Assessment of Driver Attention during a Safety Critical Situation in VR to Generate VR-based Training

Efe Bozkir  
Perception Engineering  
University of Tübingen  
efe.bozkir@uni-tuebingen.de

David Geisler  
Perception Engineering  
University of Tübingen  
david.geisler@uni-tuebingen.de

Enkelejda Kasneci  
Perception Engineering  
University of Tübingen  
enkelejda.kasneci@uni-tuebingen.de

ABSTRACT
Crashes involving pedestrians on urban roads can be fatal. In order to prevent such crashes and provide safer driving experience, adaptive pedestrian warning cues can help to detect risky pedestrians. However, it is difficult to test such systems in the wild, and train drivers using these systems in safety critical situations. This work investigates whether low-cost virtual reality (VR) setups, along with gaze-aware warning cues, could be used for driver training by analyzing driver attention during an unexpected pedestrian crossing on an urban road. Our analyses show significant differences in distances to crossing pedestrians, pupil diameters, and driver accelerator inputs when the warning cues were provided. Overall, there is a strong indication that VR and Head-Mounted-Displays (HMDs) could be used for generating attention increasing driver training packages for safety critical situations.

CCS CONCEPTS
• Human-centered computing → Empirical studies in HCI.  
• Computing methodologies → Virtual reality; Simulation environments; Perception.

KEYWORDS
Virtual Reality, Eye Tracking, Driver Attention, Driver Assistance, Driver Training, Pedestrian Safety

ACM Reference Format:

1 INTRODUCTION
Having safe driving experiences and decreasing the number of crashes are two of the most important issues when it comes to driving safety. Every year, many fatal crashes occur on roads all over the world. According to the Road Safety Annual Report in International Transport Forum 2018, most of the fatal crashes occurred on rural roads; however, the number of fatal crashes in urban roads has been increasing in more than half of the countries since 2000 [IRTAD 2018].

Apart from road or weather conditions, distracted driving can cause fatal crashes. While a total prevention is almost impossible, many crashes can be prevented by training drivers using driver assistant systems. With recent developments in the field of augmented reality (AR) and head-up display (HUD) technology, new means have become available to overlay different warnings to the driver, such as pedestrian warnings or road signs. In fact, many modern cars already employ this technology to a certain degree. The majority of studies that concentrated on driver training and the interaction between these technologies and drivers in safety critical situations used driving simulators. With the recent developments in VR and HMDs, it is possible to apply these scenarios and trainings in VR with lower cost. However, it is an open question whether VR and HMDs can be used in studying driver training and interaction for safety critical situations.

In order to assess whether VR, HMDs, and gaze-aware cues can be useful and driver attention can be increased properly in this context, we focused on an unexpected pedestrian crossing behavior at non-designated crosswalks on urban roads when the Time-to-Collision (TTC) between the driving vehicle and crossing pedestrian is very short (≈ 1.8-5 seconds). [Rasouli et al. 2017] mentioned that in this range of TTC, there is a high likelihood that joint attention between crossing pedestrian and driver happens. However, in case it does not happen, due to distracted pedestrian or driver, it is more likely that a crash will happen. In our experiments, control group did not receive any critical pedestrian warning cues, whereas the experimental group had the gaze-aware critical pedestrian warning cues. By analyzing closest distances between driving vehicles and crossing pedestrians, pupil diameter changes of drivers between baseline and risky driving timeframes, and driver performance measurements, we found that there is a strong indication that gaze-aware visual warnings for critical pedestrians help increasing the driver attention earlier in VR. Therefore, low-cost VR setups along with realistic and gaze-aware warnings can be introduced to train drivers for safety critical scenarios. Major contributions of our work are as follows: (a) Demonstrating a very critical scenario in terms of collision risk between driver and pedestrian with and without risky pedestrian warning cues in VR and (b) Evaluation of gaze-aware critical pedestrian warning cues in VR whether they increase driver attention earlier so that attention increasing VR-based training packages can be proposed and further evaluated. Since the dedicated driving scenario is highly dynamic and time-critical, the outcome of the current study can be taken as a basis for any study that includes time-dependent and safety-critical scenarios in VR.
2 RELATED WORK

Driving simulation studies have been conducted in various domains. Two of the most common issues addressed were safety and driver assistance. [Charissis and Papanastasiou 2010] introduced a novel interface for HUD over head-down display (HDD). HUDs and AR cues have been used for various purposes. [Schwarz and Fastenmeier 2017] discussed that specificity of visual warnings provided advantages in gaze, brake reaction times, passing speeds, and collision rates. [Tran et al. 2013] showed the benefits of HUDs while turning left, whereas [Rusch et al. 2014] presented positive effects of AR cues in terms of time-to-contact and gap response variation to assist elderly drivers during left-turns. In addition, [Bark et al. 2014] showed the navigational AR aid for recognizing turn locations earlier via 3D volumetric HUD. [Dijkstra et al. 2012] discussed that adaptive support in HUD for lane keeping helped drivers drive more centrally and with less lateral variation. The effect of in-car AR system for reducing collisions caused by other vehicles’ movements was presented by [Fu et al. 2013]. Additionally, increase in situational awareness using AR in automated driving for take-over scenarios was studied by [Lorenz et al. 2014] and [Langlois and Soualmi 2016], whereas classification of drivers’ take-over readiness was studied by [Braunagel et al. 2017].

