmented reality (AR), which can be presented via a head-mounted display (HMD), could potentially offer a solution for pedestrians dealing with occlusion of approaching vehicles. More specifically, we explored whether providing a live video stream of the occluded area can improve pedestrians' behaviour and perception; we also explored the concept of diminished reality, in which occlusion in the environment is replaced by direct visibility.

A fundamental question in the design of any AR system is how the augmented feedback—in this case, video feed—should be positioned. There are different reference frames to be considered. Firstly, the information could be presented at a fixed position on the screen of the head-mounted display (HMD), an approach also known as a head-locked presentation (Lebeck et al., 2017). It is also possible to present the information at a fixed distance from the user's torso, commonly referred to as a body-locked (Klose et al., 2019) or surround-fixed (Feiner et al., 1993) presentation. An alternative is to tie the information to the world, an approach also known as world-fixed (Feiner et al., 1993) or world-locked (Lebeck et al., 2017) presentation.

An advantage of head-locked and body-locked AR is the presence of information at an accessible position, which may facilitate a prompt response, especially when information in the real world lies outside the user's immediate attention (Ghasemi et al., 2021; Schinke et al., 2010; Smith et al., 2021; Tabone et al., 2023). However, a risk is that users might respond to the cues presented without successfully integrating these with task-relevant real-world cues. This could occur either because the real-world cues have not been visually identified yet, or due to the challenges in switching cognitive and accommodative attention between the augmented and real-world cues (Chen et al., 2023; Dixon et al., 2014; Kerr et al., 2012). For similar reasons, head-locked information can be difficult to use while walking (Fukushima et al., 2020) and might induce discomfort (Caserman et al., 2021; Kaufeld et al., 2022), depending on whether the information is hard-locked or soft-locked to the user's head (MagicLeap, 2020).

World-locked AR, on the other hand, offers the advantage of projecting cues at a contextually relevant location. Specifically, when a world-locked display aligns closely with the task at hand, it can offer an intuitive and seamless mode of information processing (Bauerfeind et al., 2021; Kim et al., 2018; Robertson et al., 2008; Schankin et al., 2017; Wickens, 2021; Zhao et al., 2023). One potential drawback of world-locked interfaces, however, is that the process of locating the world-locked information might consume more time for the user compared to head- or body-locked displays, where the information is directly accessible (De Oliveira Faria et al., 2020; Lee & Woo, 2023; Tabone et al., 2023).

It is noted that the classification into head-locked, body-locked, and world-locked reference frames does not completely capture the breadth of AR designs. Some propose that it is more beneficial to envisage a continuum of 'naturalism' (Pijnenburg, 2017) or 'local presence' (Rauschnabel et al., 2022), ranging from head-locked overlays, such as textual instructions or warning symbols, to the conformal presentation of virtual information that is seamlessly integrated into the real world (Kim et al., 2016; Wickens, 2021). The placement on this

continuum is influenced not only by the choice of reference frame but also by the quality of depth rendering of virtual objects, amongst others (Rauschnabel et al., 2022).

Diminished reality (DR) refers to the process of removing or modifying specific elements from the environment, usually in real-time (Cheng et al., 2022; Mann & Fung, 2002; Mori et al., 2017). DR can be considered a subtype of AR: While AR is typically used to enhance the user's perception of the real world by overlaying virtual elements onto it, DR concerns the removal or reduction of certain elements from the physical world. There are several methods for achieving DR. One is inpainting, which involves removing an object from the current image and then filling in the gap with plausible background details to 'guess' what the background should look like (Elharrouss et al., 2020). Another approach is video see-through. This method involves using a remote video camera to capture the scene beyond the occluding object, and processing the video feed, before it is displayed to the user (e.g., Meerits & Saito, 2015; Rameau et al., 2016). In automotive settings, a number of see-through displays have been incorporated. In Samsung's Safety Truck, for example, live video images were displayed on the back of the truck, effectively granting trailing drivers the ability to 'see through' the truck (Samsung, 2015; see also Zhang et al., 2018). Gomes et al. (2012) and Rameau et al. (2016) developed a 'see-through cars' system, which provided drivers an unobstructed view of the road. Other researchers have used the concept of transparency, primarily in virtual environments. In this regard, Yasuda and Ohama (2012) tried to solve the problem of poorly visible intersections by making a wall semi-transparent, so that the approaching car could be seen through the wall. Lindemann et al. (2019) examined a semi-transparent cockpit which provided drivers with the ability to see parts of the environment that are usually blocked by the car body.

In the current study, we investigated a series of AR solutions aimed at solving the occlusion problem of pedestrians. The investigated designs ranged from a head-locked video feed of the occluded area (thus low local presence according to Rauschnabel et al. 2022's framework), a solution in which the same video feed was at a fixed distance from the user on the opposite side of the road (medium local presence), to two diminished-reality solutions that blended in with the environment (high local presence). Inspired by the above-described DR solutions in the automotive domain, two DR implementations were chosen, namely a see-through video feed of the occluding vehicle and a solution in which the occluding vehicle was made semi-transparent. Given that the different positions on the local-presence dimension may carry distinct advantages and disadvantages, as delineated above, we refrained from formulating explicit hypotheses. Instead, our primary interest lies in discerning which AR condition in the pedestrian-crossing scenario shows the most and least favourable results in terms of perceived pedestrian safety, workload, comfort, and acceptance.

Due to the challenges associated with implementing AR in real-world settings, we have chosen to test the proposed solutions in a virtual environment using a high-end HMD. While this approach may have its drawbacks, such as potential disparities in attentional switching as compared to AR evaluated in real-world settings (Gabbard et al., 2019), the use of virtual reality offers various advantages. These include enhanced experimental control and the mitigation of technical issues such as difficulties in precisely anchoring objects to the world (Walter et al., 2019; Wiesner, 2019).

As indicated above, there is a challenge of divided attention between the task-intrinsic content in the world (i.e., the approaching car) or on the augmented feedback (i.e., the video feed). Prior studies have focused on directing the user's view by means of arrows (Schinke et al., 2010) or attracting it by means of bounding boxes (Chen et al., 2015; Orlosky et al., 2019), attention funnels (Renner & Pfeiffer, 2017), flickering (Schmitz et al., 2020; Waldin et al., 2017), and contrasting (Lu et al., 2012) around points of interest. In the present study, we additionally investigated whether cues in the form of a 3D arrow that continuously indicates where the task-

intrinsic information (i.e., approaching car) is located provide added value compared to just using the video feed.

Method

Participants

A total of 28 participants, ranging from 19 to 32 years of age, with an average age of 25.0 years (SD = 3.1), partook in the experiment. Participants were students and doctoral candidates from different faculties at Delft University of Technology. The recruitment process did not offer incentives and welcomed individuals regardless of their driving experience, nationality, driving-side orientation, or age. The study received approval from the Delft Human Research Ethics Committee, approval no. 1817, with each participant providing written informed consent before commencement of the experiment.

Demographic characteristics were recorded from pre-experiment questionnaires. The majority of participants were male (n = 23). The sample predominantly comprised Dutch nationals (n = 19), but also included individuals of Indian, Italian, Belgian, Russian, and Maltese nationality. Most participants held a driver's licence (n = 25), and among these, the average driving experience was 7.12 years (SD = 3.0). Furthermore, most participants were regular pedestrians in urban environments, with 12 participants walking every day and 10 walking 4–6 days per week. In terms of digital entertainment, 10 participants reported playing video games several times a week, 8 playing approximately once a month, and 10 rarely playing or not playing anymore.

Materials

The experiment was executed using Unity 2019.4.3f1 on an Alienware PC, paired with a Varjo VR-3 HMD. The PC was equipped with an Intel i7-9700K CPU and a NVIDIA GeForce RTC 2080 Ti GPU. Participants used the spacebar on the PC's keyboard to indicate their readiness to cross. The Varjo VR-3 HMD featured a high-resolution display, offering a 90 Hz refresh rate and a 115° horizontal field of view. The focus area, spanning 27° × 27°, was rendered at 70 pixels per degree on a µOLED display, providing 1920 × 1920 pixels per eye. Meanwhile, the peripheral area was rendered at about 30 pixels per degree on an LCD, producing 2880 × 2720 pixels per eye. Four SteamVR base stations facilitated the positional tracking of the Varjo HMD. Audio was delivered through a Jabra Evolve stereo headset. After each experimental session, all equipment was sanitised with alcohol wipes. The experimental setup is depicted in Figure 1.

Participant task

In the experiment, participants were instructed to press a response key when they felt safe to cross the road. More specifically, the participants were given the following task instructions: whenever they perceived crossing as safe, they were to press and hold down the spacebar on the provided keyboard. This action was to be sustained for the duration of perceived safety. Should they feel it was no longer safe to cross, they were to stop pressing the spacebar. They were allowed to engage and disengage the key as many times as they deemed necessary. Each trial was initiated with an auditory cue of "Press now", signalling participants to begin pressing the key.

Figure 1. The experimental setup.

Virtual environment
The study employed an open-source simulator (Bazilinskyy, 2020; Bazilinskyy et al. (2020b), designed using the Unity game development platform. The simulated setting is an urban city centre, with two-lane roads and static elements like buildings, parked cars, and trees (Figure 2).

