October 2019

The Promise of VR Headsets: Validation of a Virtual Reality Headset-Based Driving Simulator for Measuring Drivers’ Hazard Anticipation Performance

Ganesh Pai Mangalore
University of Massachusetts Amherst

THE PROMISE OF VR HEADSETS: VALIDATION OF A VIRTUAL REALITY HEADSET-BASED DRIVING SIMULATOR FOR MEASURING DRIVERS' HAZARD ANTICIPATION PERFORMANCE

A Thesis Presented

by

GANESH PAI MANGALORE

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE IN MECHANICAL ENGINEERING

September 2019

Mechanical and Industrial Engineering

THE PROMISE OF VR HEADSETS: VALIDATION OF A VIRTUAL REALITY HEADSET-BASED DRIVING SIMULATOR FOR MEASURING DRIVERS' HAZARD ANTICIPATION PERFORMANCE

A Thesis Presented

by

GANESH PAI MANGALORE

Approved as to style and content by:

____
Siby Samuel, Chair

____
Michael A. Knodler, Member

____
Shannon C. Roberts, Member

____
Sundar Krishnamurty, Department Head
Mechanical & Industrial Engineering

ACKNOWLEDGMENTS

This research work was carried out within the University of Massachusetts Amherst in collaboration with Arbella Insurance Human Performance Lab (HPL) and UMass Transportation Center (UMTC).

First and foremost, I would like to thank my advisor, Professor Siby Samuel for his invaluable input, ideas, and support during the course of this thesis research. I remember walking into his office one day to inquire about any available assistantship positions and this thesis is the result of his goodwill and generosity to offer me a project that day.

I would like to thank Professor Michael Knodler for his continued belief in this project and his everlasting support during the past two years. Prof. Knodler made sure that I always got the latest tools necessary to execute this thesis research. I would also like to thank Professor Shannon Roberts for giving me the opportunity to use the driving simulator at the Human Performance Laboratory. Prof. Roberts' suggestions during the proposal of this project were extremely helpful.

I would like to extend my gratitude to Professor Donald Fisher who was instrumental in many ways by providing me many useful tips and suggestions while drafting my thesis and the subsequent research papers. I would like to thank my friend and colleague, Yalda Ebadi, for helping me throughout the entire project with a wide range of tasks whilst providing excellent suggestions with a great sense of humor.

Lastly, I would like to thank my mom and dad - Roopa Pai and Prashanth Pai, as well as my grandmother, Nirmala Shenoy, for supporting my decision to study abroad and tirelessly motivating me during the entirety of my master's program.

ABSTRACT

THE PROMISE OF VR HEADSETS: VALIDATION OF A VIRTUAL REALITY HEADSET-BASED DRIVING SIMULATOR FOR MEASURING DRIVERS' HAZARD ANTICIPATION PERFORMANCE

SEPTEMBER 2019

GANESH PAI MANGALORE

B.E., N. M. A. M. INSTITUTE OF TECHNOLOGY, NITTE

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Siby Samuel

The objective of the current study is to evaluate the use of virtual reality (VR) headsets to measure driving performance. This is desirable because they are several orders of magnitude less expensive and, if validated, could greatly extend the powers of simulation. Out of several possible measures of performance that could be considered for evaluating VR headsets, the current study specifically examines drivers' latent hazard anticipation behavior both because it has been linked to crashes and because it has been shown to be significantly poorer in young drivers compared to their experienced counterparts in traditional driving simulators and in open road studies. The total time middle-aged drivers spend glancing at a latent hazard and the average duration of each glance was also compared to these same times for younger drivers using a VR headset and fixed-based driving simulator. In a between-subject design, forty-eight participants were equally and randomly assigned to one out of four experimental conditions – two young driver cohorts (18 – 21 years) and two middle-aged driver cohorts (30 – 55 years) navigating either a fixed-based driving simulator or a VR-headset-based simulator. All participants navigated six unique scenarios while their eyes were continually tracked. The proportion of latent hazards anticipated by participants which constituted the primary dependent measure was found to be greater for middle-aged drivers than young drivers across both platforms. Results also 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 difference in the magnitude of performance between middle-aged and younger drivers was the same across the two platforms. There were also no significant differences found for the severity of simulator sickness symptoms across the two platforms. The study provides some justification for the use of virtual reality headsets as a way of understanding drivers' hazard anticipation behavior.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. BACKGROUND

1.1. Overview on Driving Simulators and Virtual Reality Headsets

Over the years, driving simulators have been extensively used for various transportation, human factors, and behavioral studies (Slob, 2008). Their increased level of safety and ability to simulate real-world scenarios with a high sense of immersion have made them useful tools for studying drivers' behavior and performance in low- and high-risk scenarios, to evaluate alternative in-vehicle interface designs, and to conceptualize and design training programs (Lee et al, 2001; Godley et al, 2002; Roenker et al, 2003). The realism in the simulation of these virtual environments is particularly useful as the simulator tests can be used as a precursor to open road evaluations, thereby minimizing research expenditures and increasing the level of safety (Winn, 1999; Velev & Zlateva, 2017).

