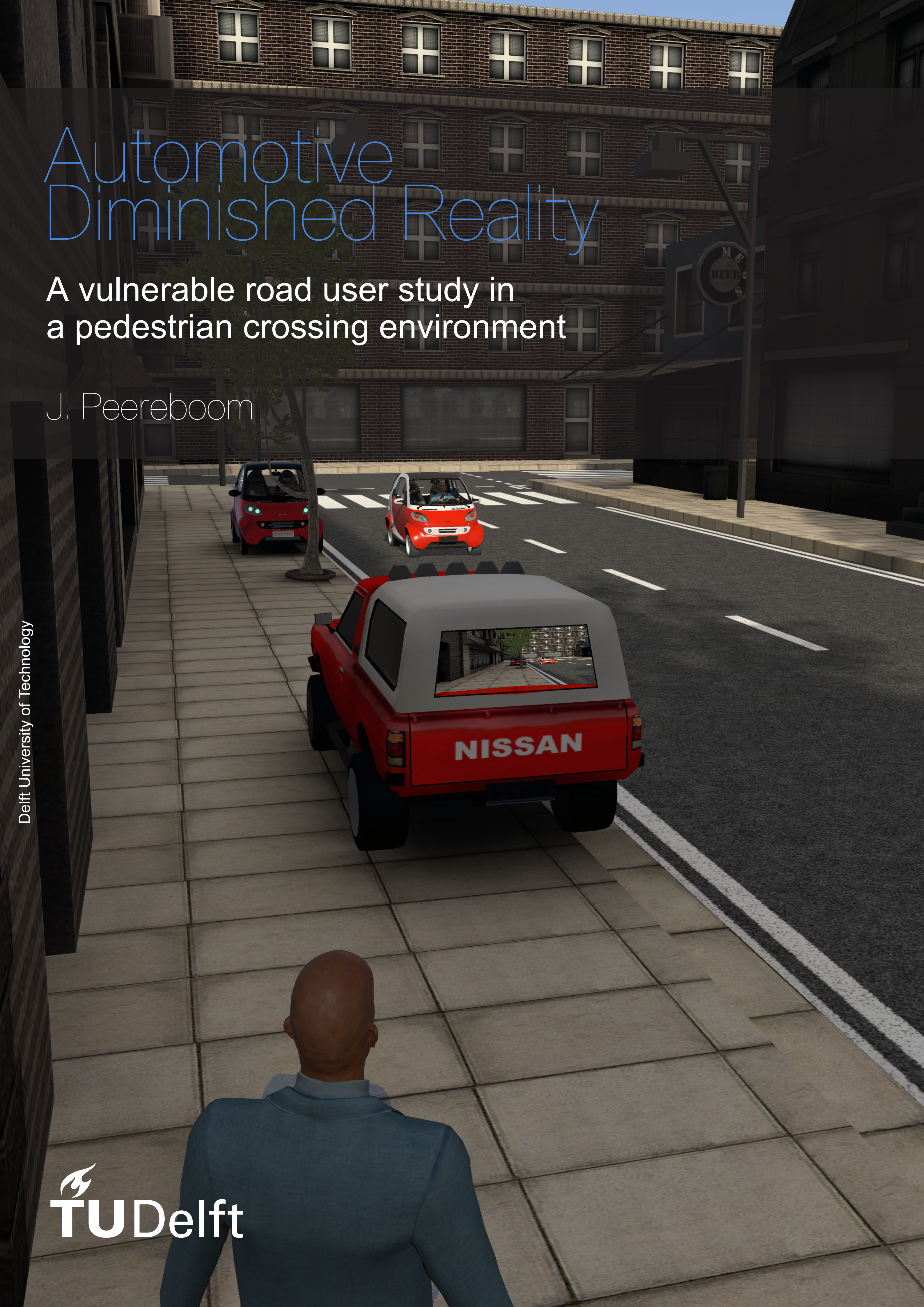
Automotive Diminished Reality

A vulnerable road user study in a pedestrian crossing environment

J. Peereboom

Delft University of Technology



TU Delft

Automotive Diminished Reality

A vulnerable road user study in a pedestrian crossing environment

by

J. Peereboom

To obtain the degree Master of Science in Mechanical Engineering
at the Delft University of Technology.

Student number:	4725212	
Supervisors:	dr. ir. J.C.F. (Joost) de Winter	TU Delft, supervisor
	dr. D. (Dimitra) Dodou	TU Delft, supervisor
	W. (Wilbert) Tabone, MSc	TU Delft, daily supervisor

TU Delft

Preface

In this thesis the effects of seven interfaces on pedestrians' crossing intentions, as measured by their perceived feeling of safety, were studied using a virtual reality setup with a head-mounted display. More specifically, the interfaces aid the pedestrian by means of camera views from unobstructed positions to help overcome the view occlusion problem that vulnerable road users suffer from. The interfaces leverage various augmented reality and diminished reality properties to convey the information to the pedestrians.

This thesis has been carried out as final part of the prerequisites to obtain the degree of Master of Science in Mechanical Engineering at the Delft University of Technology. It marks the end of my student days, and I look forward to starting a new chapter in my life.

I would like to thank my supervisors Joost de Winter, Dimitra Dodou, and Wilbert Tabone, for always being available for a question or a discussion, for providing me with feedback to improve my thesis, and for supporting me along the way. My gratitude also goes out to Vishal Onkar for helping me set up the game development program Unity and introducing me to the coupled simulator of Pavlo Bazilinskyy.

Finally, I am grateful for all participants, who have taken the time to participate in the study voluntarily. This research would not have been possible without them.

I hope you enjoy your reading.

J. Peereboom
Delft, July 2022

Abstract

Objective: In this thesis, we explore whether augmented and diminished reality interfaces, which, respectively, add and remove information from the environment, improve a pedestrian's feeling of road crossing safety, and how this information should be conveyed to the pedestrian.

Background: Literature shows that view occlusion is a prominent cause in pedestrian collisions. The research focus is currently on vehicle technology and pedestrian warning systems. Whether aiding pedestrians with camera views from unobstructed positions helps to overcome the view occlusion problem is unclear.

Methods: Twenty-eight participants engaged in a virtual reality urban road crossing scenario, in which they took on the role of a pedestrian. The pedestrian was situated on the curb and positioned such that the view on the road was largely obstructed. An autonomous vehicle approached and drove past from the left of the pedestrian. Through a head-mounted display, the participants experienced seven prototypes: baseline (i.e., no display), see-through display, transparent car, and both a head-locked and body-locked display with and without view guidance. The order in which participants encountered the prototypes was determined by a balanced Latin square, and each interface was tested by means of six trials with a non-yielding and a yielding scenario randomly selected such that in total three non-yielding and three yielding scenarios occurred in each block. The participants were instructed to continuously indicate whether they felt safe to cross by pressing a button. The interface's acceptance, workload and preference were measured with questionnaires.

Results: The participants' perceived feeling of safety revealed improved performance for all interfaces compared to the baseline condition. For the baseline condition, in which the vulnerable road user did not have access to occlusion-free information, the perceived feeling of safety was the lowest on average and decreased the earliest in the autonomous vehicle approaching phase, as well as scoring the lowest rating on acceptance. The see-through display and the transparent car interfaces, which used a combination of augmented and diminished reality properties to convey the information in a world-anchored manner to the pedestrians, achieved a higher acceptance and perceived feeling of safety than the head-locked and body-locked display interfaces.

Conclusion: A vulnerable road user's perceived feeling of safety can be increased by means of camera views from unobstructed positions to help overcome the view occlusion problem in common road crossing scenarios. This study's findings suggest a positive effect for diminished reality techniques for pedestrians, and future research could examine this technology further in more demanding scenarios.

Contents

Nomenclature

Abbreviation	Definition
AR	*Augmented Reality*: a technology that superimposes virtual (computer-generated) objects onto the real world, such that these virtual representations visually seem to coexist in the same space as the real world. An AR system is said to have the following properties: it combines real and virtual objects in a real environment, it runs interactively and in real-time, and it aligns virtual and real objects with each other (Azuma et al., 2001).
AV	*Automated Vehicle*: a vehicle capable of performing driving tasks autonomously but requiring human intervention at certain points. Automated vehicles are not to be confused with autonomous vehicles, which are vehicles capable of sensing their environment and moving safely without requiring human input (Tabone et al., 2021a).
DR	*Diminished Reality*: a set of techniques to conceal and eliminate real objects and to see through objects in order to see their occluded background (Mori et al., 2017). DR is known as a subset of AR as overlaid virtual images are used to conceal real objects, but DR applications are fundamentally different from AR in the sense that information is removed from the environment rather than added.
HMDs	*Head-Mounted Displays*: these displays are worn at a fixed distance from the eyes on the head. The display may either be a traditional video screen or a transparent one to allow see-through vision (Rolland and Hua, 2005).
HUDs	*Head-Up Displays*: HUDs present (AR) information overlaid onto and in line with the real world, eliminating the need to look away at secondary interfaces. The principle of keeping the human operator focused on the task at hand has resulted in the integration of HUDs in aircrafts and vehicles (Smith et al., 2016).
MR	*Mixed Reality*: an umbrella term for immersive technologies such as AR and VR, encompassing all variations between the real world and the virtual environment (Milgram et al., 1994).
VR	*Virtual Reality*: VR immerses users in a fully artificial digital environment, where the real world is usually blocked out as much as possible. The digital environment can be anything from virtually similar to completely different from the real world (Azuma, 1997).
VRU	*Vulnerable Road User*: road users such as pedestrians and cyclists, as well as motorcyclists and persons with disabilities or reduced mobility and orientation (Directive 2010/40/EU, 2010).

List of Figures

List of Tables

1 Introduction

1.1. Factors contributing to pedestrian collisions

Vulnerable Road Users (VRUs), such as pedestrians, constitute more than 50% of traffic-related deaths worldwide (World Health Organization, 2021). Of the road crashes among pedestrians, 63% have a car as crash opponent, and a fatality and hospitalisation rate of over 60% and 75%, respectively when the pedestrian crash occurred in an urban area (SWOV, 2020). Of all pedestrian fatalities, 74% occur at locations that are not intersections (National Highway Traffic Safety Administration, 2020). Some of the most frequent pedestrian crash types include a dart-out in the first half of the street (24%), a dart-out in the second half of the street (10%), a midblock dart (8%), walking along the roadway (7.4%), and turning-vehicle crashes (5%) (Zegeer and Bushell, 2012).

When looking into accident causation, it is reported that pedestrians are solely responsible for causing 43% of collisions, and drivers are solely responsible for 35% of collisions, with the remainder having multiple causes or caused by unknown factors (Campbell et al., 2004). Over 30% of accidents are due to missed observations, such as distraction and temporary view obstruction (European Road Safety Observatory, 2018). More specifically, pedestrians running into the road without looking (15%), stepping into the road (4.1%), stepping from between parked vehicles (7.1%), and unsafe movement (2.5%) are the main pedestrian-related accident contributing factors, while vision blockage (10.6%) is reported as a roadway and environment related factor (Campbell et al., 2004). Based on these numbers, it appears that visual occlusion is a prominent cause of pedestrian accidents.

As noted by SWOV (2020), measures that have already been applied to improve pedestrian safety include vehicle technology such as in-vehicle radar and camera systems. These are developed to detect approaching pedestrians, warn the driver, and subsequently proceed with the vehicle braking automatically. Research on future vehicle systems for active pedestrian safety argues that such systems will require not only a high pedestrian recognition performance but also an accurate analysis of the developing traffic situation and the pedestrian path trajectory (Keller and Gavrila, 2014). Systems have been developed to detect an onto-the-road-stepping pedestrian as early as possible, to increase the time budget for executing emergency braking or steering (Bartsch et al., 2012; Palffy et al., 2019).