Figure 2. The virtual street in which the experiment took place.

In this study, the participant assumed the role of a pedestrian, placed on a curb behind a parked Nissan 1400 'Bakkie' and a Ford Mustang GTO. This positioning largely impeded the view of the road, with an approaching vehicle, specifically a Smart Fortwo, approaching from the pedestrian's left. We included vehicle auditory cues such as starting and driving engine sounds, their volume contingent on distance and velocity.

The in-game camera, representing the pedestrian's perspective, was set at a height of 1.67 m, approximating the average height for men and women born in 1996 (Roser et al., 2019). The pedestrian was stationed on the 0.22 m-high curb and 2.5 m away from the road edge orthogonally. The road spanned 10 m in width. The obstructive vehicle, a Nissan pickup truck, had dimensions of 3.8 m in length, 1.7 m in height, and 1.67 m in width. The pedestrian was located 4.25 m behind the rear of the pickup truck and remained stationary throughout the duration of the simulation. A top-down review of the distances is provided in the Supplementary Material (Figure S1).

In all trials, the car started from a standstill. It accelerated and made a 90-degree left turn to then approach the participant; after 4.6 s, the car had completed this turn and had reached an approach speed of 15 km/h. In non-yielding scenarios, the vehicle maintained this 15 km/h speed until the end of the trial. In yielding scenarios, the vehicle initiated deceleration at a rate of 1 m/s^2 at a temporal marker of 14.3 s, 11.64 m from the pedestrian. The vehicle stopped at 18.4 s, positioned 6.55 m from the pedestrian. Following the halt, the vehicle remained stationary for 5 s before recommencing motion. Yielding trials consumed a total duration of 31.8 s, whereas non-yielding trials took 22.0 s to complete. A representation of the distance versus time relationship is shown in Figure 3.
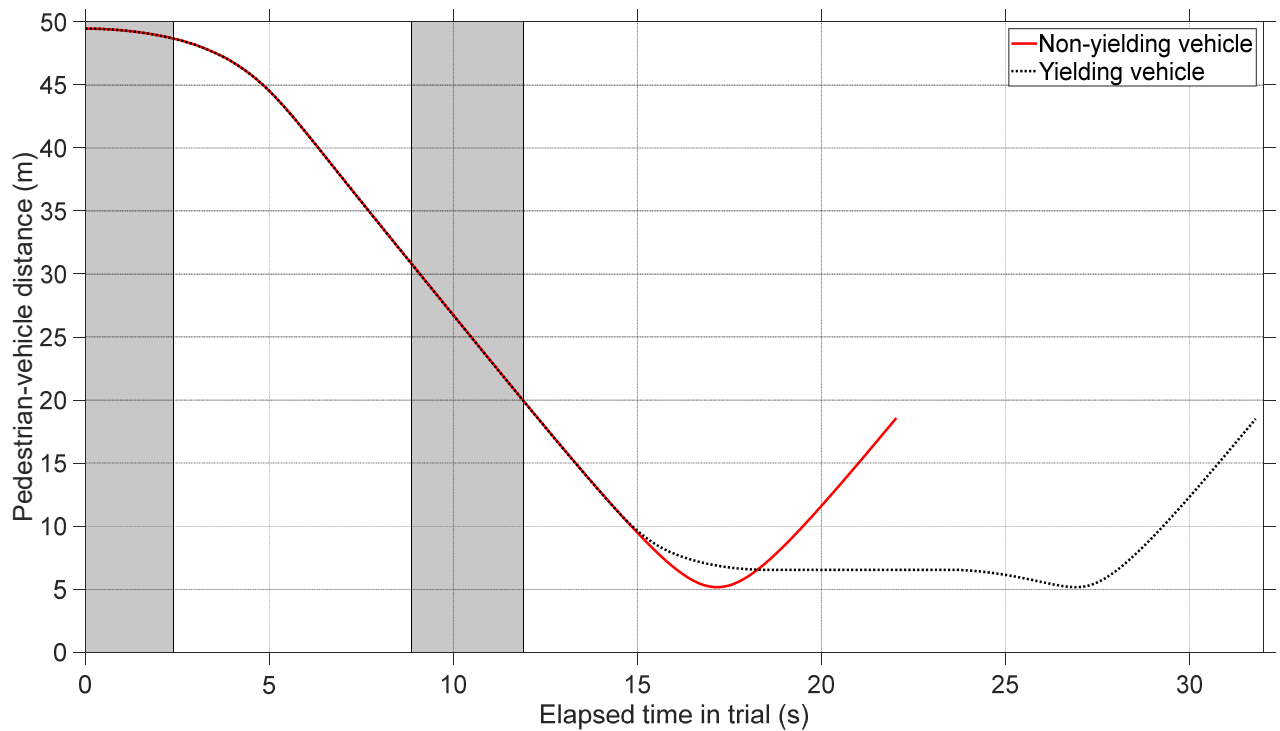
Figure 3. Distance between the approaching vehicle and pedestrian as a function of time. Grey backgrounds represent intervals where the vehicle could not be seen in the *Baseline* condition because it was fully occluded by the parked vehicle. Distances were calculated based on the Euclidean norm, using the pedestrian's location and the centre point of the vehicle.

The approaching vehicle was rendered in cyan, as this colour bears no established connotations in signalling yielding or non-yielding behaviours to pedestrians (Bazilinskyy et al., 2020a). To mimic a dart-out crossing scenario more closely, a zebra crossing was omitted from the design. The inclusion of such a crossing would implicitly suggest safety for the pedestrian to cross (in the Netherlands, traffic law mandates stopping for pedestrians poised at the curb).

Augmented Reality Designs
The experiment included a total of six AR designs, as depicted in Figure 4, with an additional condition without additional functioning as the baseline for comparisons against the other designs.

Two DR solutions were designed, a see-through display (*VideoSeeThrough*) and a semi-transparent parked vehicle (*TransparentVehicle*), along with two video feed interfaces, namely a head-locked (*VideoHead*) and a body- and world-locked version (*VideoStreet*). An additional feature was implemented in the *VideoHead* and VideoStreet designs to assist participants' views, using a continuous 'waypoint arrow' pointing to the moving vehicle (Unity, 2020). The arrow was positioned in the central top portion of the video feeds. Table 1 provides screenshots of the six AR designs and the *Baseline* tested in this experiment.

According to Rauschnabel et al. (2022)'s local-presence dimension, *VideoHead* exhibits a low local presence; the video feed is projected into the field of view but is not embedded in the world. *VideoStreet* can be described as having a medium local presence; it is present as a floating screen at a fixed position on the other side of the street, thus not truly part of the world. The two DR designs demonstrate a high local presence as the information provided is strongly embedded in the real world: *VideoSeeThrough* is presented as allowing visibility through the parked vehicle, while the transparent car also allows the participant to see through the parked vehicle.

VideoHead

VideoHead+

VideoStreet

VideoStreet+

VideoSeeThough

TransparentVehicle

Figure 4. The six augmented reality designs (screenshot resolution 2880 × 2720, consistent with the resolution of the displays in Varjo VR-3).

The positioning of the *VideoHead* display was in the upper section of the field of view, as per recommendations by Klose et al. (2019). This upper placement enabled unobstructed viewing of the road and other objects at the lower field of view. With dimensions of 8 cm both in height

and width, the display was positioned 36 cm away from the midpoint between the pedestrian's eyes. This arrangement yielded a field of view angle, both vertical and horizontal, of 12.68°.

Table 1. Overview of the six augmented reality designs.

Name	Anchoring	Video feed	Diminished reality	Guiding arrow	Local presence
VideoHead	Head-locked	Yes	No	No	Low
VideoHead+	Head-locked	Yes	No	Yes	Low
VideoStreet	Body- and world-locked	Yes	No	No	Medium
VideoStreet+	Body- and world-locked	Yes	No	Yes	Medium
VideoSeeThough	World-locked	Yes	Yes	No	High
TransparentVehicle	World-locked	No	Yes	No	High

The *VideoStreet* display was similarly positioned above road level and objects on the curb, including a bench and a waste bin, to maintain the unimpeded viewing principle. The choice of positioning was also influenced by prior research into AR in road-crossing scenarios (Tabone et al., 2021) and the contemporary traffic system, wherein traffic signals and pedestrian crosswalks typically appear perpendicular to or across the road, was a further motivation for this positioning. We consider the *VideoStreet* presentation as world-locked because it occupies fixed coordinates in the virtual world, as well as body-locked, because the participant in our experiment does not translate within the environment, and thus, the *VideoStreet* display always remains at a constant distance from the participant. The *VideoStreet* screen had a height of 300 cm and width of 400 cm, and was positioned 12.7 m distant from the pedestrian, leading to field of view angles of 13.6° vertically and 17.9° horizontally.

The *VideoSeeThrough* display, incorporated into the rear of the pickup truck, was 36 cm in height and an average width of 89 cm. When the pedestrian was positioned 4.25 m from the rear of the truck, it encompassed a vertical and horizontal field of view with angles of 4.9° and 12.0°, respectively. As for the *TransparentVehicle*, the transparency coefficient (alpha value) was set to 0.49 to generate a semi-transparent visual presentation.

