
Driving Assessment Conference

2019 Driving Assessment Conference

Jun 27th, 12:00 AM

Can Virtual Reality Headsets be Used to Measure Accurately Drivers' Anticipatory Behaviors?

Ganesh Pai Mangalore
University of Massachusetts Amherst

Yalda Ebadi
University of Massachusetts Amherst

Siby Samuel
University of Waterloo

Michael Knodler
University of Massachusetts Amherst

Donald Fisher
University of Massachusetts Amherst

Follow this and additional works at: <https://ir.uiowa.edu/drivingassessment>

Mangalore, Ganesh Pai; Ebadi, Yalda; Samuel, Siby; Knodler, Michael; and Fisher, Donald. Can Virtual Reality Headsets be Used to Measure Accurately Drivers' Anticipatory Behaviors?. In: Proceedings of the Tenth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, 24-27 June 2019, Santa Fe, New Mexico. Iowa City, IA: Public Policy Center, of Iowa, 2019: 342-348. <https://doi.org/10.17077/drivingassessment.1716>

This Event is brought to you for free and open access by the Public Policy Center at Iowa Research Online. It has been accepted for inclusion in Driving Assessment Conference by an authorized administrator of Iowa Research Online. For more information, please contact lib-ir@uiowa.edu.

CAN VIRTUAL REALITY HEADSETS BE USED TO MEASURE ACCURATELY DRIVERS' ANTICIPATORY BEHAVIORS?

Ganesh Pai Mangalore¹, Yalda Ebadi¹, Siby Samuel², Michael A. Knodler¹, Donald L. Fisher¹

¹University of Massachusetts Amherst, Amherst, MA, USA

²University of Waterloo, Waterloo, Ontario, Canada

Email: gpaimangalor@umass.edu

Summary: VR headsets are several orders of magnitude less expensive than driving simulators. Their use in research and clinical settings could explode were it shown that the results obtained with VR headsets were similar to those obtained with more standard driving simulators. Towards this end, the current study expands on a previous initial validation study of VR headsets. In particular, it has been shown in conventional driving simulation and on-road studies that middle-aged drivers glance longer at latent hazards than their younger counterparts. In this study the total time middle-aged drivers spend glancing at a latent hazard and the average duration of each glance were compared to these same times for younger drivers using a VR headset and fixed-based driving simulator. The results indicate that the middle-aged participants glanced longer than their younger counterparts on both platforms at latent hazards, as measured by the total glance duration but had no difference when measured by the average glance duration. Moreover, the magnitude of the difference between middle-aged and younger drivers was the same across the two platforms. These results are in line with previous simulator studies. There appears here a real opportunity to expand the powers of simulation using VR headsets, both for purposes of research and clinical practice.

INTRODUCTION

Various studies in transportation engineering, human factors and psychology have made use of driving simulators due their ability to simulate real world scenarios with a high sense of immersion without putting the driver at risk (Slob, 2008). With the recent penetration of virtual reality (VR) headsets in the market, they have been increasingly used for research and educational purposes (Lei et al, 2018). Due to their higher level of immersion (Johnston et al, 2018), when compared to traditional driving simulators, VR headset-based driving simulators may greatly expand the range of driving simulation studies. But first they need to be validated and, in particular, it needs to be shown that with regards to critical safety tasks, the vehicle, driver and eye behaviors recorded on a VR headset are identical to what is found on driving simulator.

In a recent study, we evaluated the use of VR headsets to measure driver's hazard anticipation behavior in comparison to a fixed-based driving simulator (Mangalore et al, 2019). Hazard anticipation behavior is important to examine because it has been linked to crashes in numerous studies (Horswill & McKenna, 2004). The study showed that the difference in the likelihood that young and middle-aged drivers glanced at latent hazards (Crundall & Pradhan, 2016) was analogous across the two platforms indicating that virtual reality headsets were as effective as driving simulators, at least in the measurement of latent hazard anticipation (Mangalore et al, 2019). We refer to this and related measures as the spatial characteristics of latent hazard anticipation glances (where the driver looked) to distinguish it from the temporal characteristics (for how long the driver looked). Importantly, the finding that younger drivers were less likely to

detect latent hazards than middle-aged drivers is consistent with past studies (Pradhan et al, 2005).