Recognizing a pedestrian, or knowing a pedestrian's path trajectory, may not always be possible, especially when stepping on the road from between parked vehicles. A hidden pedestrian, which suddenly steps onto the road, is difficult to detect by the vehicle's camera systems due to requiring a direct line of sight to work properly, making this situation critical even for highly automated vehicles (AVs) (David and Flach, 2010;Singer et al., 2021).

Accidents related to pedestrian distraction are rising due to increased smartphone usage. Pedestrian warning systems have been developed centred around mobile devices, such as smartphones and smartwatches, to send alerts to pedestrians ahead of potential collisions (Liu et al., 2015, Zadeh et al., 2017; Won et al., 2020). These systems typically require reliable and low-latency communication between vehicles, infrastructure, and pedestrians. Additionally, GPS positioning errors are a major difficulty to overcome in improving collision warning and avoidance systems.

It can be concluded that current research focuses on vehicle technology and pedestrian warning systems. We argue that aiding pedestrians with camera views from unobstructed positions will help overcome the view occlusion problem. These camera views may originate from cameras at intersections, home video surveillance cameras, vehicle cameras, and other sources. In this thesis, we hypothesise that pedestrians will make fewer unsafe crossing decisions if their vision is free from occluding objects (blind spots) as compared to situations in which their vision is blocked. The hypothesis was tested in a scenario that involved a pedestrian on the verge of crossing a street from behind a parked car, with an approaching vehicle in the blind spot of the pedestrian. The unobstructed camera view of the parked car was transmitted to the pedestrian, to aid in their crossing decision. More specifically, the view of a front-facing windscreen camera was used, since this is a device commonly found in many modern vehicles (Ziebinski et al., 2016).

1.2. View anchoring

To solve the problem of blind spots for pedestrians, we will aid pedestrians with camera views. The question arises, however, how to optimally present the visuals to the person. As technology progresses towards wearable displays in the future, such as Head-Mounted Displays (HMDs), there are multiple positioning options to present information to the user. The placement of visual information is crucial to avoid any unnecessary occlusion and distraction caused by the display (Imamov et al., 2020).

In this thesis, we assume that HMDs will be used by VRUs in the near future, so that it is a viable means of supplying information to them. An optical HMD, such as Google Glass, is a wearable device that is able to project images on the display, as well as allowing the user to see through it, similar to Augmented Reality (AR) technology.

A three-layer categorization is described by Klose et al. (2019), with head-locked, body-locked, and world-locked positioning of content for HMD usage. Head-locked positioning is such that virtual content moves together with the head movements of the user, i.e., the content is always in their sight. Body-locked content is fixed relative to the position of the user's body, and will, for example, move along with a walking user, but not with head movement. Body-locked content is comparable to a pilot's cockpit instrumentarium, or the position of elements in a car in relation to the driver. Lastly, world-locked content is anchored to objects in the (real or virtual) world, and users can walk around it or pass by.

The literature weighs several pros and cons of why either view anchoring (world-locked content) or non-anchored views (head-locked or body-locked content) are preferred. For view anchoring or world-locked content, the following arguments, or reasons against head-locked or body-locked content, are described in the literature:

- In Schankin et al., participants' reaction times to simple colour stimuli presented either in AR or in the real world were measured. Participants responded faster to AR and real-world signals when AR signals were presented world-fixed (Schankin et al., 2017).
- Head-locked content is constantly displayed in the user's field of view, which can be distracting and annoying. This holds especially true in a mobile context of use, in which attention constantly needs to be switched to or divided between the real surroundings and the virtual information (Klose et al., 2019).
- Head-locked content may cause discomfort, and it is recommended to use body-locked placement instead (Microsoft, 2020).

For non-anchored content, i.e. head-locked or body-locked content, the following arguments are given:

- Head-locked text placement is favoured in situations where constant monitoring of the central or near-peripheral visual field is needed (Tabone et al., 2021b).
- When a viewing point is not already along the principal vision axis, the eyes and the head may need to rotate to this point of interest (Imamov et al., 2020). This is not the case with non-anchored content, as information is always at hand.

Depending on where the content is rendered, it may be necessary to rotate the head and eyes, adjust eye vergence and accommodation, and shift the focus of attention (Imamov et al., 2020). From the literature, it is clear that there are arguments for both anchored and non-anchored content, but it is unclear which is optimal when used in a road crossing scenario. Therefore, in this study, we investigate whether view anchoring is preferred or non-preferred and what the effects on comfort, and perceived pedestrian safety are.

1.3. Directional guidance with head-up displays

When a HMD user is presented with anchored content, it may occur that the content is out of their sight, i.e. not within the display area of the HMD. Similarly, when non-anchored content is displayed, it can be unclear where the origin of the content is located in the real world. Following this, questions arise whether the user's view needs to be directed towards the intended visual point of focus, or the origin of the image.

The design of overlays for guiding a user of a Head-Up Display (HUD) or HMD towards real-world objects and places has been an important topic of research in studies concerning AR (Orlosky et al., 2019). A user's view can be either directed or attracted (Schmitz et al., 2020). Directing views can be done by directly prompting the user to rotate their head, or by means of 3D arrows that illustrate the direction to the point of interest. Schinke et al. (2010) compared a 3D arrow visualisation to a mini-map, and showed that participants are faster and can interpret the position of points of interest more precisely with a 3D arrow. Guidance with arrows is also described for use in pedestrian navigation (Hile et al., 2010), and used to help mechanics to re-orient towards targets behind them (Henderson and Feiner, 2011). Attracting views can be achieved by projecting bounding boxes around the points or objects of interest, such as in assembly and disassembly tasks (Chen et al., 2015). Another method of attracting views is by using colour changing objects or flickering. Waldin et al. (2017) showed that flickering in the peripheral area is easy to detect and that it can be used to attract attention without cluttering the user's visual field. Lu et al. (2012) also explored a clutter-neutral cueing method for attracting attention by using contrasting. However, Renner and Pfeiffer (2017) found that an arrow for directing was preferred and perceived significantly more rewarding over attracting attention by means of flicker.

Literature shows that directional guidance in HUDs or HMDs is used to improve performance and situational awareness compared to no guidance (Orlosky et al., 2019). Based on findings from the literature, for objects that are in view, highlighting by means of bounding boxes or flickering to guide the attention to the correct point of interest may be used. For off-screen, or outside field of view objects, 2D or 3D arrows may be used.

1.4. Diminished Reality

A solution to the identified research gap in Section 1.1, that is the occlusion problem for pedestrians, could be the introduction of Diminished Reality (DR). Diminished Reality is a set of techniques to conceal and eliminate real objects and to see through objects in order to see their occluded background. Therefore, DR is a viable option for removing occluding objects from a pedestrian's view.

In a survey by Ericsson ConsumerLab, it was concluded that consumers are ready for more frequent use of DR applications, enabling them to customise their experience of reality. The survey found that more than one third of the respondents would like to 'edit out disturbing elements around them', such as graffiti, garbage, street signs, uninteresting shop windows, and billboards (ConsumerLab, 2017). In one of the earliest DR applications, Mann and Fung (2001) accomplished to filter out advertisements at the side of a road.

DR technology has been explored and demonstrated in many domains. In the case of interior design apps, which is an eminent utilisation field of AR, virtual representations of real furniture can be placed in their desired position in a room so that the user of such an app can visualise how well it suits their needs. A practical problem is that houses generally already contain furniture, that is to be replaced by virtual furniture, and the AR experience is nullified if a real object is in the way. Therefore, DR techniques are a powerful addition to such apps, allowing for the virtual removal of real furniture (Siltanen, 2017). In other works, DR is used to remove a person from Google Street View pictures to protect his or her privacy (Flores and Belongie, 2010), visually removing a robot manipulator in manufacturing settings (Plopski et al., 2019), and gaming (Sakai et al., 2018).

DR has also been researched in automotive environments. A well-known case is a see-through visualisation that has been implemented in Samsung's Safety Truck to reduce overtaking accidents on two-lane roads (Samsung, 2015). The Safety Truck features four high brightness, large format, outdoor displays on the rear end visible to the drivers of trailing vehicles. Cameras placed at the front of the truck transmit a view of the road ahead to drivers behind the truck, enabling drivers to perceive the road situation, and to perform safe overtaking manoeuvres. The see-through effect diminishes the body of the truck by covering it with images that are occluded from the vision of trailing drivers.

Inspired by the Safety Truck, Zhang et al. (2018) investigated whether providing platoon drivers with additional visual information of the traffic environment influences their monitoring pattern and increases awareness of upcoming situations. Their results showed that when provided with front view projection, the participants spent 10% more time monitoring the road, and responded less severely to a critical situation, suggesting a positive effect of the see-through technology.

Further aiming to solve the visibility problem in overtaking manoeuvres, Rameau et al. (2016) demonstrated a real-time AR system to see through cars. Their system was capable of running up to 15 frames per second and providing realistic and seamless see-through images from the rear car's viewpoint.

Another type of see-through system is a wall see-through visualisation at blind corner intersections, as described by Yasuda and Ohama (2012). To explore if such a system prevents crossing collisions, four different visualisation levels were implemented, ranging from displaying only the direction of a intersection-approaching vehicle to a fully transparent wall. The results indicated that the lowest level of visualisation is sufficient for proper collision estimation, although it was perceived as less informative than the richer visualisations.

A transparent cockpit, i.e., car interior, aimed to increase situational awareness and driving performance was investigated by Lindemann and Rigoll (2017). This concept allowed the driver to see occluded parts of the surroundings of the car, by visually diminishing segments of the car interior. In their driving experiment, the researchers found that driving with transparent cockpit resulted in a better overall adherence to a (ideal) reference line. This finding suggests a potential for transparency based see-through systems.

The present study differs from the papers described in this section, e.g., Yasuda and Ohama (2012) and Zhang et al. (2018), in the sense that this study focuses on improving VRU safety rather than driver safety. Nonetheless, inspiration was drawn from the described prototypes, such as the see-through displays and the transparent visualisations of objects.

In this thesis, it is hypothesised that a VRU's perceived safety can be improved through DR interfaces that create a see-through effect or a transparent visualisation on an object occluding the VRU's view on the road.

1.5. Aim of the study

This study investigates novel DR interfaces and compares them to common HUD concepts. Two DR prototypes were created that rely on see-through visualisation and transparency, while a further two HUD concepts, which introduce head-locked and body-locked views, were also implemented. The DR prototypes were compared to the HUD designs. Since the DR prototypes naturally guide the pedestrian in the direction of the approaching vehicle due to their world-locked characteristics, two more variations of the HUD concepts were introduced, with a 3D arrow for directional guidance. The experiments were conducted in a VR environment using a HMD, as VR allows for an immersive, safe, controlled, and reproducible examination of the conditions (De Winter et al., 2012).

To summarise, the following hypotheses are examined in this thesis:

- Pedestrians will make fewer unsafe crossing decisions if their vision is free from occluding objects (blind spots) as compared to situations in which their vision is blocked.
- View-anchored (i.e., world-locked) content is preferred over non-anchored (head-locked and body-locked) content.
- A VRU's perceived safety can be improved by means of removing view-obstructing objects.

2
Method

2.1. Participants

Participants were recruited among students and PhD candidates at multiple faculties of the TU Delft. No incentive was offered, and people were allowed to participate regardless of their driving experience, nationality and driving side of their country of origin, or age. The study was approved by the Human Research Ethics Committee of the TU Delft, and each participant provided written informed consent before the start of the experiment.