The video feed was captured by a stationary camera unit centrally positioned at the top of the windshield of the pickup truck (Figure 5, top left). This camera was oriented to face forward in alignment with the road, and provided a field of view spanning 60°. Figure 5 shows examples of the camera view in various conditions.

Experimental Design
This study used a within-subjects design, with each participant undergoing all seven experimental conditions. Conditions were grouped into blocks, each containing six randomly allocated yielding and non-yielding scenarios to offset expectancy effects. The sequence of block presentations was managed using Bradley's (1958) balanced Latin Square method. In total, the experiment included 1176 trials, namely 28 participants who each conducted 6 trials for each of the 7 experimental conditions (6 AR designs + *Baseline*).

Questionnaires
Participants were provided with information regarding the study purpose, the procedural layout of the experiment, the data management process, and their right to abstain at any point, via an informed consent document. This was followed by the administration of a pre-experimental questionnaire, soliciting information on general demographics, commuting behaviour, and prior experiences with VR and gaming.

Figure 5. Top left: Windscreen camera placement. Top right: Close-up of the *VideoHead+* concept. Bottom left: Close-up of the *VideoSeeThrough* concept. Bottom right: Close-up of the *TransparentVehicle* concept.

Upon the completion of each trial block, participants were prompted to provide a measure of their current state of well-being, using the Misery Scale (MISC; Bos et al., 2005). In instances where a participant's MISC score was observed to be 4 or greater, a two-minute break was instituted. Following this, the NASA-TLX questionnaire (Hart & Staveland, 1988) was employed as a means of assessing workload. This questionnaire incorporated six variables: (1) mental demand, (2) physical demand, (3) temporal demand, (4) performance, (5) effort, and (6) frustration, evaluated via a 21-point scale ranging from 'perfect' to 'failure' for performance, and 'very low' to 'very high' for the other five items. Lastly, acceptance of the seven experimental conditions in terms of usefulness and satisfaction was evaluated using a questionnaire developed by Van der Laan et al. (1997). In this questionnaire, participants were asked to rate nine semantic-differential items on a five-point Likert scale. Usefulness was calculated as the mean of the following five items: 1. useful–useless; 3. bad–good; 5. effective–superfluous; 7. assisting–worthless; and 9. raising alertness–sleep-inducing, and satisfaction was calculated based on the following four items: 2. pleasant–unpleasant; 4. nice–annoying; 6. irritating–likeable; and 8. undesirable–desirable.

Due to the logistical impracticalities associated with asking participants to disengage from the VR environment in order to complete questionnaires, all post-block questionnaires were conducted within the virtual environment itself, while the participant was still wearing the HMD.

Participants were directed to verbally provide their responses, which were then recorded by the experimenter via Google Forms on a laptop. Figure 6 depicts the questionnaires used to evaluate well-being, workload, and acceptance.

Figure 6. Post-block questionnaires, displayed above the road in front of the participant.

Data analysis
Data was collected at a frequency of 50 Hz using a data logging script within Unity. No trials were excluded from the analysis.

For the analysis of keypress data, plots were created depicting the mean percentage of trials in which the key was pressed over time. The keypress time per design was computed by calculating the mean of the six trials per participant. This was calculated over the period from 8.86 to 14.30 s, namely from the moment the approaching vehicle was completely occluded by the parked vehicle until the approaching vehicle in the yielding condition began to brake. The initial seconds of the trials are less relevant to our analysis, as the approaching vehicle accelerated and executed a turn, and participants still had to press the key in response to the 'press now' instruction.

For the NASA-TLX responses, the 21-point scores were converted into percentage values. A composite score was then obtained by averaging the scores across the six items (Byers et al., 1989). For the acceptance questionnaire, the responses for Items 1, 2, 4, 5, 7, and 9 were mirrored, so that a higher score corresponds to higher usefulness/satisfaction. Next, scores were offset from the 1 to 5 scale (Figure 6) to align with the original scale of -2 to +2.

For the statistical tests, repeated-measures ANOVAs were used, with the AR condition as an independent variable. The results are presented as means with 95% confidence intervals, calculated for within-subjects designs (Morey, 2008).

Results
Figure 7 shows the key presses as a function of time for the seven conditions in the non-yielding scenario. It can be seen that there are substantial differences between conditions, with participants feeling safest to cross in the *TransparentVehicle* condition, and least safe in the *Baseline* condition. In the *Baseline* condition, participants tended to release the response key

as soon as the vehicle was behind the parked Nissan. For completeness, the same figures for the entire trial and for yielding and non-yielding vehicles separately are available in the Supplementary Material (Figures S2 and S3).
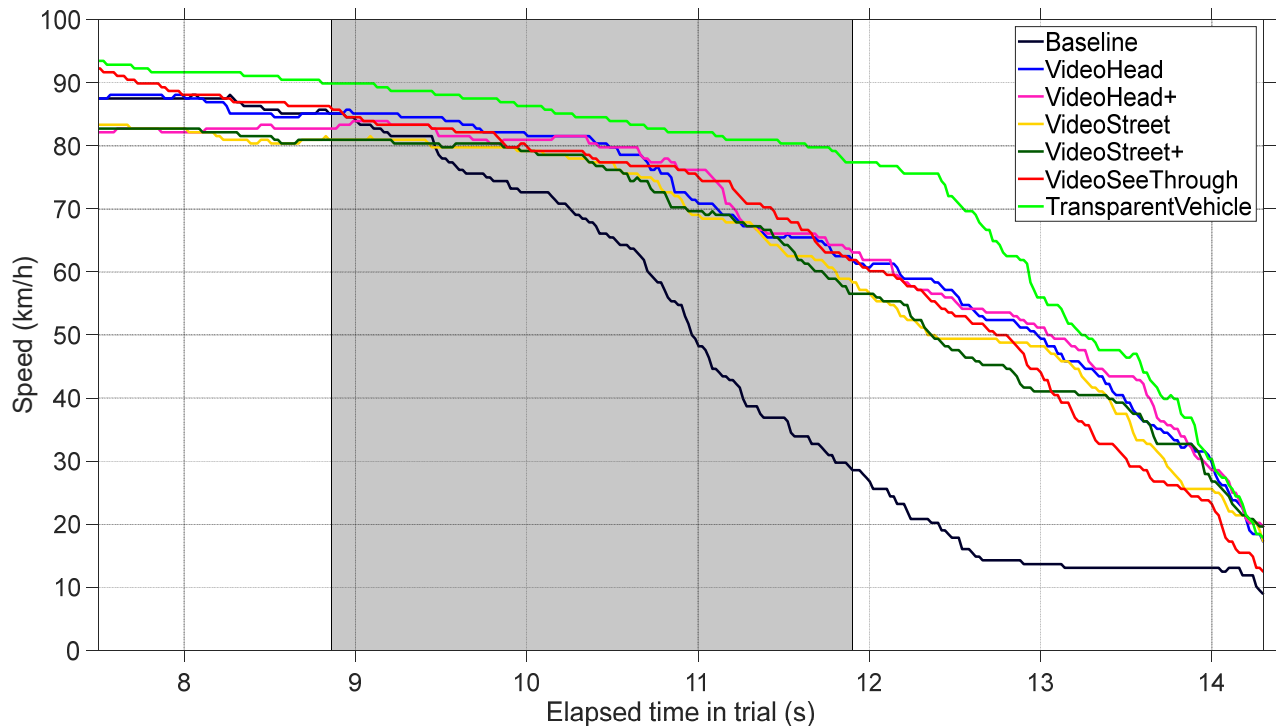
Figure 7. Key presses for the seven conditions in the non-yielding scenario. The grey background represents the interval where the approaching vehicle could not be seen in the *Baseline* condition because it was fully occluded by the parked vehicle. Across the depicted interval, the vehicle drove 15 km/h.

The perception of safety, as measured through key presses across the 8.86–14.30 s interval showed significant differences between the seven experimental conditions ($F(6,162) = 12.2$, $p = 2.90.10^{-11}$, partial $\eta^2 = 0.31$), with the highest score for the *TransparentVehicle*, and the lowest score for the *Baseline* condition (Figure 8, top left). The performance of the *VideoHead* and *VideoStreet* designs, including their guidance-included variations, was equivalent on this metric.

Usefulness (Figure 8, top middle) and satisfaction (Figure 8, top right) showed significant differences between conditions ($F(6,162) = 22.9$, $p = 1.75.10^{-19}$, partial $\eta^2 = 0.46$ and $F(6,162) = 33.3$, $p = 5.97.10^{-26}$, partial $\eta^2 = 0.55$, respectively). The *TransparentVehicle* was found to be the most useful and satisfying, followed by the *VideoSeeThrough*. A negligible difference was observed in the scores between the standard *VideoStreet* design and the one equipped with the guiding arrow, both of which attained positive scores on the scales of satisfaction and usefulness. A similar negligible difference is noted in the scores between the *VideoHead* design and the guidance-included version, although these two conditions score negatively on the satisfaction scale. Importantly, there was no substantial difference in the usefulness scores between *VideoHead* and *VideoStreet*, both with and without the guiding arrow. The *Baseline* condition attained the lowest scores, with net negative values on both the satisfaction and usefulness scales.