The aim of the current study is to continue and expand on the results of Mangalore et al. (2019) and provide further evidence to suggest the validity of VR headsets as a platform for using driving simulation to analyzed driver behaviors. In particular, in this study our focus is on measuring the temporal characteristics of the glance behaviors of drivers when traveling past latent hazards. The temporal characteristics include both the total time the driver spends glancing at a latent hazard and the duration of each glance at a latent hazard. It is important to know how long in total drivers glance at a latent hazard because drivers who look for only a short total period of time or who take very short glances are less likely to be able fully to perceive a threat, understand what the threat means, and take appropriate action (Endsley, 1995). With regard to temporal characteristics, it has been reported in previous simulator studies that middle-aged drivers spend longer in total looking at latent hazards than their younger counterparts (Urwyler et al, 2015; Crundall et al, 2012). As for the duration of individuals glances, it has been reported that as measured on a driving simulator or using video clips there are only marginally significant differences in the average glance durations of middle-aged and younger drivers (Chapman & Underwood, 1998; Chan et al, 2010). For this reason, we have considered both total glance duration and average glance duration as our dependent variables in this study. The current study aims to determine whether these two temporal characteristics (the total duration of the glances at a latent hazard and the average glance duration of each glance at a latent hazard) of young drivers and more experienced, middle-aged drivers are similar on two platforms: a VR headset-based driving simulator and a fixed-based driving simulator. If the differences between the results acquired on both platforms are small, this will further add to the evidence that VR headsets can be used to measure indices of safe driving behavior.

METHOD

Participants

A total of 50 participants (24 young drivers aged 18-21 years; 24 middle-aged drivers aged 30-55 years; and two drop-outs during the preliminary practice drive due to simulator sickness) were recruited for this study from the UMass Amherst local area. Half of the young and middle-aged drivers were randomly assigned either to a fixed-based driving simulator or a VR headset-based driving simulator resulting in four groups of drivers, each consisting of 12 drivers. The population by gender, average age and average driving experience and their respective standard deviation are listed in Table 1. There were no statistically significant differences in the average age or driving experience of the two young simulator groups or the two middle-aged simulator groups.

Table 1. Sample Characteristics

Driver Group	Age (Years)		Driving Experience (Years)		Population by Gender	
	Average	SD	Average	SD	Male	Female
Middle-Aged Simulator	38.17	7.5369	18.1522	9.6691	7	5
Young Simulator	20.25	0.8292	3.1433	1.2005	9	3
Middle-Aged Headset	39.58	8.7983	21.0142	7.5496	8	4
Young Headset	20.08	0.9538	2.6692	1.1415	8	4

Fixed-based Driving Simulator and Eye Tracker

Simulation & Controls. The Realtime Technologies (RTI) fixed-based driving simulator consists of a fully equipped 2013 Ford Fusion placed in front of five screens with 330-degree field of

view. The cab features two dynamic side-mirrors and a rear-view mirror which provide rear views of the scenarios for the participants. The simulator is equipped with a five speaker surround system for exterior noise and a two-speaker system for simulating in-vehicle noise. *Eye-Tracking Device.* The Applied Science Laboratories (ASL) MobileEye is used for eye tracking on the fixed-based driving simulator. Calibration is conducted using a 9-point calibration screen. Eye movements are recorded at 30 Hz refresh rate with an accuracy of 0.5 degrees of visual angle.