While numerous studies can be counted in the context of driver assistance, the studies include pedestrian safety, hazard anticipation, and driver training are more relevant to our work. [Rusch et al. 2013] showed that AR cueing increased the response rate for pedestrian and warning sign detection in directing driving attention to roadside hazards. The study of [Pomarjanschi et al. 2012] in a driving simulator with a maximum speed of about 30km/h showed that gaze guidance reduced number of pedestrian collisions. [Phan et al. 2016] studied three driver awareness levels of a pedestrian in a driving simulator: Perception, vigilance, and anticipation. They showed that AR cues were capable of enhancing the driver awareness in all levels. The outdoor study conducted by [Kim et al. 2018] showed that AR pedestrian warnings provided positive results on measures such as braking, distances to pedestrians, and gaze-on pedestrian travel distances. The study of [Prahdhan et al. 2005] on eye movements showed that the scanning patterns of novice drivers reflected their failure to recognize potential risks. Driving simulator studies have been used in driver training and VR as well. [Roenker et al. 2003] found out that drivers who were trained in a simulator improved their driving skills in turning into correct lane and proper signal use. Furthermore, [Fisher et al. 2007] evaluated hazard anticipation and found that trained drivers recognized the risks more often. [Lang et al. 2018] showed the effect of improvement of bad driving habits via synthesizing personalized training programs in VR. [Mangalore et al. 2019] assessed drivers’ hazard anticipation across VR and driving simulators to evaluate the usage of VR headsets and justified that VR headsets could be used for measuring driving performance. [Ju et al. 2019] studied personality traits on sacrifice decisions including pedestrians during VR-based driving. While the studies which include driving simulators and hazardous situations showed great potential for driver training, it is an open question whether visual cues for critical situations in VR can increase driver attention properly, so that VR-based training packages can be proposed and synthesized for safety critical situations.

3 EXPERIMENT

We focused on driver behavior in a very critical scenario when pedestrians tried to cross the road with TTC was between ≈ 1.8-5 seconds in VR. In this range of TTC, there is a high likelihood that pedestrian or joint attention occurs [Rasouli et al. 2017]. However, if it does not occur, the outcome can be fatal. Our experiment included a control group that did not receive any cues, and an experimental group that received gaze-aware critical pedestrian cues. Our major hypothesis is that if the gaze-aware warning cues can successfully increase the driver attention earlier in the safety critical situation in VR, similar low-cost VR setups along with adaptive warnings could be proposed for driver training for these situations.

3.1 Participants

16 volunteer participants (4 female, 12 male) whose ages range from 25 to 50 (M = 31) and driving experiences range from 5 to 30 years (M = 12) participated in the experiment. Participants were separated into two groups. A control group, receiving no critical pedestrian warning cues, and an experimental group, receiving the warning cues.

3.2 Apparatus

HTC-Vive along with Pupil-Labs Binocular Add-on [Kassner et al. 2014], which has binocular 120hz eye tracking cameras and clip-on rings, Logitech G27 Steering Wheel and Pedals, and Phillips headphones were used to create driving setup. Eye tracking was measured using the open-source hmd-eyes of Pupil-Labs with Pupil Service version 1.7. Virtual city was created using Unity3D game engine. For the environment, vehicles, and pedestrians, we purchased and used models from Urban City Pack, City Park Exterior Props, Traffic Sign Set, Modern People, Traffic Cars, Realistic Car HD 02, Realistic Car Controller, Simple Waypoint System, and Playmaker asset packages. We designed the main roads long and straight so that the drivers would have opportunity to speed up as they want and drive naturally. Example scenes from our virtual environment are shown in Figure 1.

(a) Cockpit of Driving Car
(b) Main Road

Figure 1: Example Scenes from VR Environment

The dedicated setups were run on a PC equipped with an NVIDIA Titan X graphics card with 12GB memory, a 3.4GHz Intel i7-6700 processor, and 16GB of RAM.

Since the visual warning cues for experimental group are very important in our setup, Figure 2 shows a pedestrian model with and without warning cues.