Self-reported workload (Figure 8, bottom left) also showed significant differences between conditions ($F(6,162) = 6.80$, $p = 1.88.10^{-6}$, partial $\eta^2 = 0.20$), with the *TransparentVehicle* showing the lowest workload, and the *VideoHead* showing the highest. The MISC scores

(Figure 8, bottom middle) showed significant differences between conditions ($F_{(6,162)} = 3.52$, $p = 0.003$, partial $\eta^2 = 0.12$), with the highest discomfort for the *VideoHead* condition.

Upon completion, the participants ranked the AR solutions. The mean ranks (Figure 8, bottom right) showed significant differences between conditions, $F_{(6,162)} = 21.4$, $p = 2.07.10^{-18}$, partial $\eta^2 = 0.44$. Participants preferred the *Baseline* condition over the *VideoHead* display, while the *TransparentVehicle* emerged as the most favoured by a considerable margin, followed by the *VideoSeeThough* display. The *VideoHead* designs were the least preferred.

It should also be noted that the seven mean values shown in Figure 8 were found to correlate with each other. For example, the mean keypress rate correlated strongly with mean self-reported usefulness ($r = 0.87$), and mean perceived satisfaction strongly correlated with mean perceived usefulness ($r = 0.82$), mean perceived workload ($r = -0.92$), and the mean preference rank ($r = -0.97$).
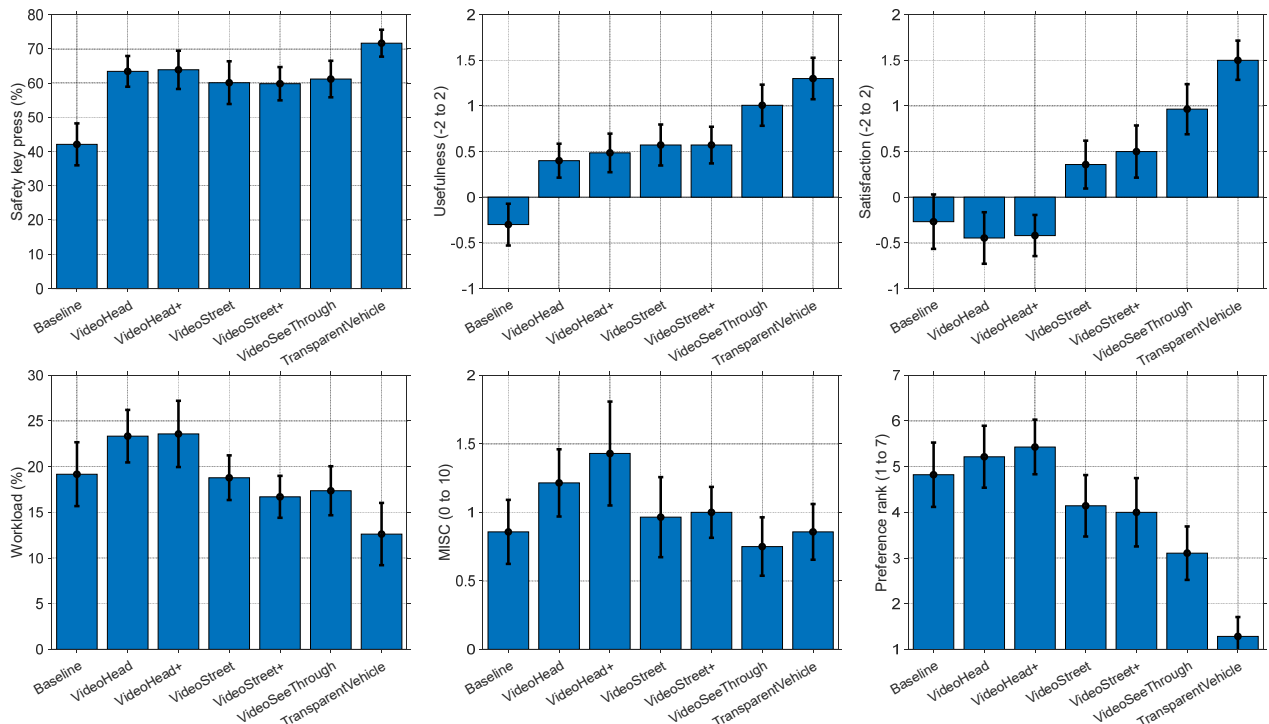
*Figure 8.* Means over the participants with 95% confidence intervals for the seven experimental conditions, for the dependent variables of this study.

Head Movement Analysis
In the conditions that featured high local presence, namely *Baseline*, *VideoSeeThrough*, and *TransparentVehicle*, participants predominantly glanced leftward, evident by a yaw angle approximating 125 degrees, aligning with the DR interfaces within the virtual environment (Figure 9, left).

On the contrary, engagement with the *VideoStreet* displays, which were positioned across the road, resulted in participants mainly focusing straight ahead with a yaw angle near 175 degrees, while intermittently looking leftward to spot the approaching vehicle (Figure 9, right).

The histogram plot of the *VideoHead* conditions reveals less defined peaks, with a peak at 130 degrees suggesting that participants predominantly focused on the parked vehicle or directly across the road at 180 degrees (Figure 9, middle).
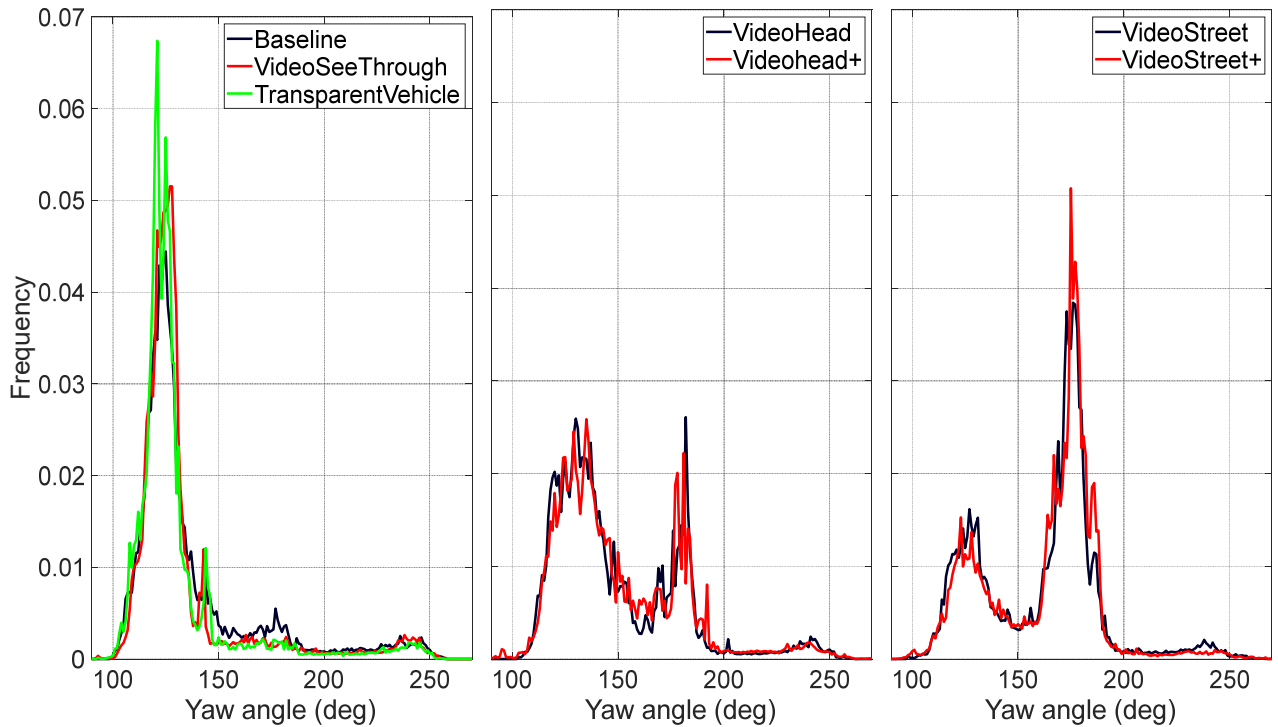
Figure 9. Distribution of yaw head rotation angles analysis during the vehicle's approaching phase (0 to 14.3 s). Note that a yaw angle of 180 degrees indicates the participant gaze across the road, with angles lesser or greater than 180 degrees suggesting a leftward or rightward gaze, respectively.

Discussion

This study employed a virtual reality setup with a HMD to examine the impact of six distinct AR designs on participants' perceived safety when crossing. The results show that pedestrians' perceived safety can increase by eliminating obstructions to their view. Particularly, in the *Baseline* condition, participants released the key earlier than with the other conditions, reflecting a diminished sense of safety. This may be attributed to participants potentially feeling cautious, given the view-blocking parked vehicle. It should be acknowledged here that the subjective feeling of safety does not necessarily imply objective safety. Nevertheless, the results clearly demonstrate that eliminating occlusion through innovative augmented reality solutions has the potential to positively impact participants' perceptions, as measured both during the experimental trials via the response key and subsequent to trial completion via a questionnaire.