VR Headset-Based Driving Simulator

Simulation & Eye-Tracking Device. The Tobii Pro Integrated HTC Vive VR headset is a head-mounted display (HMD) which is integrated with Tobii Eye Tracking. It provides a 110 field-of-view with a display resolution of 1080×1200 at a 90 Hz refresh rate. The eye tracker provides eye-tracking with an accuracy 0.5 degrees of visual angle at a 120 Hz refresh rate.

Controls. The Logitech G29 Driving Force features a powerful dual-motor force feedback to simulate the force effects required for an accurate response from the driver, along with good steering control. It also consists of a floor unit with integrated throttle, brake, and clutch pedals.

Scenarios

Unity 3D and SimCreator were used to create identical scenarios for, respectively, the VR headset-based driving simulator and fixed-based driving simulator respectively. The signage, lane markings and speed limits were similar on both platforms. The scenarios are briefly described in Table 2. The latent threats never actually materialized in the scenarios in order not to oversensitize the participants.

Table 2. Scenario Descriptions (The latent threats are in bold.)

Scenario	Scenario Description (The driver is referred to as ‘D’)
Right Turn	D approaches a stop-sign controlled four-way intersection with a travel lane in either direction. The driver is expected to turn right at the intersection. There is a crosswalk at the intersection and a pedestrian approaching the crosswalk is obscured by a block of buildings on the right.
Left-Turn Truck	D approaches a four-way intersection with two travel lanes in either direction, with cross traffic controlled by stop signs. In the left lane, a truck is attempting to make a left turn. The truck blocks the driver’s view of any oncoming traffic from the opposing lanes.
Obscured Crosswalk	There is a truck parked on the right side of a two-lane roadway right before a crosswalk. As D nears the truck and tries to pass from its left side, a vehicle approaches in the opposing lane .
Pedestrian Island	D is in the right lane while approaching a T-intersection. Only the stem of the T is controlled by a stop sign. In the left lane, a line of vehicles waits to turn left. The median to the left of the line accommodates a pedestrian island at the crosswalk. A pedestrian on this island is obscured by the line of vehicles
Obscuring Vegetation	D is approaching a stop sign controlled T-intersection with one travel lane in either direction. There is a pedestrian at the crosswalk which lies further beyond the intersection to the driver’s right side. Vegetation obscures the stop sign and also the driver’s view of the crosswalk.
Stop Ahead	D is travelling on a road curving to the right and approaching a stop sign controlled intersection for the driver (but not cross traffic). At the beginning of the curve, a Stop Ahead sign exists and the Stop Sign at the end of the curve is partially obscured by vegetation.

Driver Behavior Questionnaire

This study utilizes the North American version of the Driver Behavior Questionnaire (DBQ) which was originally developed in the United Kingdom. DBQ is a widely used tool to measure driving behaviors linked to collision risks (Reason et al, 1990). Three subscales were identified for the 24 items listed in the questionnaire in the form of questions, namely, Error (E), Lapses

(L), and Violations (V). Each participant rated an item on a scale of 0 to 5 (rarely to always), based on how often they engaged in the behavior mentioned in that item. For example, “Try to pass another car that is signaling a left turn” is an Error related item and a participant who rarely engages in such behavior would rate this item as ‘0’. Each DBQ has 8 items for each subscale. In our experiment we computed the average score for each subscale based on each participants’ responses for each of the driving experience groups (Young and Middle-Aged) on both platforms.

Experimental Design

The experimental design was a 2 x 2 mixed design with platform (fixed-based driving simulator or VR headset-based driving simulator) and driving experience (young or middle-aged) as the two between-subject factors and scenario as the within subject factor. The order of the scenarios was counterbalanced across and within groups using a balanced Latin Square method (0 1949).