During the experiment, participants had to indicate their willingness to cross, and thereby their perceived safety to cross, by pressing a button. The participant's task instruction was explained at the end of the pre-experiment questionnaire and described as follows:

Each time you feel safe to cross, please do the following:

1) *Press and hold down the spacebar on the keyboard in front of you;*
2) *Keep pressing as long as you feel safe to cross;*
3) *Release if you feel it is no longer safe to cross.*

You may press and release the button as many times as you want.

At the start of every trial, a "Press now" sound instructed the participants to start pressing the button.

2.2. Virtual environment

The virtual environment was created in the open source simulator of Bazilinskyy et al. (2020b), which uses the game development platform Unity. It is a model of a city centre with a network of two-lane roads and static objects such as buildings, parked cars, and trees.

Fig. 2.1: The virtual street in which the experiment took place

The pedestrian was situated on the curb behind a parked Nissan 1400 Bakkie and a Ford Mustang GTO, and positioned such that the view on the road was largely obstructed. The AV, a Smart Fortwo, approached from the left of the pedestrian. Vehicle sounds, such as a starting engine sound and a driving engine sound, were implemented. The volume of the driving sound depended on distance and velocity.

Participants took on the role of the pedestrian and were equipped with a HMD such that they were free to look around in the virtual environment. The pedestrian's in-game camera was placed at a height of 1.67 m to align with an average height for men and women born in 1996 (Roser et al., 2013). The pedestrian was positioned on the 0.22 m high curb, and 2.5 m orthogonally away from the edge of the road. The width of the road was 10 m. The model of the view-obstructing vehicle (a Nissan truck) had a length of 3.8 m, a height of 1.7 m, and a width of 1.67 m. The pedestrian was at a distance of 4.25 m from the back of this vehicle. The pedestrian could not move through the environment.

In the design of the experiment, the AV drove at a speed of 15 km/h, and acceleration and deceleration of 1 m/s^2. Higher vehicle speeds were pilot-tested and resulted in an unrealistic driving speed as when perceived via the HMD. In yielding scenarios, the car started to slow down after 14.3 s at a distance of 11.64 m from the pedestrian. A full stop was reached at 18.4 s, 6.55 m away from the pedestrian. Distances were computed by taking the Euclidean norm of the pedestrian location and the AV car's centre. A detailed top-down overview of the distances at full stop is provided in Appendix A. The AV stood still for 5 s before driving off again. In total, yielding trials took 32.8 s and non-yielding trials lasted 22 s. The distance over time is shown in Fig. 2.2.

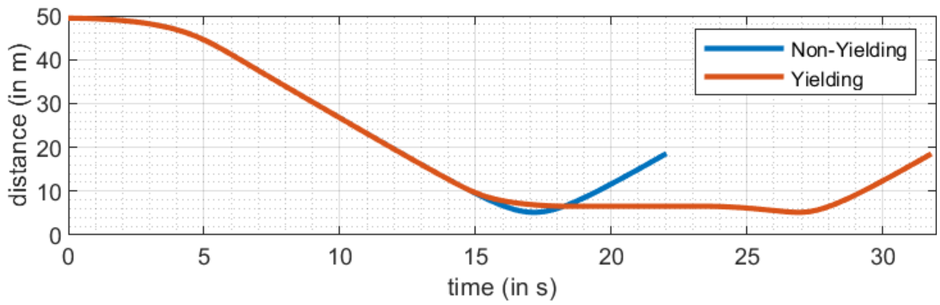


Fig. 2.2: Euclidean distance between AV and pedestrian as a function of time

For the approaching AV vehicle, a cyan colour was chosen because cyan has no association with communicating either yielding or non-yielding to pedestrians (Bazilinskyy et al., 2020a). Furthermore, a zebra crossing was not used to increase the resemblance with a dart-out crossing situation. Additionally, the presence of a zebra crossing would suggest that it is safe to cross, as in The Netherlands approaching vehicles are obliged to stop when a pedestrian stands on the curb and is about to cross a zebra.

2.3. Prototypes

The study featured a total of seven unique interfaces. The developed prototype interfaces are shown from the participant's perspective in Table 2.1. High resolution pictures from the field of view of the participant and a video showcasing the prototypes are provided in Appendix B.

A control condition without an interface served as the baseline to which the other interfaces were compared. Two DR interfaces (see-through display and transparent car) and two AR interfaces (head-locked and body-locked displays) were developed to compare the world-locked aspects of the DR interfaces to the body-locked or head-locked characteristics of the AR interfaces. For the two AR interfaces, a variation with view guidance was implemented to help aiding the view of the participant by continuously pointing to the AV with a 'waypoint arrow' (Unity, 2020).

The head-locked display was placed in the centre top section of the field of view since it is recommended to place text in top positions in complex environments that require constant monitoring (Klose et al., 2019). Top positioning also leaves the road and other objects at the bottom part of the field of view still viewable without occlusion. The same reasoning was used for the placement of the body-locked display: across the road and above objects on the curb such as a bench and a waste bin.

Baseline

See-Through Display

Transparent Car

Head-Locked Display

Head-Locked Display with guidance

Body-Locked Display

Body-Locked Display with guidance

Table 2.1: Overview of the developed prototype interfaces.

The see-through display, visually integrated in the back of the Nissan truck, had a height of 36 cm and an average width of 89 cm. With the pedestrian at a distance of 4.25 m from the back of the truck, this results in vertical and horizontal field of view angles of $4.85°$ and $11.95°$, respectively.

Fig. 2.3: Detailed view of the see-through display.

For the transparent car prototype, the alpha value was set to 0.49 for a semi-transparent appearance.

Fig. 2.4: Detailed view of the transparent car.

The head-locked display had a height and width of 8 cm, and it was placed at a distance of 36 cm from the middle point between the pedestrian's eyes, resulting in vertical and horizontal field of view angles of $12.68°$. The body-locked display had a height and width of 300 cm and 400 cm, respectively. The body-locked display was placed at a distance of 12.7 m away from the pedestrian, leading to a vertical view angle of $13.47°$ and a horizontal view angle of $17.9°$.

The video input for the see-through, head-locked, and body-locked displays originated from a camera object placed at the centre top of the Nissan truck's windscreen. This camera was facing forward, parallel with the road, and it had a field of view of $60°$.

Fig. 2.5: Windscreen camera placement.

The waypoint arrow for aiding the participant's view was programmed to continuously point to the AV. It was placed in the centre top position of the head-locked display and the body-locked display. A demonstration is provided by Fig. 2.6, where the AV is approaching from the left and the pedestrian is looking straight ahead.

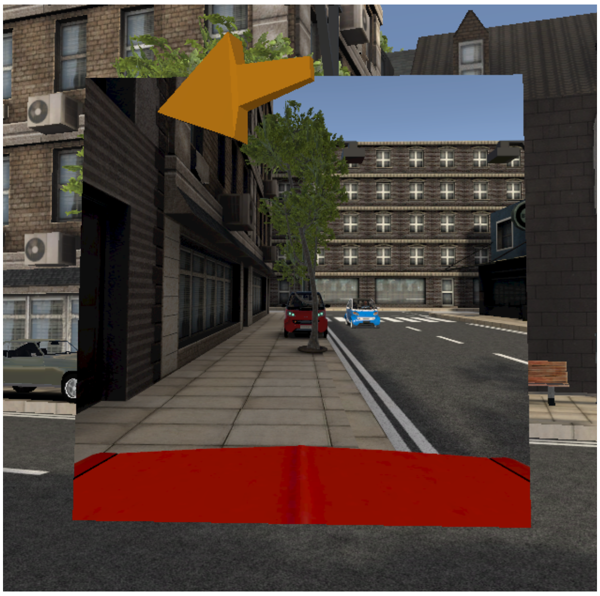


Fig. 2.6: Waypoint arrow demonstration in the head-locked display.

2.4. Questionnaires

Before the start of the experiment, participants were informed of the purpose of the study, the experimental procedure, the right to withdraw, data handling, and prevention measures related to COVID-19 by means of an informed consent form. Subsequently, a pre-experiment questionnaire containing general demographic, traffic, VR and gaming experience questions was filled out.

When wearing a VR headset, participants might experience discomfort and a maximum usage duration of 1 hr is advised (Kidwell, 2018). During the experiment, after each block of trials, the participant was first asked to indicate their well-being using a misery scale (MISC) (Bos et al., 2005). When a MISC score equal to 4 or higher was registered, a short, 2-minute break was called. Secondly, Hart and Staveland's NASA-TLX questionnaire was used for measuring workload (Hart and Staveland, 1988). An assessment of the participant's workload was made by using six items: (1) mental demand, (2) physical demand, (3) temporal demand, (4) performance, (5) effort, and (6) frustration, which were answered on a scale with 21 gradations ranging from "perfect" to "failure" for the performance item and "very low" to "very high" for the other items. Lastly, acceptance of the interface, in terms of usefulness and satisfaction, was measured using Van Der Laan's questionnaire (Van Der Laan et al., 1997). Participants responded to the statement "Your judgments of this interface are ..." for nine items. Individual item scores run from -2 to +2, but in our implementation, participants were asked to call a number between 1 and 5. The responses were afterwards transformed back to the original scale.

The post-block questionnaires were completed in the virtual environment, thus while the participant was wearing the HMD, because it was unfeasible to let participants remove the HMD and leave the virtual environment for completing the questionnaires. This would have required the participant to re-orientate in the real world, which induces a "break-in-presence (BIP)" (Jerald, 2015). Moreover, according to (Schwind et al., 2019), completing questionnaires in VR reduces a study's duration and the participants' disorientation. To accomplish this, participants were asked to read out the number corresponding to their answer, which was immediately registered on a laptop using Google Forms. Fig. 2.7 demonstrates the well-being, workload, and acceptance questionnaires as displayed to the participant.

Fig. 2.7: Post-block questionnaires, displayed above the road in front of the participant. Questions were verbally answered by reading out the number corresponding to their answer.

After the experiment, participants filled in the IGroup Presence Questionnaire (IPQ) about their experiences in the virtual environment and level of presence (Schubert, 2003). The IPQ is a 14-item questionnaire, containing questions about spatial presence, involvement, and realism, which were answered on 7-point scales. The IPQ was chosen over Witmer's presence questionnaire (Witmer et al., 2005) as the former provides the highest reliability in a reasonable time frame (Schwind et al., 2019).

2.5. Block and trial design

The experiment was of within-subject type as each participant encountered every condition. Each condition was assigned to a block, and the condition was tested by means of six trials. To counter learning effects within a block, a non-yielding and a yielding scenario were randomly selected such that in total three non-yielding and three yielding scenarios occurred in each block. To mitigate order-effects, such as fatigue, and condition carry-over effects, a balanced Latin square with $n = 7$ was used to design the order of blocks according to the method by (Bradley, 1958). A Latin square size of 7 conditions requires 14 rows to be balanced, and thus at least 14 participants conducting an equal amount of experiments. To ensure sufficient data for the analysis, it was determined to have 28 participants take part in the experiment.