There were large differences in the results between the six different AR conditions. In general, all six solutions can be considered effective compared to the *Baseline*, in terms of objective key presses and subjective usefulness. That is, participants, to a greater or lesser extent, made use of the additional information that made the occluded vehicle visible. However, when considering the results on the dimension of low local presence (*VideoHead*), medium local presence (*VideoStreet*), to high local presence (*VideoSeeThrough*, *TransparentVehicle*) (Rauschnabel et al., 2022), higher local presence was more advantageous. The transparent car, which was fully conformal and did not introduce new elements to the environment, received particularly high scores. This result is consistent with the proximity compatibility interface-design principle (Wickens & Carswell, 1995), which posits that when task elements need to be processed concurrently, it is advantageous to present corresponding cues together and integrated, rather than separately. In this respect, it is useful that the estimation could be made by directly looking at the vehicle rather than elsewhere, such as on a separate screen.

The head-locked interfaces did not receive satisfactory ratings, and they appeared to be less preferred when compared to the *Baseline*. This observation can be linked to literature within the introduction which suggests that head-locked displays may engender discomfort (Caserman et al., 2021; Kaufeld et al., 2022). In our case, the substantial degree of accommodation necessary for the *VideoHead* screen, which was presented at a close distance, could be an auxiliary explanatory factor.

Our analysis of head movements supports the notion that AR designs with a high degree of local presence permit participants to focus their attention at a single location, thereby obviating the need for attention division. While the *VideoHead* conditions somewhat alleviate the issue of divided attention between the video feed and the car (given that the video feed remains constantly in view), attentional switching is still evident. While displaying information on the opposing side of the roadway constitutes a familiar and demonstrably useful approach, evidenced, for instance, by pedestrian traffic signals (Tabone et al., 2023), a notable limitation is that it necessitates that pedestrians allocate their attention between the information across the road and the oncoming vehicle. The result is a pronounced bimodal distribution in the frequency of head movements (see Figure 9, right).

Our study established that information fully integrated within the environment yielded superior results compared to supplementary displays in the environment. This outcome appears to contradict a study by Tabone et al. (2023), where they found that a head-locked display was preferred over cues projected on the road or the approaching vehicle itself. A plausible explanation is that, in the study by Tabone and colleagues, the head-locked information comprised an unambiguous and dependable message (text: "danger! vehicle is approaching" / "safe to cross" or a red/green pedestrian traffic light) which pedestrians could use to determine whether or not to cross the road. Such explicit information could be particularly advantageous when the pedestrian is visually distracted, or has not yet identified the approaching vehicle in the environment. In our case, however, the extra screens did not provide any information that was not available in the DR solutions. Another factor is that the head-locked displays in Tabone et al. (2023) were presented in a CAVE-based simulator rather than a head-mounted display, which reduces the likelihood of discomfort (Kim et al., 2012; Pala et al., 2021).

Although previous research has demonstrated that directional arrows in AR may have a beneficial effect on performance, this has primarily been established in the context of tasks requiring navigation or spatial orientation (Gabbard et al., 2019; Liu et al., 2021; Markov-Vetter et al., 2020). In our experiment, the guiding arrow did not have any meaningful impact on the dependent variables. A plausible explanation is that the vehicle always came from the left, and the participant, who conducted a total of 42 trials, eventually became aware of the vehicle's origin. A second explanation concerns divided attention. As mentioned above, the augmented screens entailed that the participant had to divide attention between the screen and the approaching traffic. The introduction of an arrow, in addition to the screen, would necessitate even more need for divided attention, something participants may have preferred to avoid by ignoring the arrow.

The present experiment was executed in a virtual environment to ensure a controlled comparison of conditions and to minimise the risk of technological glitches. Designs of AR solutions with minimal local presence, such as the *VideoHead* design, would demand the least complex technology for real-life applications (Rauschnabel et al., 2022). In our case, a live video stream would need to be relayed and then displayed on the HMD. The *VideoStreet* design necessitates marginally more advanced technology. Here, an HMD-embedded camera could determine certain world features like the ground surface and road layout, and the screen could be presented at a predetermined height above this surface. The implementation of the two DR solutions requires more intricate tracking technology, where a parked vehicle is recognised by a camera, following which the model is either rendered transparent or augmented with a screen, while ensuring an accurate representation of the user's viewing

perspective. Despite the fact that previous DR applications have been demonstrated in real-world automotive contexts, and the concept of essentially rendering objects invisible presents an intriguing notion, the technological prerequisites along with the potential risks of artefacts, such as jitter or perspective distortions, are substantial (Overmeyer et al., 2023; Rameau et al., 2016; Wilmott et al., 2022).

A limitation of our study is that it examined fixed display sizes and positions. Investigating displays of different dimensions could provide additional insight, as the current video feed displays were relatively small. Furthermore, a more optimal positioning of the screen, especially in regard to the distance from the user, may aid in reducing feelings of discomfort. Additionally, in our experiment, the pedestrian remained stationary on the curb and could only visually scan the environment. A more naturalistic setup could be explored, such as a CAVE-based simulator, allowing measurement of pedestrian's body movement as part of the perception-action cycle. It should be noted that in this case, extra thought needs to be given to the design of the *VideoStreet* and *VideoSeeThrough* designs, in regard to the perspective of the screens. Further enhancing the realism of the crossing situation could involve incorporating other vehicles. Such vehicles would increase participants' visual demands and require divided attention.

Conclusion

This study shows the promise of removing visual occlusions through AR in enhancing pedestrians' feeling of safety. The findings indicate a user preference for AR solutions that exhibit a high degree of 'local presence'—in this case, diminished reality solutions using see-through video and transparent vehicle renderings. These designs also demonstrated the least demand for divided attention. It is important to consider, however, the technological feasibility and complexity of implementing such solutions in a real-world scenario.

Acknowledgments

Data availability

The repository with the version of the simulation used in the study is available at https://github.com/bazilinskyy/coupled-sim/tree/diminished-reality-experiment.

References

Bálint, A., Labenski, V., Köbe, M., Vogl, C., Stoll, J., Schories, L., Amann, L., Sudhakaran, G. B., Huertas Leyva, P., Pallacci, T., Östling, M., Schmidt, D., & Schindler, R. (2021). *Use case definitions and initial safety-critical scenarios* (Report No. D2.6. Project SAFE-UP).

Bastani Zadeh, R., Ghatee, M., & Eftekhari, H. R. (2018). Three-phases smartphone-based warning system to protect vulnerable road users under fuzzy conditions. *IEEE Transactions on Intelligent Transportation Systems*, *19*, 2086–2098. https://doi.org/10.1109/TITS.2017.2743709

Bauerfeind, K., Drüke, J., Schneider, J., Haar, A., Bendewald, L., & Baumann, M. (2021). Navigating with augmented reality – How does it affect drivers' mental load? *Applied Ergonomics*, *94*, 103398. https://doi.org/10.1016/j.apergo.2021.103398

Bazilinskyy, P. (2020). coupled-sim. https://github.com/bazilinskyy/coupled-sim

Bazilinskyy, P., Dodou, D., & De Winter, J. (2020a). External human-machine interfaces: Which of 729 colors is best for signaling 'please (do not) cross'? *Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics*, 3721–3728, Toronto, ON. https://doi.org/10.1109/SMC42975.2020.9282998

Bazilinskyy, P., Kooijman, L., Dodou, D., & De Winter, J. C. F. (2020b). Coupled simulator for research on the interaction between pedestrians and (automated) vehicles. *Proceedings of the Driving Simulation Conference Europe*, Antibes, France. https://repository.tudelft.nl/islandora/object/uuid:e14ae256-318d-4889-adba-b0ba1efcca71

Bos, J. E., MacKinnon, S. N., & Patterson, A. (2005). Motion sickness symptoms in a ship motion simulator: effects of inside, outside, and no view. *Aviation, Space, and Environmental Medicine*, *76*, 1111–1118.

Bradley, J. V. (1958). Complete counterbalancing of immediate sequential effects in a Latin Square design. *Journal of the American Statistical Association*, *53*, 525–528. https://doi.org/10.1080/01621459.1958.10501456

Byers, J. C., Bittner, A. C., Jr., & Hill, S. G. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A. Mital (Ed.), *Advances in industrial ergonomics and safety, Vol. I* (pp. 481–485). London: Taylor & Francis.