Procedure

After informed consent was obtained from the participants, a Pre-Study questionnaire and a Driver Behavior Questionnaire were administered to record data related to demographics, driver experience, and drivers’ tendency to engage in aggressive behavior while driving. Next the participants were outfitted with an eye-tracker and calibration was done to ensure accurate eye-tracking data. The participants on the VR headset-based driving simulator were given a short tutorial on different aspects of the headset and steering wheel. Participants were given basic instructions such as to follow on-screen/audio instructions and maintain the posted speed limit. Then they were asked to drive through a preliminary practice drive for the next five minutes, after which, a set of six counterbalanced scenarios were introduced to the participants with a gap of 30 seconds between loading each scenario. The study lasted 45 minutes on average.

Dependent Variables

In each of the recorded videos, a set of ‘launch zones’ and ‘target zones’ were generated for each scenario (Samuel & Fisher, 2015). A *target zone* is defined as an area(s) of the roadway from where potential threats may emerge. A *launch zone* is defined as that area of the roadway where the drivers should scan at least once towards the target zone to successfully identify the presence of any potential threats. The term *glance* in this experiment is used to refer to one or more sequential fixations on the target zone when the participant is in the launch zone in a particular scenario. Each frame includes an indication of where the driver is looking in the frame. In a frame-by-frame tracking of the recorded videos (one frame = 33 milliseconds), every sequence of frames in which the driver is looking at the target zone from the launch zone is recorded as a glance. A participant usually makes more than one glance in the scenario where he or she successfully detected the latent hazards. The *total glance duration* is the sum of the duration of all glances made by a participant in a scenario at a latent hazard, while the *average glance duration* is the mean duration of all glances at the latent hazard.

RESULTS

In our previous study, we analyzed the latent hazard anticipation behavior of the drivers based on a binary scoring method (‘1’ – successful detection; ‘0’ – unsuccessful detection). In this experiment we considered only those glances that were scored as ‘1’ and calculated the total glance duration and average glance duration of all gazes towards the target zone made by each participant for each of the scenarios.

Analysis of Anticipatory Glances. A 2×2 factorial [2 driving experience groups: Young & Middle-aged; 2 Platforms: VR headset-based simulator and fixed-based driving simulator] ANOVA was performed separately for the total glance duration and average glance duration for each scenario for each participant, $n = 48$, $\alpha = 0.05$. Analysis of the total glance duration indicated no main effect of platform ($F = 2.309$; p -value = 0.130; $\eta^2 = 0.010$) or interaction between experience and platform ($F = 2.733$; p -value = 0.1; $\eta^2 = 0.012$). There was a main effect of driving experience ($F = 19.9$; p -value < 0.005; $\eta^2 = 0.084$). These results are consistent with those reported in previous studies which state that middle-aged drivers' glance in total at potential threats longer than their younger counterparts (Urwyler et al, 2015; Crundall et al, 2012). For average glance duration, there was no interaction between experience and platform ($F = 0.042$; p -value = 0.838; $\eta^2 = 0.0002$) or main effect of platform ($F = 3.42$; p -value = 0.066; $\eta^2 = 0.015$) or of driving experience ($F = 3.429$; p -value = 0.065; $\eta^2 = 0.015$). These results are consistent with those reported in previous studies which state that the difference in the average glance duration between younger and middle-aged drivers is marginal or non-existent (Chapman & Underwood, 1998; Chan et al, 2010).