2.6. Materials and Equipment

The experiment was conducted with Unity 2019.4.3f1 on a Alienware PC coupled with a Varjo VR-3 HMD. The PC had a Intel i7-9700K CPU and a NVIDIA GeForce RTC 2080 Ti GPU. The participants used the spacebar of the PC's keyboard to indicate their willingness to cross. The Varjo VR-3 features a display with human-eye resolution at a refresh rate of 90 Hz and a horizontal field of view of 115°. The focus area (27° x 27°) was rendered at 70 pixels per degree on a µOLED display, with 1920 x 1920 pixels per eye. The peripheral area was rendered at over 30 pixels per degree on a LCD, with 2880 x 2720 pixels per eye. Positional tracking of the Varjo HMD was achieved by four SteamVR base stations. A Jabra Evolve stereo headset was used for sounds. After each experiment the equipment was cleaned with alcohol wipes. The experiment setup is shown in Fig. 2.8.

Fig. 2.8: The experimental setup and a participant.

2.7. Data analysis

All data was recorded at a frequency of 50 Hz by means of a data logging script in Unity and saved as .csv files. The data was post-processed in MATLAB R2021b. No data was excluded. Graphs and tables were constructed containing the participants' means and standard deviations of the safety button scores and questionnaire responses, separated per condition, i.e. interface, and yielding or non-yielding scenario.

For the safety button press data, plots were constructed that described the mean button press percentage per interface over time and over distance between AV and pedestrian. The safety button presses per interface, for each non-yielding and yielding scenario, were computed by calculating the mean of the three trials per participant, followed by calculating the mean of all participants.

For Hart and Staveland's NASA-TLX questionnaire, the 21-point scores were transformed to percentages, and a composite score was obtained by taking the mean of the six items (Byers et al., 1989). In Van der Laan's acceptance questionnaire items 3, 6, and 8 were mirrored and the scale was converted back to the original -2 to +2 scale. For the IPQ, a similar approach as Vasconcelos-Raposo et al. (2016) was followed for a descriptive analysis of the presence questionnaire items.

3
Results

Between 22 March 2022 and 7 April 2022, 28 participants between 19 and 32 years old ($M = 24.96$; $SD = 3.1$) took part in the study and completed a total of 1176 trials (28 participants $\times$ 7 conditions $\times$ 6 repetitions per condition). Table 3.1 provides an overview of the respondents' sociodemographic characteristics as obtained from the pre-experiment questionnaire. The 25 participants with a driver's licence had a mean driving experience of 7.12 years ($SD = 3.0$).

Subject	Response category	*n*
Gender	Male	23
	Female	5
Nationality	Dutch	19
	Indian	3
	Italian	3
	Belgian	1
	Russian	1
	Maltese	1
Seeing aids	No	23
	Yes, glasses	4
	Yes, contact lenses	1
Driver's licence	Yes	25
	No	3
Vehicle driving	Every day	1
	4-6 days per week	1
	1-3 days per week	5
	Once a month to once a week	7
	Less than once a month	8
	Never	3
Driving kilometrage	Between 1 and 5,000 km	13
	Between 5,001 and 10,000 km	6
	Between 10,001 and 20,000 km	3
Pedestrian walking in urban environments	Every day	12
	4-6 days per week	10
	1-3 days per week	5
	Once a month to once a week	1
Video gaming experience	Playing several times a week	10
	Playing approximately once a month	8
	Playing rarely or not anymore	10
VR experience	Used VR headsets very often	3
	Have used VR headsets a few times	17
	Never worn a VR headset	8

Table 3.1: Overview of the pre-experiment questionnaire responses. *n* = number of respondents.

In Fig. 3.1 and Fig. 3.2, the button presses, with which the participants indicated if they felt safe to cross, for the seven interfaces are shown in the non-yielding and yielding scenario, respectively. Data per interface was averaged over trials and participants.

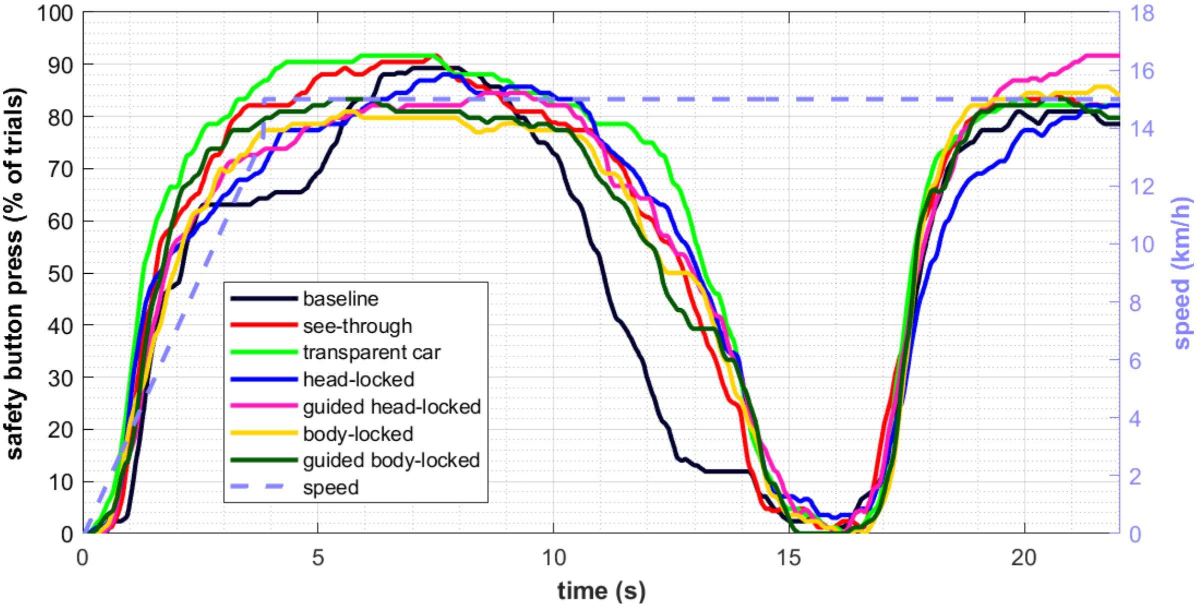


Fig. 3.1: Safety button presses for the seven interfaces in the non-yielding scenario, with AV speed.

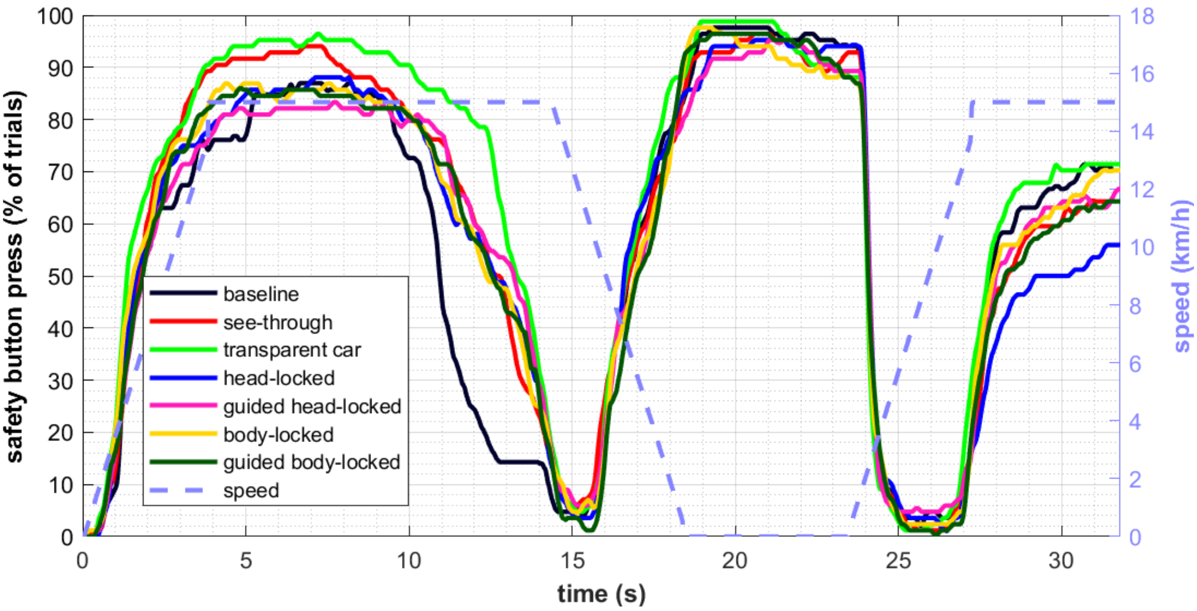


Fig. 3.2: Safety button presses for the seven interfaces in the yielding scenario, with AV speed.

In yielding scenarios, the car started to slow down after 14.3 s and was capable of stopping before crossing the pedestrian's position on the curb. Good pedestrian crossing behaviour is therefore indicated by having the safety button pressed up until 14.3 s for both yielding and non-yielding scenarios.

In the approaching phase (< 14.3 s) of both scenarios, on average, participants indicate that they felt less safe to cross and also let go of the button earlier in the baseline condition compared to the other interfaces. The transparent car and the see-through interfaces score better than the other interfaces in terms of percentage pressed. No differences are observed after the car passed the pedestrian in the non-yielding scenario, or when the car stopped and drove away again in the yielding scenario, which is illustrated in the graphs for $time > 14.3$ s.

In Fig. 3.3 and Fig. 3.4 the button presses are again shown, this time as a function of distance between the AV and the pedestrian. The distance ranges from the farthest away AV location (49.4 m), to the minimum distance where the AV drives past the pedestrian (5.2 m). Data per interface was again averaged over trials and participants.

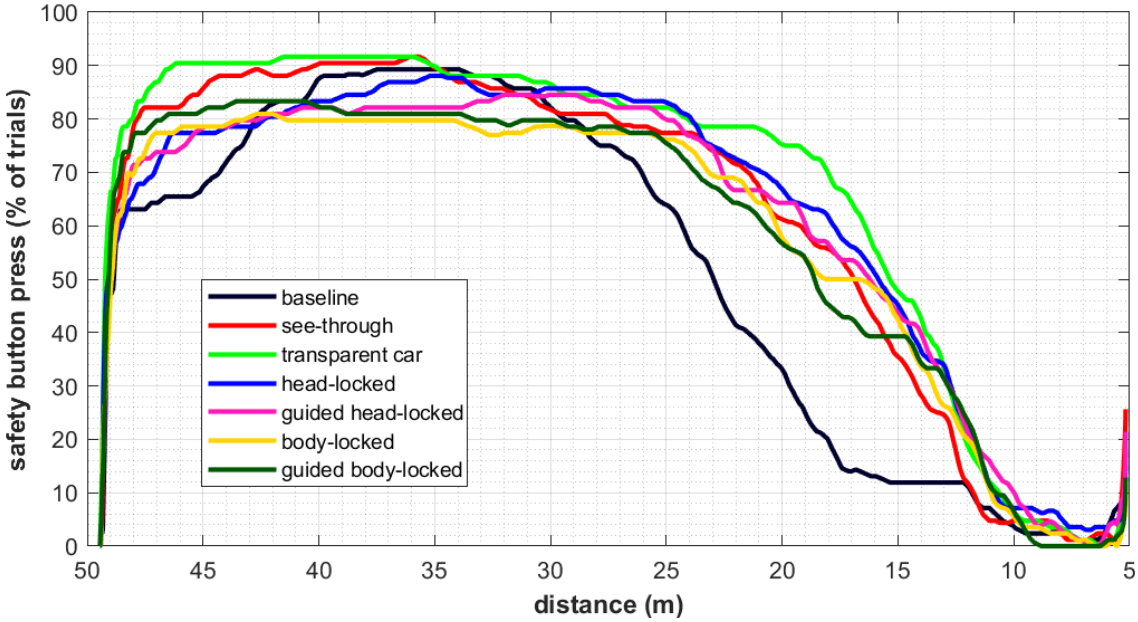


Fig. 3.3: Safety button presses as a function of distance, non-yielding scenario.