3.3 Procedure

In the beginning of the experiment, participants were informed about the purpose and scope of the experiment orally. They had
the opportunity to stop and cancel the experiment anytime. At the end of the experiment, participants filled a small questionnaire about demographic and qualitative information. The experiment consisted of two phases. For both phases, participants were given written instructions before starting. In the first phase, participants acclimated the setup. This phase did not include any pedestrians or dynamic objects apart from the driver’s car; no data were collected during this phase. Generally, this phase lasted in 5-10 minutes, although if participants had not felt comfortable, they could have continued driving. Once they felt comfortable with the setup, they continued to the second phase.

In the beginning of the second phase, 2D calibration with 16 points using hmd-eyes of Pupil-Labs was performed. After calibration success, participants started the experiment. The starting location of the driving vehicle was in the beginning of the main road, where a critical pedestrian crossing happened. Since there was no intersection until the end of this road, all of the participants were required to drive until the end. At the end of the road, they could have turned left or right and continued driving, however our data analyses did not concentrate on the data acquired after the turn, since they could have encountered with different scenarios. The speed limit of the driven road was 90km/h, and participants were supposed to realize this by traffic signs. The driving vehicle was also equipped with maximum speed warning.

The critically crossing pedestrian scenario was as follows. At the beginning of each run, two occurrences of a critical pedestrian were generated along with other non-critical pedestrians on the side walks. The critically crossing pedestrian was determined at random, as active and proceeded to dangerously cross the street before the driving vehicle. Both of these occurrences had dedicated gaze-aware warning cues. Pedestrians were not located in the beginning of the road, so that the drivers had the opportunity to speed-up or slow-down until the crossing. Pedestrian warnings were activated for the experimental group when the distance between front of the driving vehicle and critical pedestrians became ≈ 77m. The crossing pedestrian started crossing the road from the right side when the distance between vehicle and pedestrian was \( d_{critical} \approx 45m \). We assumed that drivers would obey the speed limit (90km/h) and also drive faster than 30km/h. This way, parameter of \( d_{critical} \) helps to map expected TTC to \( 1.8s \leq TTC \leq 5s \) interval. Ray-casting \cite{Roth1982} method was used to map gaze signal, which was obtained from Pupil-Labs software, from 2D canvas to 3D environment by the help of Unity3D colliders \cite{Unity2019} that were attached to virtual objects. Once the drivers’ gaze signal in 3D environment was closer than 5 meters to the pedestrians for \( \approx 0.85 \) seconds, the cues were deactivated. Therefore, the cues became gaze-aware. Since the control group did not receive cues, the timeframes consisted of different milestones for each group. \( t_w \) and \( t_m \) correspond to start of the critical pedestrian warning and start of the pedestrian movement respectively. For the control group, baseline driving corresponds to \( [t_m - \delta t, t_m] \), whereas for the experimental group, it is \( [t_w - \delta t, t_w] \). \( [t_m, t_m + \delta t] \) is the risky driving timeframe for both groups. Setting different values of \( \delta t \) means changing the durations of the timeframes.

### 3.4 Measurements

The metrics analyzed were the closest distances between the crossing pedestrians and the driving vehicles, driver performance measurements including inputs on accelerator and brake pedals, and pupil diameter changes between baseline and risky driving timeframes. Particularly, since the critically crossing pedestrian is only safety critical for the driving vehicle inside of the driven lane, we took the closest distance in this lane. Driver inputs on pedals are also indicators of perception and reflect the smoothness of the driving experience as well. Lastly, pupil enlargement corresponds to increase in cognitive load \cite{Appel2018}. Pupil diameter values were fetched from Pupil-Labs software in pixel units. For smoothing and normalization, we applied Savitzky-Golay filter \cite{Savitzky1964}, and divisive baseline correction using baseline duration of 0.5 seconds and median \cite{Mathot2018}.

### 3.5 Hypotheses

Our hypotheses are based on the driver attention and actions. Since the experimental group was provided with the risky pedestrian cues, we expected that the closest distances between the crossing pedestrians and the driving vehicles for the experimental group would be more than the control group. In addition, when the visual cues were provided to the drivers, we expected that they would understand the criticality earlier, and their cognitive load would increase earlier. Pupil dilation is one of the indicators of the cognitive load increase, therefore we expected that pupil dilation of the experimental group would happen earlier. Furthermore, cues would affect driver inputs on accelerator and brake pedals, hence it was expected that experimental group drivers would take their foot off the accelerator earlier and perform smoother braking behavior than the control group drivers. In all, we expected that the experimental group would perform safer and smoother driving experience than the control group.

### 4 RESULTS

Analyses for the distances, driver performance measurements and pupil diameters during baseline driving and risky driving timeframes were calculated using MATLAB and are as follows.