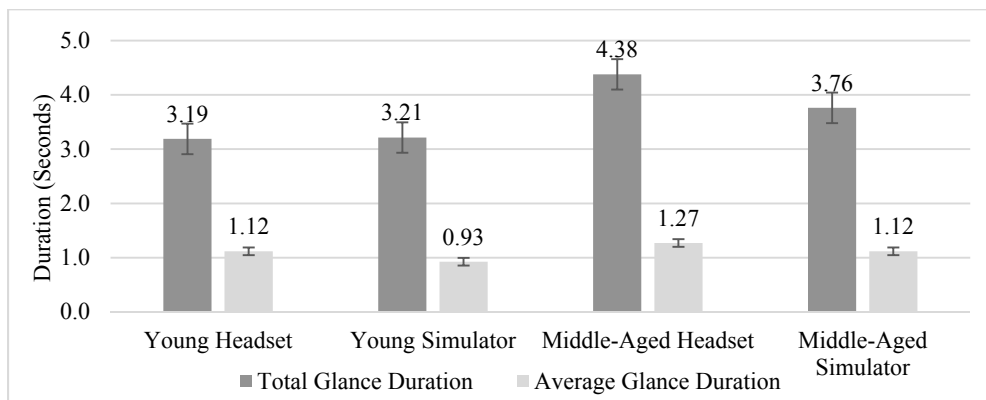


Figure 1. The mean average glance duration and mean total glance duration for each driver group

Driver Behavior Questionnaire. A 2×2 factorial [2 driving experience groups: Young and Middle-aged; 2 Platforms: VR headset-based simulator and fixed-based driving simulator] ANOVA was performed for the average scores for Error, $n = 48$, $\alpha = 0.05$. No interaction between experience and platform ($F = 0.322$; p -value = 0.573; $\eta^2 = 0.005$) or main effect of platform ($F = 0.158$; p -value = 0.394; $\eta^2 = 0.017$) or of driving experience ($F = 0.055$; p -value = 0.816; $\eta^2 = 0.001$) were observed. Similar results were observed for the average scores for Lapse and Violation.

CONCLUSION

The current study seeks to add more evidence in support of using VR headsets to measure driver performance (total glance duration and average glance duration of anticipatory glances) in safety critical tasks where normally a fixed-based driving simulator might be used to do such. In particular, the results showed that middle-aged drivers spent a longer time glancing at latent hazards than did young drivers on both the VR headset-based and fixed based driving simulators. With this in mind, it is also important to note that the average glance duration was the same among young and middle-aged drivers across both platforms. Had the middle-aged drivers' average glance duration at the latent hazards been longer than those of younger drivers, the middle-aged drivers would potentially have compromised their safety. Both results are in line with results from previous research conducted on driving simulators and on-road studies that

demonstrated that while middle-aged drivers gaze longer at latent hazards, i.e., have a longer total glance duration, there may only be marginal or no differences in terms of their average glance duration when compared to the younger drivers. Most importantly, the results from the ANOVA models for total glance duration and average glance duration showed that there was no impact of platform on performance for either the young or the middle-aged drivers.

With every between-subject design, there exists a possibility for certain confounds to arise, such as, the overrepresentation in one group of drivers who tend to engage in aggressive, aberrant driving behavior. In order to determine whether such confounds were present, a Driver Behavior Questionnaire was administered in the study. Results show no indication of such confounds with no significant effect in questionnaire responses across all platforms and driving experience groups.

The study has some important limitations as noted here. The current study utilized a between-design experiment to test the performance of drivers on a VR-based driving simulator and fixed-based driving simulator. Maintaining complete homogeneity is difficult despite random assignment. Other measures of driving performance may also be considered for validation of a platform (e.g., various vehicle measures such as the lane deviation, speed etc.) in future studies.

To summarize our results, the current study provides more evidence that VR headsets can be used to effectively measure driver performance, specifically the temporal characteristics of anticipatory glance behaviors. Together with the results from the previous study showing that the spatial characteristics of the glance behaviors of drivers in the presence of latent hazards can be measured as effectively on VR headsets as they can on driving simulators, it suggests that VR headsets can potentially be used to measure a wide range of safety critical behaviors, not only hazard anticipation behaviors. Such additional behaviors are known to include hazard mitigation behaviors as well as attention maintenance behaviors (Fisher et al, 2017). The previous study also showed that VR headsets, like driving simulators, did not to generate more than minimal simulator sickness (Mangalore et al, 2019). The range of applications in which VR headset-based driving simulators could now be employed is greatly expanded. They could be used to study scenarios which are going to require the participation of multiple drivers, for example, scenarios in which each driver was using different levels of automation. They could be used for training novice drivers or older drivers on a widespread basis, something that is not possible with more expensive fixed-based driving simulators. They could be used during licensure to evaluate drivers crash avoidance skills. The opportunities are many and the impact could potentially be equally large.