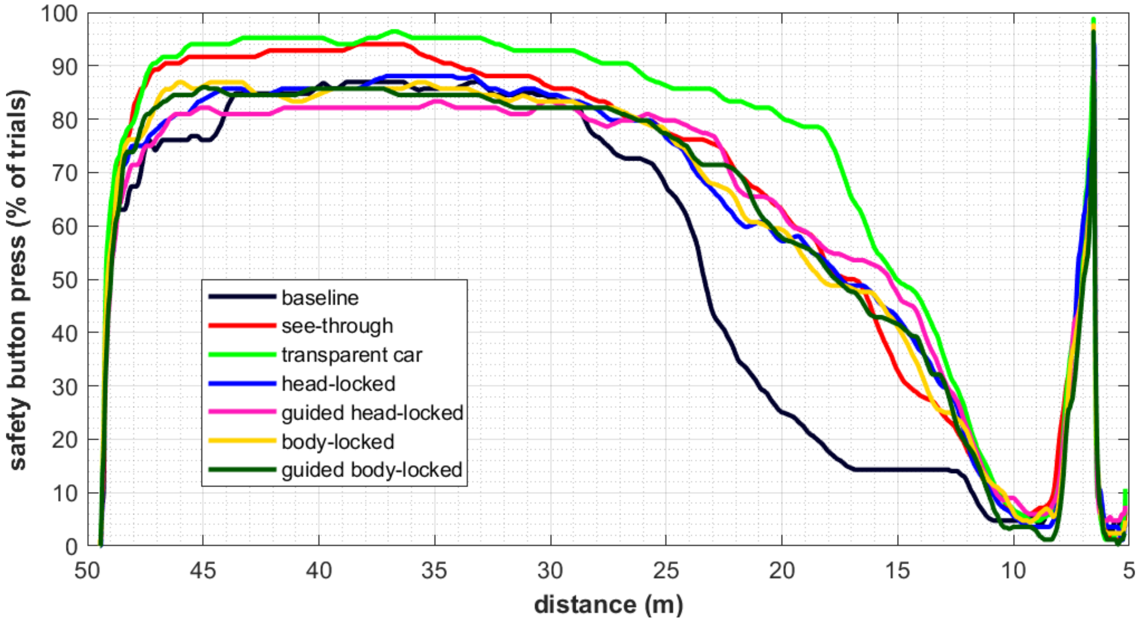


Fig. 3.4: Safety button presses as a function of distance, yielding scenario.

In the yielding scenario, the AV initiated braking at a distance of 11.64 m from the pedestrian. Good pedestrian crossing behaviour, for both yielding and non-yielding scenarios, is indicated by pressing the safety button up until this distance is reached.

For the baseline, the perceived level of pedestrian safety started decreasing at the farthest distance compared to the other interfaces, as observed by the downturn of presses at around 30 m from the pedestrian for the baseline versus 25 m for the other interfaces. The button-press percentages for the transparent car interface only started to decline for distances shorter than 20 m.

The findings from both the button presses over time and distance charts are confirmed by the mean safety button press percentages per interface in Table 3.2 below, where the mean M and standard deviation SD of the button presses of both yielding and non-yielding conditions are combined for the approaching phase of the AV (< 14.3 s).

Interface	M	SD
Baseline	0.5537	0.29
See-Through Display	0.6722	0.27
Transparent Car	0.7363	0.25
Head-Locked Display	0.6513	0.24
Head-Locked Display with guidance	0.6400	0.23
Body-Locked Display	0.6367	0.24
Body-Locked Display with guidance	0.6402	0.24

Table 3.2: Mean (SD) SafetyButton press percentage.

The perceived safety is highest with the transparent car interface, the lowest scoring interface is the baseline. The head-locked and body-locked interfaces, including the guidance variants, perform equal on this metric.

The mean acceptance scores per interface, rated on the sub-scales usefulness and satisfaction, are shown in Fig. 3.5 as error bars. The transparent car interface was experienced as most useful and satisfying, followed by the see-through interface. There is a negligible difference in the scores of the regular body-locked interface and the body-locked interface with guidance, and both score net positive on the scales of satisfaction and usefulness. The same negligible difference is observed in the scores of the head-locked interface and the head-locked interface with guidance, but score negative on the scale of satisfaction. The usefulness score does not differ meaningfully for the head-locked, guided head-locked, body-locked and guided body-locked interfaces. The baseline scores the worst, with both satisfaction and usefulness scores net negative.

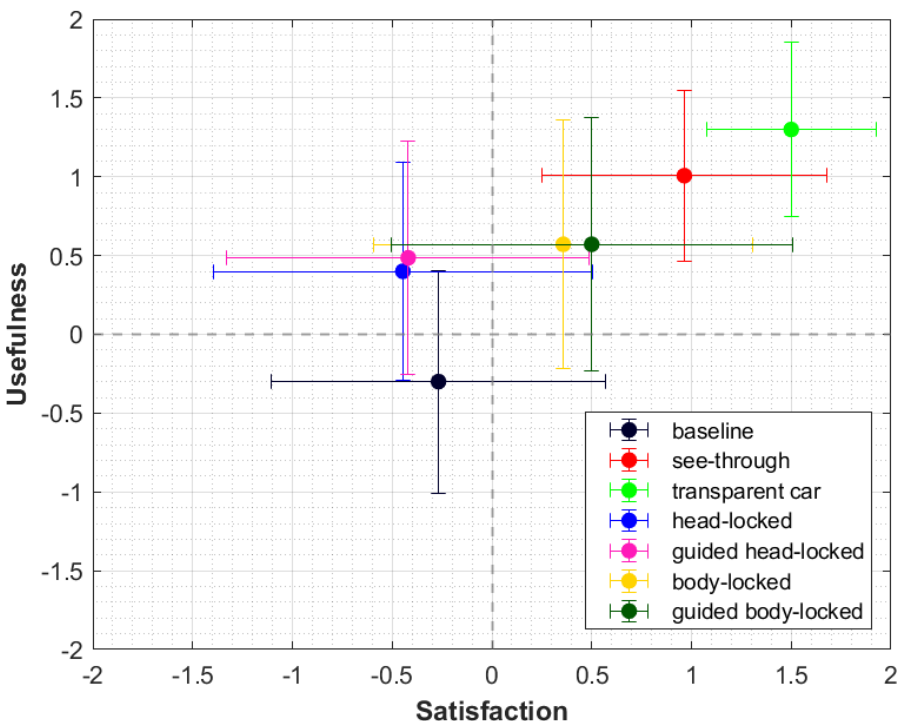


Fig. 3.5: Mean acceptance scores per interface.

A box chart of the MISC scores per interface is shown by Fig. 3.6. The misery scale shows increased uneasiness after the trials in which participants experienced the head-locked display interfaces. The lowest score is allocated to the see-through display, with a mean rating of 0.75 ($SD = 1.14$).

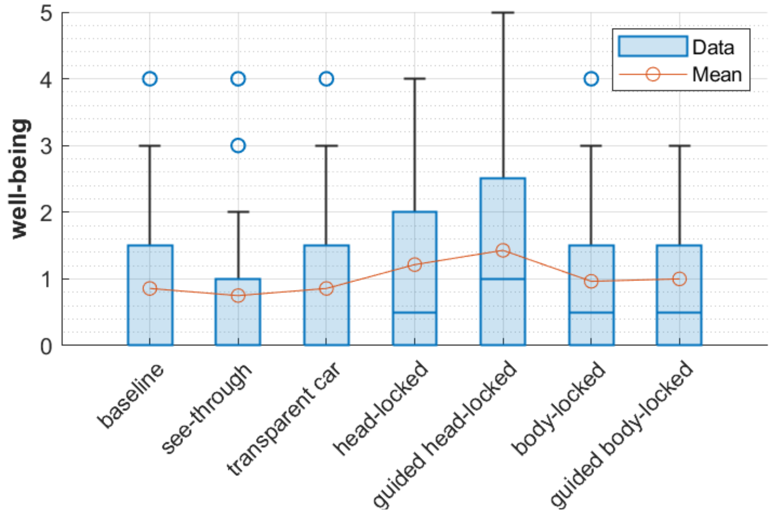


Fig. 3.6: Misery scale scores, used to indicate well-being. The red line-connected circles represent the means, the blue lines indicate the median values, and the interquartile range is defined by the error bars.

The workload scores, obtained from the NASA-TLX questionnaire, are presented in the box chart of Fig. 3.7. The mean experiment workload for all interfaces is 19.6 ($SD = 14.22$, $median = 17.86$). The transparent car interface required the least amount of effort with a mean workload of 11.1 ($SD = 11.66$). An average workload is observed for the see-through display and the body-locked interfaces. The head-locked interfaces score highest, followed by the baseline. It is notable that the variant with guidance scored lower for the body-locked interface, while it scored higher for the head-locked interface.

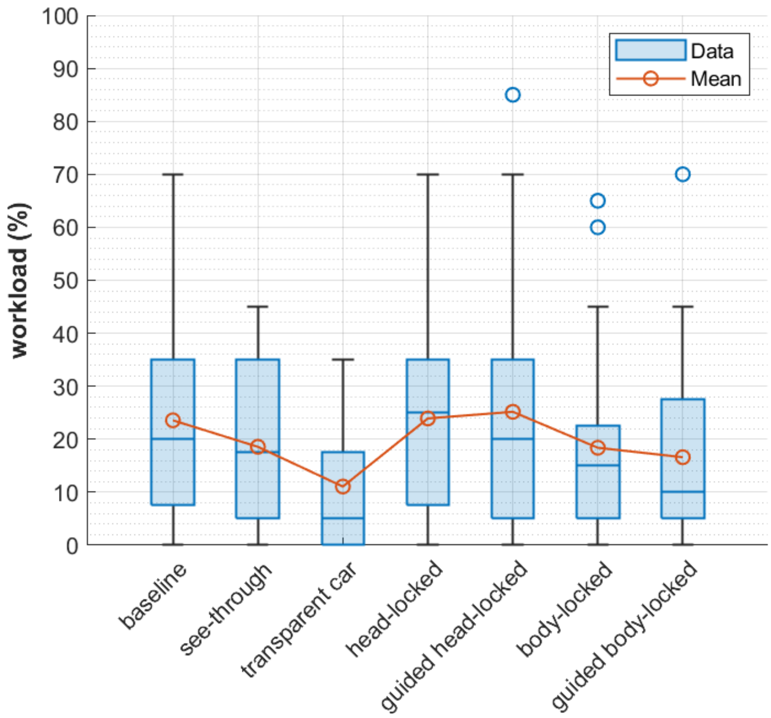


Fig. 3.7: Mean workload scores per interface. The red line-connected circles represent the means, the blue lines indicate the median values, and the interquartile range is defined by the error bars.

In Fig. 3.8 and Fig. 3.9 the participants' mean yaw head rotations over time are shown for the non-yielding and yielding scenario, respectively. Note that an angle of 180 degrees signifies the participant looking straight across the road and that angles smaller or larger than 180 degrees imply looking left or right, respectively.