#### 4.1 Closest Distance to Crossing Pedestrian

We measured the closest distances between the crossing pedestrians and driving vehicles until pedestrians completed half of their trajectories, since during the second half, the pedestrians were not safety critical to the driving vehicles anymore. Figure 3 shows the results for this metric. We applied two sample T-test with alpha level of 0.05 and found significant difference between two groups with \( p \approx 0.00059 \) (Cohen’s \( d \approx 2.21 \)). One of the participants in...
the control group hit the pedestrian, and the experiment was terminated at that moment. In addition, since the velocities of the vehicles were not fixed, the difference in distances could vary. However, the deceleration trend in the experimental group started from $t_{br}$, which is a strong indication that they acknowledged the critical situation earlier than the control group and behaved accordingly. Overall, it is clear that the experimental group participants drove safer than the control group participants.

**Figure 3: Closest Distance to Crossing Pedestrian - Experiment Group Relationship**

### 4.2 Driver Performance Measurements

Driver inputs on accelerator and brake pedals are the two main indicators of safe and smooth driving. Therefore, we analyzed the normalized driver inputs on accelerator and brake during different durations of baseline and risky driving timeframes. First, we applied paired T-test with alpha level of 0.05 between baseline and risky driving timeframes using normalized mean accelerator inputs. As expected, significant differences for experimental group even for very short durations (e.g. $\delta t \approx 50\text{ms}, p = 0.0158, \text{Cohen's } d \approx 1.12$) were found. However, significant differences were found for the control group starting from $\delta t \approx 1.4s (p = 0.0495, \text{Cohen's } d \approx 0.84$). Figure 4 shows the dedicated analyses. Finding significant differences in shorter $\delta t$ values means that the drivers acknowledged the critical situation earlier. Therefore, it is a significant indicator that visual pedestrian cues helped drivers drive safely even during a very dangerous situation. Furthermore, we analyzed braking behaviors by analyzing whether participants performed full brake, since the braking happens in very short time. In total, five of the participants in the control group performed full brake, whereas none of the participants in the experimental group did this. This indicates that visual cues also helped to have smoother driving experience.

**Figure 4: Accelerator Inputs - Driving Timeframe Relationship**

### 4.3 Pupil Diameter

Since pupil dilation is one of the indicators of cognitive load increase, we analyzed normalized pupil diameters of the drivers in the same way as accelerator inputs between baseline and risky timeframes using paired T-test with alpha level of 0.05. Since HMDs and VR offer controlled illumination, we expected that pupil dilation would happen due to the increase in cognitive load, and pupil diameters of the experimental group would increase earlier than the control group. Analyses showed that significant difference in pupil diameters between baseline and risky timeframes for the experimental group starts from $\delta t \approx 1.4s (p = 0.048, \text{Cohen's } d \approx 0.85)$, whereas it starts from $\delta t \approx 2.4s (p = 0.0489, \text{Cohen's } d \approx 0.84)$ for the control group. Figure 5 shows the results. Overall, there is a strong indication that cues for the critical pedestrian increased cognitive load of the experimental group earlier so that they behaved accordingly.

**Figure 5: Pupil Diameter - Driving Timeframe Relationship**

### 5 CONCLUSION

We introduced a VR driving simulation environment and a safety critical pedestrian crossing to study whether VR setups and gaze-aware cues can increase driver attention in critical situations despite the prevalent disadvantages, such as narrow field-of-view, low resolution or weight of HMDs, so that low-cost VR-based training for safety critical situations can be proposed and further evaluated. To the best of our knowledge, this is the first work that assesses VR setups using gaze-aware cues for safety critical situations in driving by analyzing eye tracking and performance metrics. We found significant differences in the distances to crossing pedestrians, accelerator inputs, and pupil diameters between baseline and risky timeframes. Results indicate that driver attention can be increased earlier with minimalistic gaze-aware cues properly in safety critical situations in VR. Most of the previous work on driving simulation and training were done using physical driving simulators. However, VR setups can decrease cost of implementation and time. Overall, we suggest that driver attention increasing training packages can be introduced in VR. Since many modern cars have different warnings for safety critical situations, VR could be used to assess these systems and train people to get acclimated with them as well.

As future work, detailed eye-tracking analyses, a study to generate better attention grabbing cues, and a driver training study for critical situations to assess whether drivers improve their bad driving habits by VR-based training can be done.
Assessment of Driver Attention in VR for Driver Training

SAP '19, September 19–20, 2019, Barcelona, Spain

REFERENCES


Karlin Bark, Cuong Tran, Kikuo Fujimura, and Victor Ng-Thow-Hing. 2014. Personal Navi: Benefits of an Augmented Reality Navigational Aid Using a See-Thru 3D Volumetric HUD. https://doi.org/10.1145/2582001


Michelle Rusch, Mark Schall, John Lee, Jeffrey Dawson, and Matthew Rizzo. 2014. Augmented reality cues to assist older drivers with gap estimation for left-turns. Accident Analysis & Prevention 71 (01 2014), 210ÂŞ221. https://doi.org/10.1016/j.aap.2014.05.020