Caserman, P., Garcia-Agundez, A., Gámez Zerban, A., & Göbel, S. (2021). Cybersickness in current-generation virtual reality head-mounted displays: Systematic review and outlook. *Virtual Reality*, *25*, 1153–1170. https://doi.org/10.1007/s10055-021-00513-6

Chen, C. J., Hong, J., & Wang, S. F. (2015). Automated positioning of 3D virtual scene in AR-based assembly and disassembly guiding system. *The International Journal of Advanced Manufacturing Technology*, *76*, 753–764. https://doi.org/10.1007/s00170-014-6321-6

Chen, W., Song, J., Wang, Y., Wu, C., Ma, S., Wang, D., Yang, Z., & Li, H. (2023). Inattentional blindness to unexpected hazard in augmented reality head-up display assisted driving: The impact of the relative position between stimulus and augmented graph. *Traffic Injury Prevention*, *24*, 344–351. https://doi.org/10.1080/15389588.2023.2186735

Cheng, Y. F., Yin, H., Yan, Y., Gugenheimer, J., & Lindlbauer, D. (2022). Towards understanding diminished reality. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, New Orleans, LA. https://doi.org/10.1145/3491102.3517452

De Oliveira Faria, N., Gabbard, J. L., & Smith, M. (2020). Place in the world or place on the screen? Investigating the effects of augmented reality head-up display user interfaces on drivers' spatial knowledge acquisition and glance behavior. *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, 762–763. Atlanta, GA. https://doi.org/10.1109/VRW50115.2020.00232

Dixon, B. J., Daly, M. J., Chan, H. H. L., Vescan, A., Witterick, I. J., & Irish, J. C. (2014). Inattentional blindness increased with augmented reality surgical navigation. *American Journal of Rhinology & Allergy*, *28*, 433–437. https://doi.org/10.2500/ajra.2014.28.4067

Elharrouss, O., Almaadeed, N., Al-Maadeed, S., & Akbari, Y. (2020). Image inpainting: A review. *Neural Processing Letters*, *51*, 2007–2028. https://doi.org/10.1007/s11063-019-10163-0

European Road Safety Observatory. (2018). Traffic safety basic facts 2018 - pedestrians. European Commision. https://ec.europa.eu/transport/road_safety/system/files/2021-07/bfs2018_pedestrians.pdf

Feiner, S., MacIntyre, B., Haupt, M., & Solomon, E. (1993). Windows on the world: 2D windows for 3D augmented reality. *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology*, 145–155, Atlanta, GA. https://doi.org/10.1145/168642.168657

Fukushima, S., Hamada, T., & Hautasaari, A. (2020). Comparing world and screen coordinate systems in optical see-through head-mounted displays for text readability while walking. *Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality*, 649–658, Porto de Galinhas, Brazil. https://doi.org/10.1109/ISMAR50242.2020.00093

Gabbard, J. L., Smith, M., Tanous, K., Kim, H., & Jonas, B. (2019). AR DriveSim: An immersive driving simulator for augmented reality head-up display research. *Frontiers in Robotics and AI, 6*, 98. https://doi.org/10.3389/frobt.2019.00098

Ghasemi, Y., Singh, A., Kim, M., Johnson, A., & Jeong, H. (2021). Effects of head-locked augmented reality on user's performance and perceived workload. *Proceedings of the*

Human Factors and Ergonomics Society Annual Meeting, *65*, 1094–1098. https://doi.org/10.1177/1071181321651169

Gomes, P., Olaverri-Monreal, C., & Ferreira, M. (2012). Making vehicles transparent through V2V video streaming. *IEEE Transactions on Intelligent Transportation Systems*, *13*, 930–938. https://doi.org/10.1109/TITS.2012.2188289

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139–183. http://doi.org/10.1016/S0166-4115(08)62386-9

Hunter, W. W., Stutts, J. C., Pein, W. E., & Cox, C. L. (1996). *Pedestrian and bicycle crash types of the early 1990's* (No. FHWA-RD-95-163). McLean, VA: Federal Highway Administration.

Kaufeld, M., Mundt, M., Forst, S., & Hecht, H. (2022). Optical see-through augmented reality can induce severe motion sickness. *Displays*, *74*, 102283. https://doi.org/10.1016/j.displa.2022.102283

Kerr, S. J., Rice, M. D., Lum, G. T. J., & Wan, M. (2012). Evaluation of an arm-mounted augmented reality system in an outdoor environment. *Proceedings of the 2012 Southeast Asian Network of Ergonomics Societies Conference*, Langkawi, Malaysia. https://doi.org/10.1109/SEANES.2012.6299589

Kim, H., Gabbard, J. L., Anon, A. M., & Misu, T. (2018). Driver behavior and performance with augmented reality pedestrian collision warning: An outdoor user study. *IEEE Transactions on Visualization and Computer Graphics*, *24*, 1515–1524. https://doi.org/10.1109/TVCG.2018.279368

Kim, H., Miranda Anon, A., Misu, T., Li, N., Tawari, A., & Fujimura, K. (2016). Look at me: Augmented reality pedestrian warning system using an in-vehicle volumetric head up display. *Proceedings of the 21st International Conference on Intelligent User Interfaces*, 294–298, Sonoma, CA. https://doi.org/10.1145/2856767.2856815

Kim, K., Rosenthal, M. Z., Zielinski, D., & Brady, R. (2012). Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task. *Proceedings of the 2012 IEEE Virtual Reality Workshops*, 143–144, Costa Mesa, CA. https://doi.org/10.1109/VR.2012.6180922

Klose, E. M., Mack, N. A., Hegenberg, J., & Schmidt, L. (2019). Text presentation for augmented reality applications in dual-task situations. *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, 636–644, Osaka, Japan. https://doi.org/10.1109/VR.2019.8797992

Lebeck, K., Ruth, K., Kohno, T., & Roesner, F. (2017). Securing augmented reality output. *Proceedings of the 2017 IEEE Symposium on Security and Privacy*, 320–337, San Jose, CA. https://doi.org/10.1109/SP.2017.13

Lee, H., & Woo, W. (2023). Exploring the effects of augmented reality notification type and placement in AR HMD while walking. *Proceedings of the 2023 IEEE Conference Virtual Reality and 3D User Interfaces*, 519–529, Shanghai, China. https://doi.org/10.1109/VR55154.2023.00067

Lindemann, P., Eisl, D., & Rigoll, G. (2019). Acceptance and user experience of driving with a see-through cockpit in a narrow-space overtaking scenario. *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, 1040–1041, Osaka, Japan. https://doi.org/10.1109/VR.2019.8798069

Liu, B., Ding, L., & Meng, L. (2021). Spatial knowledge acquisition with virtual semantic landmarks in mixed reality-based indoor navigation. *Cartography and Geographic Information Science*, *48*, 305–319. https://doi.org/10.1080/15230406.2021.1908171

Liu, Z., Pu, L., Meng, Z., Yang, X., Zhu, K., & Zhang, L. (2015). POFS: A novel pedestrian-oriented forewarning system for vulnerable pedestrian safety. *Proceedings of the 2015 International Conference on Connected Vehicles and Expo*, 100–105. Shenzhen, China. https://doi.org/10.1109/ICCVE.2015.63

Lu, W., Duh, B.-L. H., & Feiner, S. (2012). Subtle cueing for visual search in augmented reality. *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality*, 161–166. Atlanta, GA. https://doi.org/10.1109/ISMAR.2012.6402553

MagicLeap. (2020). 5.1 Head-locked content - Unity. https://ml1-developer.magicleap.com/en-us/learn/guides/head-locked-content-tutorial-unity

Mann, S., & Fung, J. (2002). EyeTap devices for augmented, deliberately diminished, or otherwise altered visual perception of rigid planar patches of real-world scenes. *Presence: Teleoperators & Virtual Environments*, *11*, 158–175. https://doi.org/10.1162/1054746021470603

Markov-Vetter, D., Luboschik, M., Islam, A. T., Gauger, P., & Staadt, O. (2020). The effect of spatial reference on visual attention and workload during viewpoint guidance in augmented reality. *Proceedings of the 2020 ACM Symposium on Spatial User Interaction*, 10, Virtual Event. https://doi.org/10.1145/3385959.3418449

Meerits, S., & Saito, H. (2015). Real-time diminished reality for dynamic scenes. *Proceedings of the 2015 IEEE International Symposium on Mixed and Augmented Reality Workshop*, 53–59. Fukuoka, Japan. https://doi.org/10.1109/ISMARW.2015.19

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64. https://doi.org/10.20982/tqmp.01.1.p042

Mori, S., Ikeda, S., & Saito, H. (2017). A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects. *IPSJ Transactions on Computer Vision and Applications*, *9*, 1–14. https://doi.org/10.1186/s41074-017-0028-1

Orlosky, J., Liu, C., Kalkofen, D., & Kiyokawa, K. (2019). Visualization-guided attention direction in dynamic control tasks. *Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct*, 372–373. Beijing, China. https://doi.org/10.1109/ISMAR-Adjunct.2019.000-9

Overmeyer, L., Jütte, L., & Poschke, A. (2023). A real-time augmented reality system to see through forklift components. *CIRP Annals*. https://doi.org/10.1016/j.cirp.2023.03.010

Pala, P., Cavallo, V., Dang, N. T., Granié, M.-A., Schneider, S., Maruhn, P., & Bengler, K. (2021). Is the street-crossing behavior with a head-mounted display different from that behavior in a CAVE? A study among young adults and children. *Transportation Research Part F: Traffic Psychology and Behaviour*, *82*, 15–31. https://doi.org/10.1016/j.trf.2021.07.016

Palffy, A., Kooij, J. F. P., & Gavrila, D. M. (2023). Detecting darting out pedestrians with occlusion aware sensor fusion of radar and stereo camera. *IEEE Transactions on Intelligent Vehicles*. https://doi.org/10.1109/TIV.2022.3220435

Pijnenburg, J. (2017). *Naturalism: Effects of an intuitive augmented reality interface property in the display of automated driving status* (MSc thesis). Delft University of Technology.