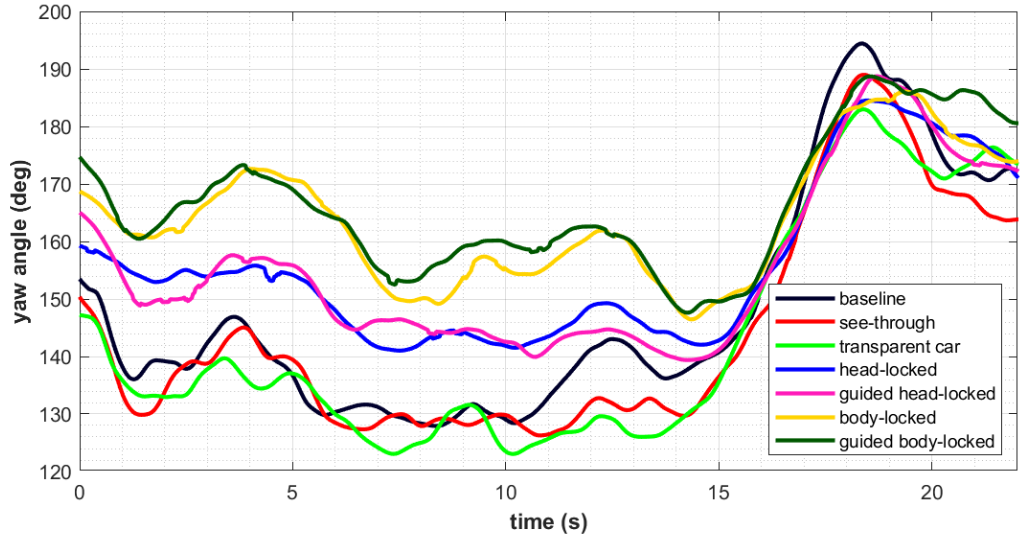


Fig. 3.8: Yaw head rotation angles over time, non-yielding scenario.

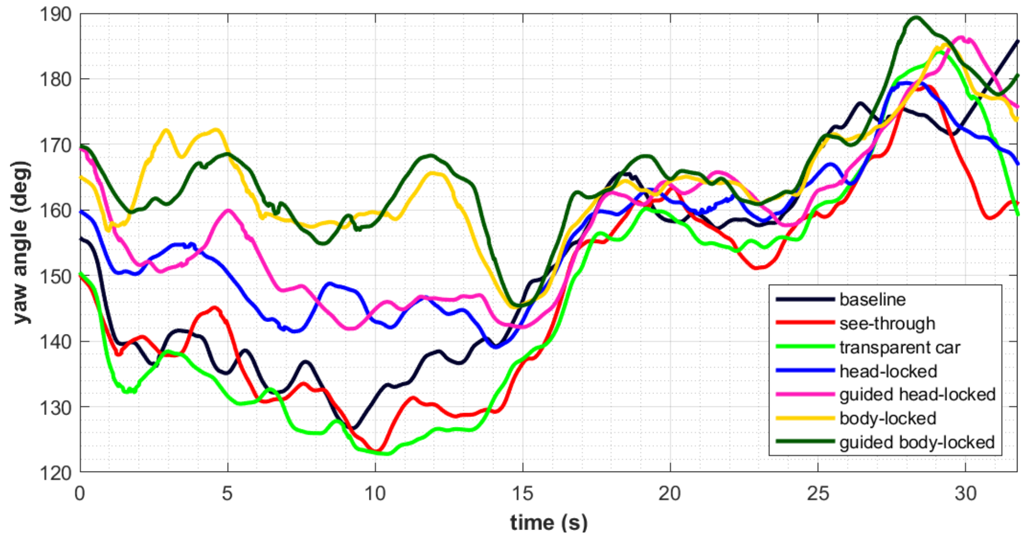


Fig. 3.9: Yaw head rotation angles over time, yielding scenario.

As these figures indicate where participants look during the experiment, they confirm that the interfaces have been used. This is especially prominent when looking at the AV approaching phase in the first 14.3 seconds, and doing an analysis of the occurrence frequency of the yaw angles in that phase (see Fig. 3.10).

For the baseline, see-through, and transparent car interfaces, participants look to the left with a yaw angle of approximately 125 degrees. This corresponds with the location of the interfaces within the virtual environment. For the body-locked and guided-body locked interfaces, which were positioned across the road, the participants mainly look straight with a yaw angle of 175 degrees, while also looking left occasionally to try and spot the approaching AV. In the frequency plot of the head-locked and guided head-locked interfaces there are less distinct peaks. The higher peak at 130 degrees indicates the participants looked at the parked car more than that they looked at the free space or across the road at 180 degrees.

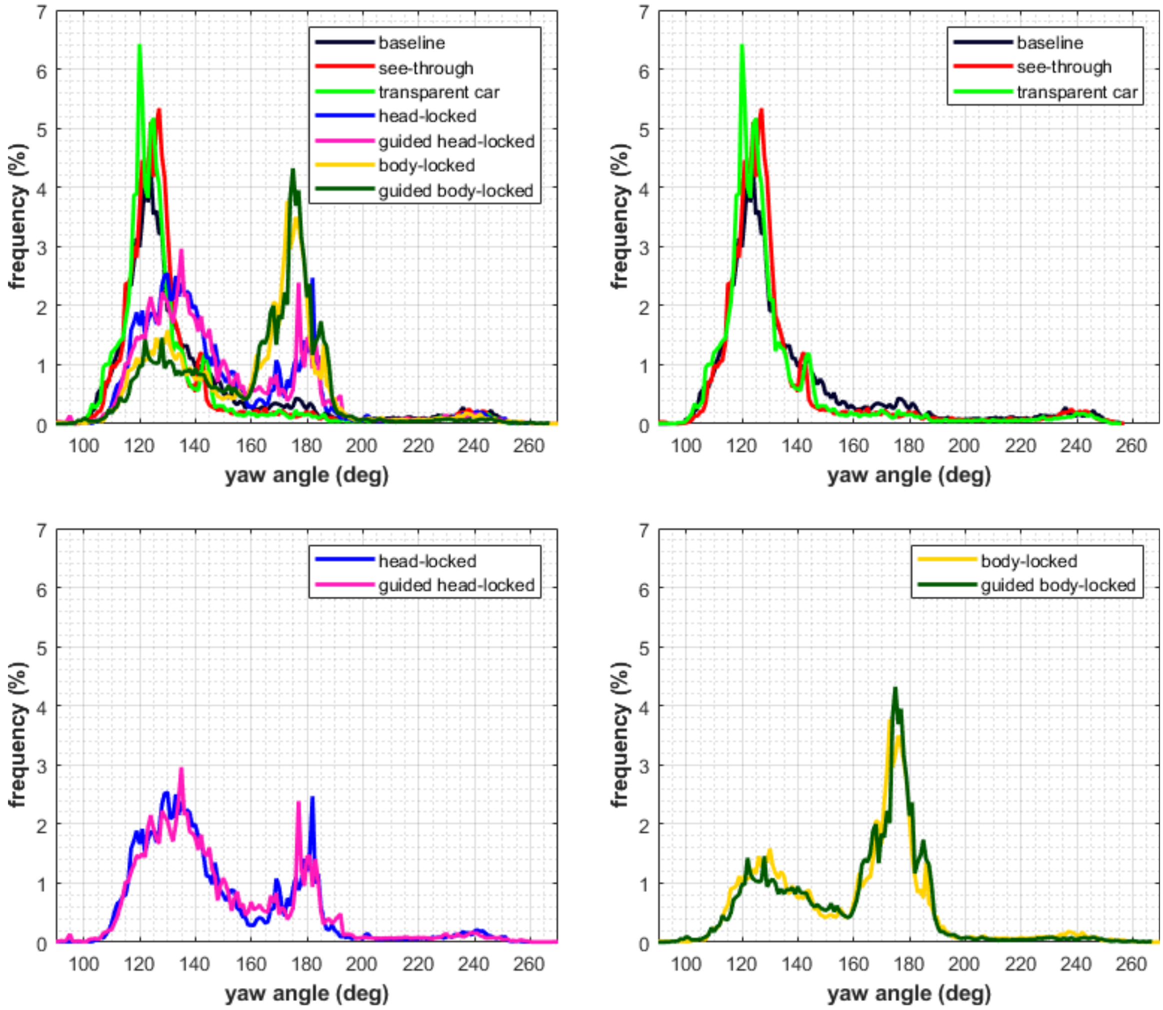


Fig. 3.10: Yaw head rotation angles analysis for the AV approaching phase.

The mean completion time of the VR experiment per participant was 30.95 min ($SD = 2.42$). Since the yielding trials and non-yielding trials have a fixed duration, all of the variance is due to completion time of the questionnaire scene. Table 3.3 shows the mean completion time in seconds per block of the questionnaire scene.

Block number	M	SD
1	198.10	54.19
2	106.84	19.05
3	97.42	31.60
4	97.36	43.93
5	78.04	20.02
6	78.10	25.41
7	69.43	24.19

Table 3.3: Mean (SD) questionnaire scene completion time in seconds.

A strong learning effect is observed which results in a 64.9% decrease in post-block questionnaire completion time from first block to last. The mean completion time was 103.61 s.

Table 3.4 summarises the participants' experiences and level of presence according to the IPQ scale. Individual items are scored on a 7-point scale.

category:	Spatial presence	Involvement	Realness	Global presence
item 1:	4.3 (1.72)	3.6 (1.70)	3.7 (1.27)	
item 2:	5.3 (1.28)	3.9 (1.80)	4.3 (1.70)	
item 3:	5.2 (1.42)	3.4 (1.91)	2.9 (1.22)	5.4 (1.23)
item 4:	5.0 (1.37)	4.6 (1.26)	1.7 (0.79)	
item 5:	2.3 (1.38)			
total:	22.1 (3.23)	15.5 (3.38)	12.4 (2.57)	5.4 (1.23)

Table 3.4: Mean (SD) IPQ responses measuring level of presence.

At last, participants completed a questionnaire asking them to rank the prototypes according to their preferences. The results (Table 3.5) show that the baseline, contrary to the acceptance scores shown earlier, was not the least preferred interface. In fact, the baseline is ranked as fifth choice, scoring higher than the head-locked display and head-locked display with guidance, which ranked sixth and seventh respectively. The most preferred interface, by a large margin, is the transparent car, followed by the see-through display. There are no clear differences between the head-locked and body-locked displays with and without guidance.

Interface	Mean rank	First choice	Second choice	Third choice	Fourth choice	Fifth choice	Sixth choice	Seventh choice
Baseline	4.82	0	2	6	3	9	0	8
See-Through Display	3.11	0	13	6	6	0	2	1
Transparent Car	1.29	25	1	1	0	0	1	0
Head-Locked Display	5.21	0	1	5	4	3	7	8
Head-Locked Display with guidance	5.43	0	2	1	3	5	11	6
Body-Locked Display	4.14	1	5	3	6	8	3	2
Body-Locked Display with guidance	4	2	4	6	6	3	4	3

Table 3.5: Number of participants per response option to the question "Which interface do you prefer the most?"

4 Discussion

This study investigated the effects of seven interfaces on participants' crossing intentions, as measured by their perceived feeling of safety, using a VR setup with a HMD. The results support our main hypothesis, namely that a VRU's perceived safety can be improved by means of removing view obstructing objects.

The participants' button presses, with which the participants showed if they felt safe to cross, indicated improved performance for all interfaces compared to the baseline condition. This finding is conforming with literature that shows that tools for aiding pedestrians in a crossings environment yield an increased feeling of safety than without such a tool (e.g., Mok et al., 2021; De Clercq et al., 2019). However, this study's first hypothesis, that pedestrians will make fewer unsafe crossing decisions if their vision is free from occluding objects (blind spots) as compared to situations in which their vision is blocked, can not be rejected nor supported by our data. At the end of the AV approaching phase (around 14.3 s), all participants released the safety button and only started pressing again once it was safe to cross. In other words, there are no differences, as no unsafe behaviour was observed. Regardless, individual differences are observed in the sense that in the baseline condition, the button is released earlier compared to the other interfaces, and with that the perceived feeling of safety drops earlier. An explanation for this is that in the baseline condition participants may feel extra cautious and agnostic, since their view is blocked. This can be explained by the fact that pedestrians tend to rely on implicit communication (e.g., Moore et al., 2019; Dey and Terken, 2017), such as the car's motion, which was a mostly unavailable cue for the pedestrians in the baseline condition.