REFERENCES

- Chan, E., Pradhan, A. K., Pollatsek, A., Knodler, M. A., & Fisher, D. L. (2010). Are driving simulators effective tools for evaluating novice drivers' hazard anticipation, speed management, and attention maintenance skills? *Transportation research part F: traffic psychology and behaviour*, 13(5), 343-353.
- Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, 27(8), 951-964.

- Crundall, D., Chapman, P., Trawley, S., Collins, L., Van Loon, E., Andrews, B., & Underwood, G. (2012). Some hazards are more attractive than others: Drivers of varying experience respond differently to different types of hazard. *Accident Analysis & Prevention*, 45, 600-609.
- Crundall, D., & Pradhan, A. (2016). Hazard Avoidance: Definitions and a framework for research in novice drivers' skills in dealing with roadway hazards. Handbook of Teen and Novice Drivers (Fisher, D., Caird, J., Horrey, W., Trick, L. eds.), CRC Press, Boca Raton, 61-74.
- Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human factors*, 37(1), 65-84.
- Fisher, D. L., Young, J., Zhang, L., Knodler, M., & Samuel, S. (2017). Accelerating Teen Driver Learning: Anywhere, Anytime Training. *Behavior Research Methods*, 37, 379-384.
- Horswill, M. S., & McKenna, F. P. Drivers' hazard perception ability: Situation awareness on the road. In *A Cognitive Approach to Situation Awareness* (S. Banbury & S. Tremblay, eds), Aldershot, UK, Ashgate, 2004, pp.155-175.
- Johnston, A. P., Rae, J., Ariotti, N., Bailey, B., Lilja, A., Webb, R., ... & McGhee, J. Journey to the centre of the cell: Virtual reality immersion into scientific data. *Traffic*, (2018). 19(2): 105-110.
- Lei, X., Zhang, A., Wang, B., & Rau, P. L. P. Can Virtual Reality Help Children Learn Mathematics Better? The Application of VR Headset in Children's Discipline Education. In *International Conference on Cross-Cultural Design*, 2018. pp. 60-69.
- Mangalore, G.P., Ebadi, Y., Samuel, S., Knodler, M., & Fisher, D.L. (2019, Accepted). Can Virtual Reality Headsets Be Used to Measure Accurately Drivers' Hazard Anticipation Performance: The Promise of VR Headsets. *Transportation Research Record: Journal of the Transportation Research Board*.
- Pradhan, A. K., Hammel, K. R., DeRamus, R., Pollatsek, A., Noyce, D. A., & Fisher, D. L. Using eye movements to evaluate effects of driver age on risk perception in a driving simulator. *Human factors*, 2005. 47(4): 840-852.
- Reason, J., Manstead, A., Stradling, S., Baxter, J., & Campbell, K. (1990). Errors and violations on the roads: a real distinction?. *Ergonomics*, 33(10-11), 1315-1332.
- Samuel, S., & Fisher, D. L. Evaluation of the minimum forward roadway glance duration. *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2518: 9-17.
- Slob, J.J. State-of-the-art Driving Simulators, A Literature Survey. DCT Report. Eindhoven University of Technology, Department Mechanical Engineering, Control Systems Technology Group, 2008. 107: 1-19.
- Urwyler, P., Gruber, N., Müri, R. M., Jäger, M., Bieri, R., Nyffeler, T., ... & Nef, T. (2015). Age-dependent visual exploration during simulated day-and night driving on a motorway: a cross-sectional study. *BMC geriatrics*, 15(1), 18.
- Williams, E. J. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 1949. 2(2): 149-168.