Rameau, F., Ha, H., Joo, K., Choi, J., Park, K., & Kweon, I. S. (2016). A real-time augmented reality system to see-through cars. *IEEE Transactions on Visualization and Computer Graphics*, *22*, 2395–2404. https://doi.org/10.1109/TVCG.2016.2593768

Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a framework for augmented and virtual reality. *Computers in Human Behavior*, *133*, 107289. https://doi.org/10.1016/j.chb.2022.107289

Renner, P., & Pfeiffer, T. (2017). Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. *Proceedings of the 2017 IEEE Symposium on 3D User Interfaces*, 186–194. Los Angeles, CA. https://doi.org/10.1109/3DUI.2017.7893338

Robertson, C. M., MacIntyre, B., & Walker, B. N. (2008). An evaluation of graphical context when the graphics are outside of the task area. *Proceedings of the 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, 73–76. Cambridge, UK. https://doi.org/10.1109/ISMAR.2008.4637328

Roser, M., Appel, C., & Ritchie, H. (2019). Human height. https://ourworldindata.org/human-height

Rudenko, A., Palmieri, L., Herman, M., Kitani, K. M., Gavrila, D. M., & Arras, K. O. (2020). Human motion trajectory prediction: A survey. *The International Journal of Robotics Research*, *39*, 895–935. https://doi.org/10.1177/0278364920917446

Samsung. (2015). The Safety Truck could revolutionize road safety. https://news.samsung.com/global/the-safety-truck-could-revolutionize-road-safety

Schankin, A., Reichert, D., Berning, M., & Beigl, M. (2017). The impact of the frame of reference on attention shifts between augmented reality and real-world environment. *Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality*, 25–30. Nantes, France. https://doi.org/10.1109/ISMAR-Adjunct.2017.24

Schinke, T., Henze, N., & Boll, S. (2010). Visualization of off-screen objects in mobile augmented reality. *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, 313–316, Lisbon, Portugal. https://doi.org/10.1145/1851600.1851655

Schmitz, A., MacQuarrie, A., Julier, S., Binetti, N., & Steed, A. (2020). Directing versus attracting attention: Exploring the effectiveness of central and peripheral cues in panoramic videos. *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces*, 63–72. Atlanta, GA. https://doi.org/10.1109/VR46266.2020.00024

Smith, M., Gabbard, J. L., Burnett, G., Hare, C., Singh, H., & Skrypchuk, L. (2021). Determining the impact of augmented reality graphic spatial location and motion on driver behaviors. *Applied Ergonomics*, *96*, 103510. https://doi.org/10.1016/j.apergo.2021.103510

SWOV. (2020). Pedestrians [Fact sheet]. https://swov.nl/en/fact-sheet/pedestrians

Tabone, W., Happee, R., Yang, Y., Sadraei, E., García, J., Lee, Y. M., Merat, N., & De Winter, J. (2023). *Immersive insights: Evaluating augmented reality interfaces for pedestrians in a CAVE-based experiment* [preprint]. ResearchGate. https://www.researchgate.net/publication/370160064_Immersive_Insights_Evaluating_Augmented_Reality_Interfaces_for_Pedestrians_in_a_CAVE-Based_Experiment

Tabone, W., Lee, Y. M., Merat, N., Happee, R., & De Winter, J. (2021). Towards future pedestrian-vehicle interactions: Introducing theoretically-supported AR prototypes. *Proceedings of the 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 209–218, Leeds, United Kingdom.. https://doi.org/10.1145/3409118.3475149

Unity. (2020). Arrow waypointer. https://assetstore.unity.com/packages/tools/particles-effects/arrow-waypointer-22642

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

Waldin, N., Waldner, M., & Viola, I. (2017). Flicker observer effect: Guiding attention through high frequency flicker in images. *Computer Graphics Forum*, *36*, 467–476. https://doi.org/10.1111/cgf.13141

Walter, M., Wendisch, T., & Bengler, K. (2019). In the right place at the right time? A view at latency and its implications for automotive augmented reality head-up displays. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association* (pp. 353–358). Cham: Springer. https://doi.org/10.1007/978-3-319-96074-6_38

Wickens, C. (2021). Attention: Theory, principles, models and applications. *International Journal of Human–Computer Interaction*, *37*, 403–417. https://doi.org/10.1080/10447318.2021.1874741

Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors*, *37*, 473–494. https://doi.org/10.1518/001872095779049408

Wiesner, C. A. (2019). *Increasing the maturity of the Augmented Reality Head-Up-Display* (Doctoral dissertation). Technische Universität München.

Wilmott, J. P., Erkelens, I. M., Murdison, T. S., & Rio, K. W. (2022). Perceptibility of jitter in augmented reality head-mounted displays. *Proceedings of the 2022 IEEE International Symposium on Mixed and Augmented Reality*, 470–478, Singapore. https://doi.org/10.1109/ISMAR55827.2022.00063

Won, M., Shrestha, A., Park, K.-J., & Eun, Y. (2020). SaferCross: Enhancing pedestrian safety using embedded sensors of smartphone. *IEEE Access*, *8*, 49657–49670. https://doi.org/10.1109/ACCESS.2020.2980085

World Health Organization. (2023). Pedestrian safety: a road safety manual for decision-makers and practitioners, 2nd ed. https://www.who.int/publications/i/item/9789240072497

Yasuda, H., & Ohama, Y. (2012). Toward a practical wall see-through system for drivers: How simple can it be? *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality*, 333–334. Atlanta, GA. https://doi.org/10.1109/ISMAR.2012.6402600

Yue, L., Abdel-Aty, M., Wu, Y., Zheng, O., & Yuan, J. (2020). In-depth approach for identifying crash causation patterns and its implications for pedestrian crash prevention. *Journal of Safety Research*, *73*, 119–132. https://doi.org/10.1016/j.jsr.2020.02.020

Zhang, B., Wilschut, E. S., Willemsen, D. M. C., Alkim, T., & Martens, M. H. (2018). The effect of see-through truck on driver monitoring patterns and responses to critical events in truck platooning. In N. Stanton (Ed.), *Advances in Human Aspects of Transportation. AHFE 2017* (pp. 842–852). Cham: Springer. https://doi.org/10.1007/978-3-319-60441-1_81

Zhao, Y., Stefanucci, J., Creem-Regehr, S., & Bodenheimer, B. (2023). Evaluating augmented reality landmark cues and frame of reference displays with virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, *29*, 2710–2720. https://doi.org/10.1109/TVCG.2023.3247078
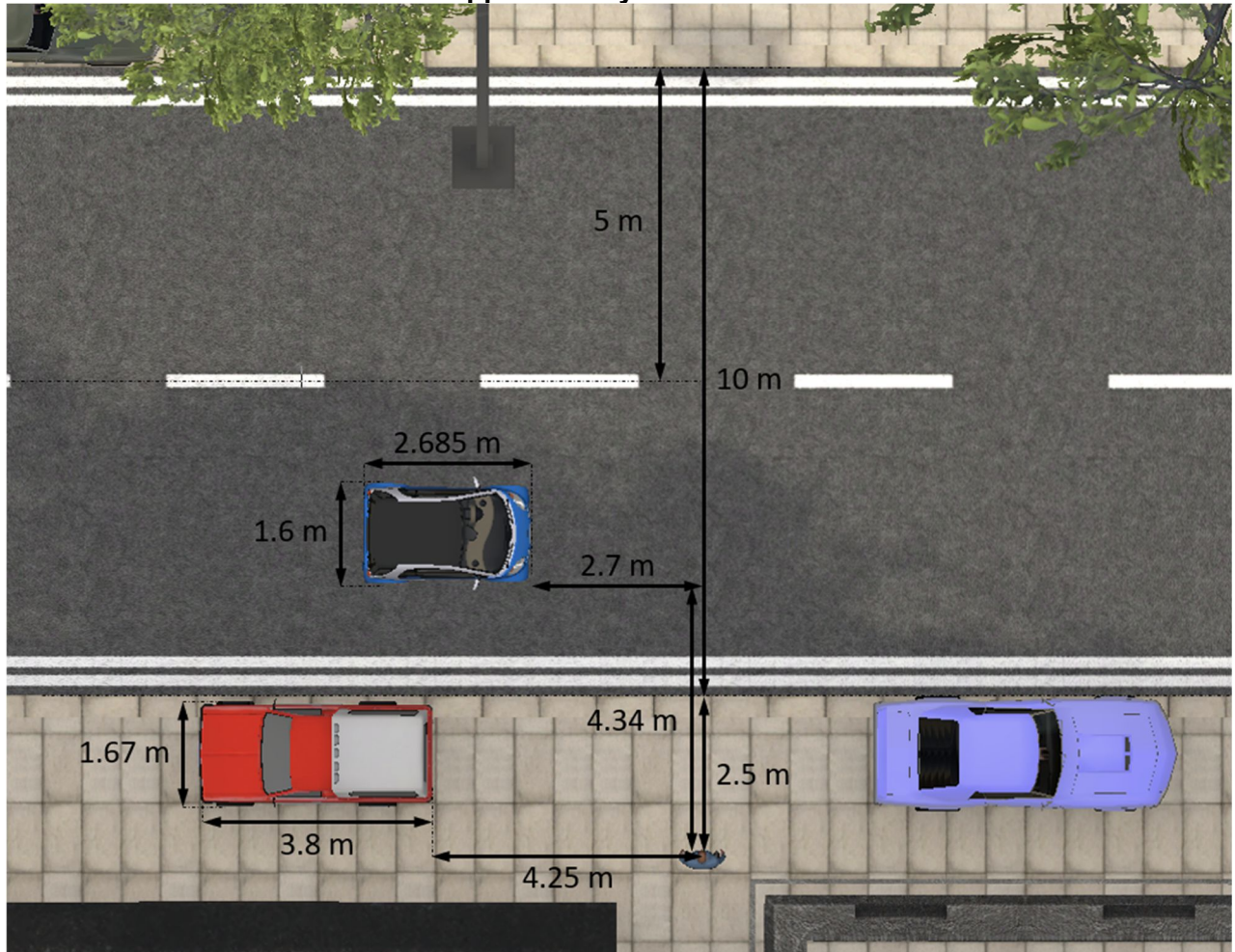
Figure S1. Top-down overview of the virtual environment. In this instance, the vehicle is positioned at the stopping point in the yielding scenario.