Our second hypothesis, which argues that view-anchored (i.e., world-locked) content is preferred over non-anchored (head-locked and body-locked) content, is supported by the results of the acceptance and preference questionnaires. For the non-anchored interfaces, an explanation may be that participants had to constantly switch between the interface and the environment, which can be distracting and annoying. This result has also been found in previous research (Klose et al., 2019). We also confirmed that head-locked content causes more discomfort than body-locked and world-locked content, which has previously been stated by Microsoft (2020). Another explanation for why anchored interfaces, represented by the transparent car and the see-through display, were preferred over the non-anchored interfaces in the current experiment, is that the anchored content was placed in a natural position, slightly below the horizon, in a place that required the least amount of head- or eye-movements to and from the environment. This claim is backed by the lower workload score for the (world-locked) transparent car prototype compared to the head-locked and body-locked prototypes. This argument is contradicted by approximately equal workload scores for the see-through display and the body-locked prototypes, yet the increased workload for the see-through display can be explained by the relatively small display size compared to the body-locked prototypes.

Regarding the waypoint arrow used in this study for the head-locked and body-locked displays with view guidance, we believe that providing contextual guidance for more than just the one vehicle in this experiment is complex to achieve unambiguously. Similar to Tabone et al. (2021a), we argue that it is better to let a VRU interpret non-anchored HUD content on their own, without extra guidance in the form of an arrow. The results indicate that the waypoint arrow for view guidance had no meaningful effect, as metrics on perceived safety, acceptance, workload, and preference show near equal scores for their non-guided interface equivalent.

5

Limitations and recommendations

The technical difficulties and practical feasibility of real life implementations of the prototypes used in this experiment were not considered during the design process. Many technical challenges need to be overcome, such as establishing a vehicle to pedestrian communication system for transmitting images from the front-facing windscreen camera. Furthermore, difficulties regarding wearable optical HMDs, i.e. AR glasses, are persistent even after years of development, such as image brightness and battery life, and mass adoption has not (yet) occurred. Other cases regarding AR that need to be dealt with in future traffic are privacy, invasiveness, user-friendliness, and inclusiveness (Tabone et al., 2021a).

A second limitation is that only fixed size interfaces and positions were considered. It would be interesting to explore differently sized see-through displays, as it can be argued that the dimensions of the see-through display were small compared to the head-locked and body-locked displays, or the transparent car visualisation. Furthermore, circular AR see-through spots with varying levels of transparency and diameters can also be implemented, similar to Lindemann and Rigoll (2017).

A third limitation was the fixed driving speed setting of the AV of 15 km/h. While this was perceived as a realistic driving speed in the simulator, it makes comparing to real life crossing scenarios hard. A comment from a participant was that a decisive factor to commit to crossing was not the interface, but vehicle speed. Therefore, investigating varying levels of AV speed seems worthwhile, while it also allows for exploring different temporal windows of possible interaction with the AR interfaces.

While eye tracking on the used HMD was available, it was chosen not to use it. In the study of Mok et al. (2021), the eye tracking proved unreliable as it had to be recalibrated often during the experiment, and thus it was chosen to eliminate eye tracking to decrease our VR system complexity. As a consequence, the participant's effort based on eye movements, i.e. constantly switching from and to a HUD or DR interface, could not be measured. Furthermore, the yaw angle analysis, which investigated where participants look during the experiment, could have been more accurate if an aggregate yaw angle, from both head rotations and eye rotations, had been used.

Lastly, we hypothesised in internal discussions, based on the definition of world-locked anchoring (displacement- and rotation-locked to world objects), that in a scenario in which a world-locked interface is assisting a pedestrian to make safe crossing decisions, it helps with looking in the right direction, i.e. the direction from which a vehicle is approaching. The mean yaw head rotation data suggest that this may the case, but we did not explicitly test this hypothesis. A future study could investigate this effect further.

Several recommendations for future research are plausible from the limitations in diversity or realistic resemblance of this experiment. It could be worthy to explore the interaction of elderly people with AR interfaces, as the participants in this study were predominantly young adults.

Secondly, in our experiment, the pedestrian was fixed on the curb, and could only look around. A (simulated) approach at typical walking speed to the curb edge can be explored, similar to Kaleefathullah et al. (2020), as this enhances the resemblance with a real life crossing scenario, allowing to measure the pedestrian's approaching hesitancy and other behaviour. Such an experiment, in which the pedestrian is free to move around or constrained along a trajectory, would also allow for the implementation of body-locked interfaces which move along with the pedestrians' position.

Further improving the realism of the crossing situation would be to add distraction vehicles (e.g., Mok et al., 2021; De Clercq et al., 2019). Distraction vehicles would add to the participants' workload and should therefore increase task engagement. In our experiment, workload was low as the participant's task was not demanding. Therefore, there is room to experiment with more variation and taxing elements.

Lastly, since the baseline condition was included in the balanced Latin square design for determining the order of blocks, there were cases in which the baseline condition came at the end for some participants. As the baseline served as the reference condition, it sometimes posed difficulty for answering the (subjective) questionnaires, especially with Van Der Laan's acceptance questionnaire. It is therefore recommended to include a test trial at the beginning to allow the participants to familiarise with the environment.

There are also two recommendations for future researchers to adopt practices of this study.

The well-being MISC scores are low, despite prolonged use of a VR HMD of approximately 30.95 min. Considering the fact that seven post-block questionnaires were also conducted in VR, which took about 12 min in total, substantial time savings were achieved by not constantly having to take off the HMD between blocks to do the questionnaire on paper. Furthermore, this decreases the risk of damaging the apparatus, as well as reducing the number of "break-in-presence (BIP)". If a VR experiment is around 30 min, we recommend keeping the HMD on for the full duration of the experiment, but nonetheless monitor well-being regularly.

The game development platform Unity offered a great freedom to design this VR experiment, and coupled with C# programming to create a continuous loop of trials, a complete and convincing experience was delivered to the participants. For this reason, it is recommended to use Unity as a development engine in VR studies.

6
Conclusion

The results of this research confirm that a VRU's perceived feeling of safety can be increased by means of camera views from unobstructed positions to help overcome the view occlusion problem in common road crossing scenarios. For the baseline condition, in which the VRU did not have access to occlusion-free information, the perceived feeling of safety was the lowest on average and decreased the earliest in the approaching phase of the AV, as well as scoring the lowest rating on acceptance. An explanation for this is that in the baseline condition participants may felt extra cautious and agnostic, since their view was blocked. The see-through display and the transparent car interfaces, which use a combination of AR and DR properties to convey the information in a world-anchored manner to the pedestrians, achieved a higher acceptance and perceived feeling of safety than the head-locked and body-locked display interfaces. Furthermore, a lower workload was achieved for the transparent car interface. These findings suggest a positive effect for diminished reality techniques for pedestrians, and future research could examine this technology further in more demanding scenarios.

References

Azuma, R. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, *6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE computer graphics and applications*, *21*(6), 34–47. https://doi.org/10.1109/38.963459

Bartsch, A., Fitzek, F., & Rasshofer, R. (2012). Pedestrian recognition using automotive radar sensors. *Advances in Radio Science*, *10*, 45–55. https://doi.org/10.5194/ars-10-45-2012

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

Bazilinskyy, P., Kooijman, L., Dodou, D., & De Winter, J. (2020b). Coupled simulator for research on the interaction between pedestrians and (automated) vehicles. *Proceedings of the DSC*. http://resolver.tudelft.nl/uuid:e14ae256-318d-4889-adba-b0ba1efcca71

Bos, J., Mackinnon, S., & Patterson, A. (2005). Motion sickness symptoms in a ship motion simulator: Effects of inside, outside and no view. *Aviation, Space, and Environmental Medicine*, *76*, 1111–8.

Bradley, J. (1958). Complete counterbalancing of immediate sequential effects in a latin square design. *Journal of the American Statistical Association*, *53*(282), 525–528. https://doi.org/10.1080/01621459.1958.10501456

Byers, J., Bittner, A., & Hill, S. (1989). Traditional and raw task load index (tlx) correlations: Are paired comparisons necessary? *Advances in Industrial Ergonomics and Safety I: Taylor and Francis.*, *1*, 481–485.

Campbell, B. J., Zegeer, C. V., Huang, H. H., & Cynecki, M. J. (2004). A review of pedestrian safety research in the united states and abroad. https://popcenter.asu.edu/sites/default/files/problems/pedestrian_injuries/PDFs/campbell_etal_2004.pdf

Chen, C., Hong, J., & Wang, S. (2015). Automated positioning of 3d virtual scene in ar-based assembly and disassembly guiding system. *The International Journal of Advanced Manufacturing Technology*, *76*(5), 753–764. https://doi.org/10.1007/s00170-014-6321-6

ConsumerLab, E. (2017). 10 hot consumer trends 2017. https://www.ericsson.com/en/reports-and-papers/consumerlab/reports/10-hot-consumer-trends-2017

David, K., & Flach, A. (2010). Car-2-x and pedestrian safety. *IEEE Vehicular Technology Magazine*, *5*(1), 70–76. https://doi.org/10.1109/MVT.2009.935536

De Clercq, K., Dietrich, A., Núñez Velasco, J., De Winter, J., & Happee, R. (2019). External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. *Human factors*, *61*(8), 1353–1370. https://doi.org/10.1177%2F0018720819836343

De Winter, J., Leeuwen, P., & Happee, R. (2012). Advantages and disadvantages of driving simulators: A discussion.

Dey, D., & Terken, J. (2017). Pedestrian interaction with vehicles: Roles of explicit and implicit communication. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 109–113. https://doi.org/10.1145/3122986.3123009

Directive 2010/40/EU. (2010). *The framework for the deployment of intelligent transport systems in the field of road transport and for interfaces with other modes of transport.* European Parliament, Council of the European Union. http://data.europa.eu/eli/dir/2010/40/oj

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

Flores, A., & Belongie, S. (2010). Removing pedestrians from google street view images. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 53–58. https://doi.org/10.1109/CVPRW.2010.5543255

Hart, S., & Staveland, L. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Human mental workload* (pp. 139–183). https://doi.org/10.1016/S0166-4115(08)62386-9

Henderson, S., & Feiner, S. (2011). Exploring the benefits of augmented reality documentation for maintenance and repair. *IEEE Transactions on Visualization and Computer Graphics*, *17*(10), 1355–1368. https://doi.org/10.1109/TVCG.2010.245

Hile, H., Liu, A., Borriello, G., Grzeszczuk, R., Vedantham, R., & Košecka, J. (2010). Visual navigation for mobile devices. *IEEE MultiMedia*, *17*(2), 16–25. https://doi.org/10.1109/MMUL.2010.5

Imamov, S., Monzel, D., & Lages, W. (2020). Where to display? how interface position affects comfort and task switching time on glanceable interfaces, 851–858. https://doi.org/10.1109/VR46266.2020.00110

Jerald, J. (2015). The vr book: Human-centered design for virtual reality. https://doi.org/10.1145/2792790.2792802

Kaleefathullah, A., Merat, N., Lee, Y., Eisma, Y., Madigan, R., Garcia de Pedro, J., & De Winter, J. (2020). External human-machine interfaces can be misleading: An examination of trust development and misuse in a cave-based pedestrian simulation environment. https://doi.org/10.1177/0018720820970751

Keller, C., & Gavrila, D. (2014). Will the pedestrian cross? a study on pedestrian path prediction. *Intelligent Transportation Systems, IEEE Transactions on*, *15*, 494–506. https://doi.org/10.1109/TITS.2013.2280766

Kidwell, E. (2018). How long in vr is too long? https://www.androidcentral.com/how-long-vr-too-long

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

Lindemann, P., & Rigoll, G. (2017). Examining the impact of see-through cockpits on driving performance in a mixed reality prototype. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct*, 83–87. https://doi.org/10.1145/3131726.3131754

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

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

Mann, S., & Fung, J. (2001). Videoorbits on eye tap devices for deliberately diminished reality or altering the visual perception of rigid planar patches of a real world scene. *EYE*, *3*, P3.