The vehicle commenced from a stationary position and accelerated to 15 km/h, executing a 90° left turn to approach the pedestrian. In the *TransparentVehicle* condition, the vehicle was fully visible from the start of the trial. In contrast, for the *Baseline* condition, the vehicle was not fully visible at the start of the trial. For the remaining conditions, the vehicle was visible but may have appeared small, as its depiction was confined to the virtual displays and the vehicle was still far away. These differences in the designs may explain variations in keypress rates early in the trial (i.e., before 5 s), with the *Baseline* and *TransparentVehicle* conditions eliciting the fewest and most key presses, respectively. It was also the case that for 38 of the 84 yielding trials in the *Baseline* condition, due to a programming error, the vehicle accelerated faster in the beginning of the trial (5 m/s^2) than in the other trials (1 m/s^2), as a result of which it reached its target speed more quickly. This offset has been corrected in the figure below, and these trials have been excluded in the visualisation of the yielding trials between 0 and 7.5 s. This anomaly does not apply to the non-yielding trials.

Figures S2 and S3 show that there were no discernible differences between conditions in pressing the key after the vehicle had passed the occluding vehicle, i.e., from ~ 15 s onwards. A plausible explanation is that the vehicle was consistently visible and served as the dominant cue for determining whether or not to press the key. Also, the video displays were designed such that the vehicle was not shown after it had passed by (i.e., the displays simulated a non-rotating camera mounted on the parked pickup truck and did not provide a rear-facing view).
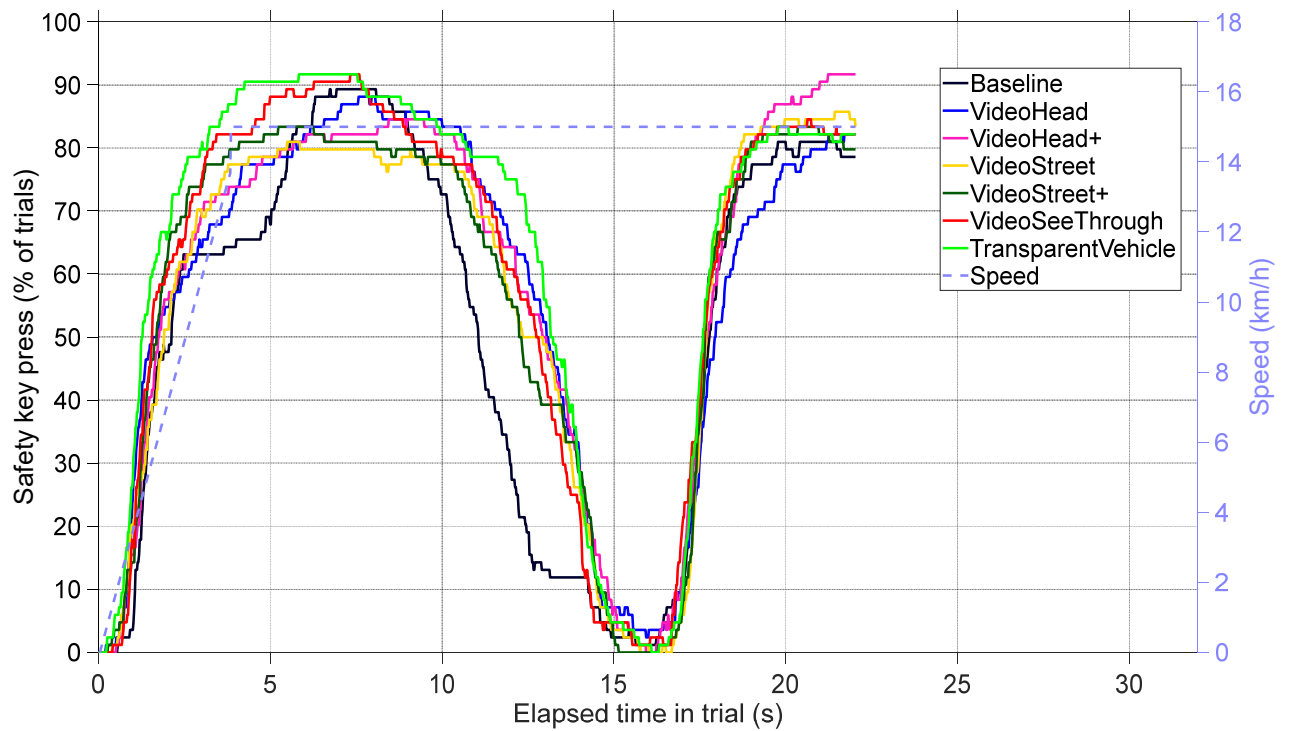
Figure S2. Key presses for the seven conditions in the non-yielding scenario, with vehicle speed.
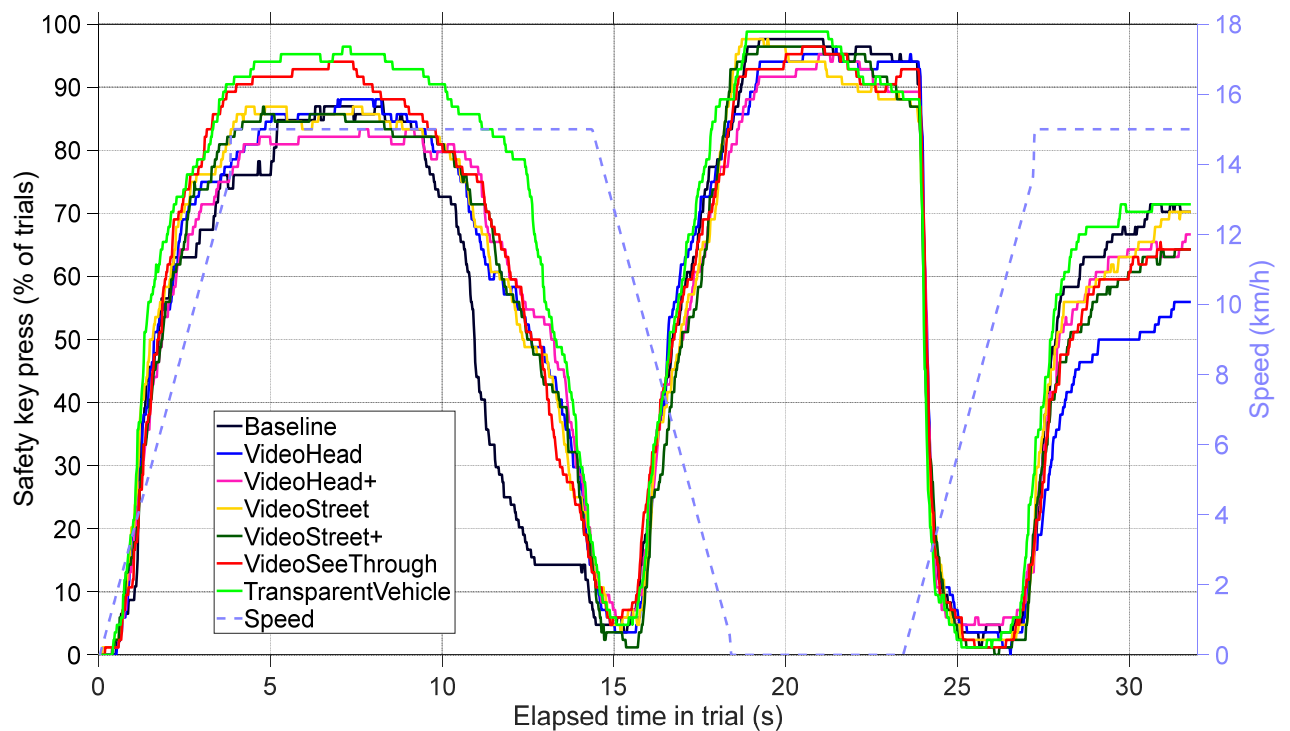

Figure S3. Key presses for the seven conditions in the yielding scenario, with vehicle speed.

23