Microsoft. (2020). Mixed reality - comfort. https://docs.microsoft.com/en-us/windows/mixed-reality/design/comfort

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies*, *2351*. https://doi.org/10.1117/12.197321

Mok, C., Bazilinskyy, P., & De Winter, J. (2021). Stopping by looking: A driver-pedestrian interaction study in a coupled simulator using head-mounted displays with eye-tracking.

Moore, D., Currano, R., Strack, G., & 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. https://doi.org/10.1145/3342197.3345320

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

National Highway Traffic Safety Administration. (2020). Traffic safety facts: 2018 data. pedestrians. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812850

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

Palffy, A., Kooij, J., & Gavrila, D. (2019). Occlusion aware sensor fusion for early crossing pedestrian detection, 1768–1774. https://doi.org/10.1109/IVS.2019.8814065

Plopski, A., Taylor, A., Carter, E., & Admoni, H. (2019). InvisibleRobot: Facilitating robot manipulation through diminished reality. *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 165–166. https://doi.org/10.1109/ISMAR-Adjunct.2019.00-55

Rameau, F., Ha, H., Joo, K., Choi, J., Park, K., & Kweon, I. (2016). A real-time augmented reality system to see-through cars. *IEEE Transactions on Visualization and Computer Graphics*, *22*(11), 2395–2404. https://doi.org/10.1109/TVCG.2016.2593768

Renner, P., & Pfeiffer, T. (2017). Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. *2017 IEEE symposium on 3D user interfaces (3DUI)*, 186–194. https://doi.org/10.1109/3DUI.2017.7893338

Rolland, J., & Hua, H. (2005). Head-mounted display systems. *Encyclopedia of optical engineering*, *2*. https://doi.org/10.1081/E-EOE-120009801

Roser, M., Appel, C., & Ritchie, H. (2013). Human height. *Our World in Data*. https://ourworldindata.org/human-height

Sakai, S., Yanase, Y., Matayoshi, Y., & Inami, M. (2018). D-ball: Virtualized sports in diminished reality. *Proceedings of the First Superhuman Sports Design Challenge: First International Symposium on Amplifying Capabilities and Competing in Mixed Realities*. https://doi.org/10.1145/3210299.3210305

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. *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, 25–30. 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, 313–316. 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. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 63–72. https://doi.org/10.1109/VR46266.2020.00024

Schubert, T. (2003). The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Z. für Medienpsychologie*, *15*(2), 69–71. https://doi.org/:10.1026//1617-6383.15.2.69

Schwind, V., Knierim, P., Haas, N., & Henze, N. (2019). Using presence questionnaires in virtual reality. https://doi.org/10.1145/3290605.3300590

Siltanen, S. (2017). Diminished reality for augmented reality interior design. *The Visual Computer*, *33*(2), 193–208. https://doi.org/10.1007/s00371-015-1174-z

Singer, T., Stuckert, A., Erkan, A., Hoffmann, D., & Khanh, T. (2021). Behaviour of automated vehicles in critical situations with pedestrians - accident prevention through additional light signals.

Smith, M., Gabbard, J. L., & Conley, C. (2016). Head-up vs. head-down displays: Examining traditional methods of display assessment while driving. *Proceedings of the 8th international conference on Automotive User Interfaces and Interactive Vehicular Applications*, 185–192. https://doi.org/10.1145/3003715.3005419

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

Tabone, W., De Winter, J., Ackermann, C., Bärgman, J., Baumann, M., Deb, S., Emmenegger, C., Habibovic, A., Hagenzieker, M., Hancock, P., et al. (2021a). Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. *Transportation Research Interdisciplinary Perspectives*, *9*, 100293. https://doi.org/10.1016/j.trip.2020.100293

Tabone, W., Lee, Y. M., Merat, N., Happee, R., & De Winter, J. (2021b). Towards future pedestrian-vehicle interactions: Introducing theoretically-supported ar prototypes, 209–218. 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., 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), 1–10. https://doi.org/10.1016/S0968-090X(96)00025-3

Vasconcelos-Raposo, J., Bessa, M., Melo, M., Barbosa, L., Rodrigues, R., Teixeira, C., Cabral, L., & Sousa, A. (2016). Adaptation and validation of the igroup presence questionnaire (ipq) in a portuguese sample. *Presence Teleoperators & Virtual Environments*, *25*(3), 191–203. https://doi.org/10.1162/PRES_a_00261

Waldin, N., Waldner, M., & Viola, I. (2017). Flicker observer effect: Guiding attention through high frequency flicker in images. *Computer Graphics Forum*, *36*(2), 467–476. https://doi.org/10.1111/cgf.13141

Witmer, B., Jerome, C., & Singer, M. (2005). The factor structure of the presence questionnaire. *Presence: Teleoperators & Virtual Environments*, *14*(3), 298–312. https://doi.org/10.1162/105474605323384654

Won, M., Shrestha, A., Park, K., & 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. (2021). Road traffic injuries. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

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

Zadeh, R., Ghatee, M., & Eftekhari, H. (2017). Three-phases smartphone-based warning system to protect vulnerable road users under fuzzy conditions. *IEEE Transactions on Intelligent Transportation Systems*, *19*(7), 2086–2098. https://doi.org/10.1109/TITS.2017.2743709

Zegeer, C. V., & Bushell, M. (2012). Pedestrian crash trends and potential countermeasures from around the world [Safety and Mobility of Vulnerable Road Users: Pedestrians, Bicyclists, and Motorcyclists]. *Accident Analysis & Prevention*, *44*(1), 3–11. https://doi.org/10.1016/j.aap.2010.12.007

Zhang, B., Wilschut, E., Willemsen, D., Alkim, T., & Martens, M. (2018). The effect of see-through truck on driver monitoring patterns and responses to critical events in truck platooning. *Advances in Human Aspects of Transportation*, *597*, 842–852. https://doi.org/10.1007/978-3-319-60441-1_81

Ziebinski, A., Cupek, R., Erdogan, H., & Waechter, S. (2016). A survey of adas technologies for the future perspective of sensor fusion. *International Conference on Computational Collective Intelligence*, 135–146. https://doi.org/10.1007/978-3-319-45246-3_13

A

Top-down overview of the virtual environment

Fig. A.1 provides measurements relative to pedestrian location and the locations of the AV and parked vehicle. The AV in this instance is at the full stop location of the yielding scenario. The AV drives in the centre of the 5 m wide lane, about 5 m orthogonally away from the pedestrian on the curb.

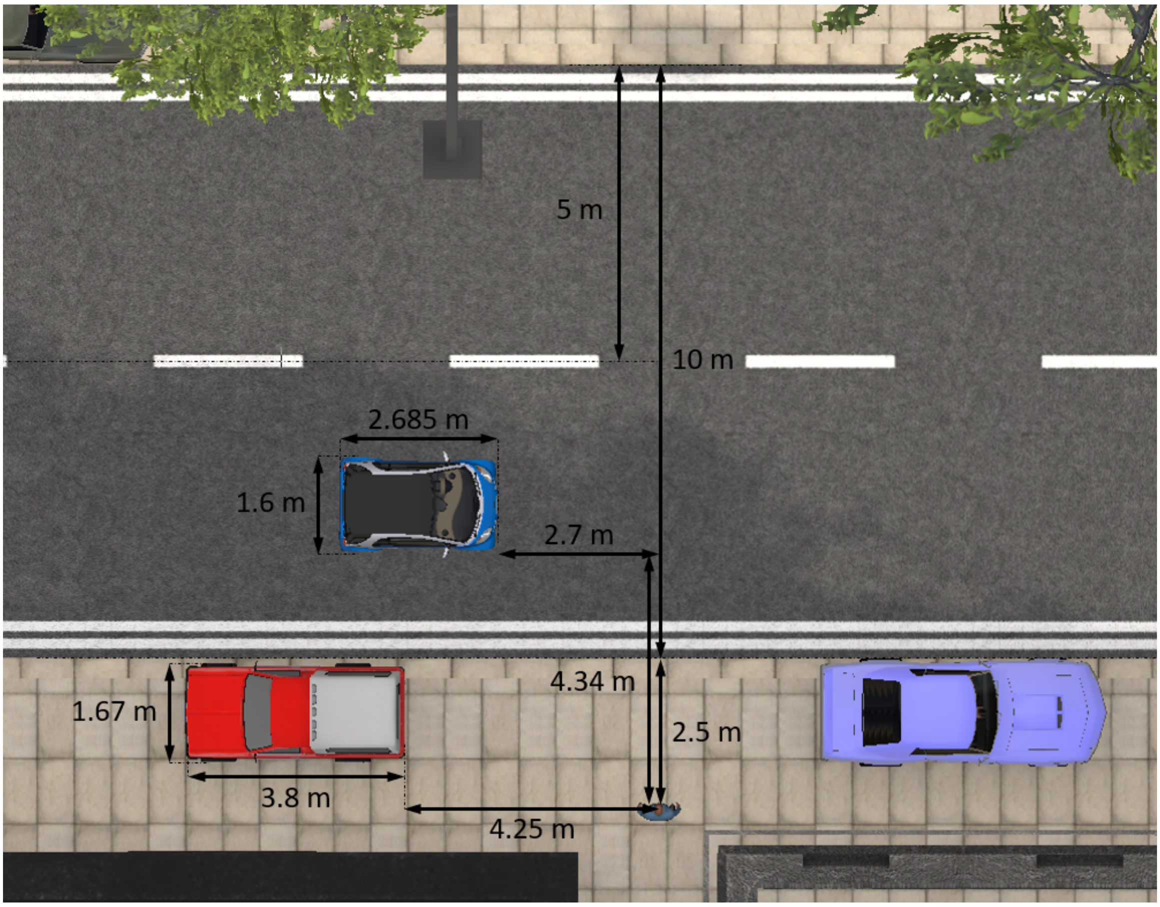


Fig. A.1: Top-down overview of the virtual environment

Note that these measurements were taken manually within Unity, and may be prone to small errors. Therefore, Fig. A.1 serves for illustrative purposes only. The distances used for calculations in this thesis were computed by taking the Euclidean norm of the pedestrian location and the AV car's centre, as these positions were recorded automatically by the logging script.

B
Prototypes

In this appendix, high resolution pictures of the prototype interfaces are shown from the perspective of the participant, i.e., HMD user. A demo video of the interfaces in the non-yielding trials can be found here: https://dropbox.com/

Fig. B.1: Baseline prototype

Fig. B.2: See-through display prototype

Fig. B.3: Transparent car prototype

Fig. B.4: Head-locked display prototype

Fig. B.5: Head-locked display prototype, with view guidance

Fig. B.6: Body-locked display prototype

Fig. B.7: Body-locked display prototype, with view guidance