In the past few years, the market has been saturated with a wide variety of VR headsets such as Oculus Rift, Nintendo Wii U VR, and HTC Vive, among others, which have been used for research, training and educational purposes (Pulijala et al, 2018; Oagaz et al, 2018; Lei et al, 2018). The ambiguity of a 3D environment is eliminated in VR headsets and the true experience of that 3D environment which cannot be achieved in non-VR headset platform is possible (DeLuca & Deluca, 2003; Marks et al, 2014). VR headsets allow the user to experience virtual worlds with higher resolution graphical quality, regulated visual flow with a high sense of realism when compared to environments presented on conventional driving simulators. Additionally, VR headsets have more flexibility and portability which is not the case with most driving simulators.

Using Virtual Reality headsets, users can use the engaging, immersive virtual worlds to learn rich and complex content while enhancing their technical, creative and problem-solving skills (Burns, 2012). By executing optimized, intelligent designs with systematic delivery, a user can grasp more complex concepts (Darken & Silbert, 1996). This makes training programs aimed at drivers and pilots to be greatly enhanced by introduction of Virtual Reality, by not only making it possible to measure participants' behavioral responses more effectively, but by also making considerable cost reductions on infrastructure, equipment and their accompanying technical support (McComas et al, 2002; Velev & Zlateva, 2017). VR headsets can also be used to review certain expensive designs and concepts more effectively, for example, by combining 3D models along with VR headsets, an architect or contractor can walk through a simulated virtual space of a structural design before the expensive real-life construction of that structure begins (Hilfert & König, 2016). However, VR headsets do have their own disadvantages. VR headsets are known to cause a phenomenon called the 'Screen door effect' which can be described as a black grid over the original image while displaying a virtual world. The Oculus Rift headset, when worn close to the eyes of the user, has been known to cause a screen door effect. It is unclear whether the VR headset (HTC Vive) used in this research causes this phenomenon. Ghosting is another phenomenon where faded trails appear behind moving objects. This has again been detected during the use of the Oculus Rift (Desai et al, 2014). Prolonged use of VR headsets could also cause physical discomfort which may affect the user's experience of the virtual environment. This may lead to the user developing a negative attitude towards VR use in general. It should not be assumed that physical ergonomics are simply due to the poor design of VR peripherals, since VR peripherals are developing fast, although it is worth noting that sophisticated models may not be cost effective (Nichols, 1999). In a study utilizing the Oculus Rift virtual reality headset, simulator sickness was a strong factor in modulating people's gaming experiences using the Rift, though it was found that simulator sickness did not always significantly diminish the participants' immersive experiences. With that in mind, it is pivotal to consider the effect of simulator sickness during the development of virtual worlds (Tan et al, 2015).

In the driving safety research domain, such headsets have been used to train hazard anticipation behavior in young drivers (Agrawal et al, 2018). The aim of the current study is to perform an initial validation of VR headsets as a platform for driving simulation since they offer better immersion (Johnston et al, 2018), additional portability, and much lower costs while maintaining the level of safety provided by traditional simulators. As such they could greatly extend the use of simulators in science and engineering, possibly making the study of 100s of drivers in mixed traffic environments a real possibility. However, at least two concerns stand in the way. First, there is a lack of documented research that specifically examines the ability of these headsets to measure driving performance and do so as well as traditional fixed-based driving simulators. Second, there is a concern that VR headsets can lead to simulator sickness (Munafo et al, 2017; Jensen & Konradsen, 2018).

1.2. Hazard Anticipation

With respect to the first concern about VR headsets, several aspects of driver performance could have been examined for such a validation study. In this experiment, we specifically focus on a higher order cognitive skill – latent hazard anticipation. In the literature, hazard anticipation is described as a collection of driver behavioral attributes such as the awareness and knowledge of traffic risks, the ability to scan and understand hazardous situations which may result in crashes, the ability to anticipate latent hazards from the current field of view, and finally the capacity to adopt the necessary actions to safely navigate the roadway by mitigating risks (Vlakveld, 2011; McDonald et al, 2015). Researchers have learned that it is important to differentiate between hazards that are visible and those that are not visible or have not materialized but can easily be anticipated (Borowsky et al, 2013). This is perhaps best understood using examples: An example of a visible hazard is a vehicle in the opposing lane crossing over into the driver's lane. An example of a hazard that is not visible, but can be anticipated, is a pedestrian in a crosswalk hidden by a stopped vehicle in a travel lane. An example of a hazard that has not materialized, consider a vehicle driving through a residential area, on a two-lane roadway with a hidden driveway on the right side. The driveway is obscured by vegetation and any potential hazard coming onto the road from the driveway is also obscured. To minimize any potential conflicts, the driver would need to identify and continuously scan the driveway for any potential hazards that may emerge until safely passing through that area of the roadway (Mehranian, 2013).

There are two reasons we focus on hazard anticipation. On the one hand, there is a consensus that young, novice drivers lack the ability to acquire and assess information relevant to the recognition of risks on the road ahead (Mayhew & Simpson, 1995; Fisher et al, 2002; Lee et al, 2008; Romoser et al, 2013). A driving simulator study by Pradhan et al. (2005), reported that while 69.59% of older, experienced drivers engaged in behaviors indicative of successful latent hazard detection in the scenarios, only 25.82% of the younger, inexperienced drivers and 40.14% of the younger, experienced drivers depicted such behaviors. In summary, hazard anticipation has been shown repeatedly to be significantly poorer in young drivers than more experienced drivers (Pradhan et al, 2005), and therefore can serve as a standard for comparing the performance of VR headsets with other measures of latent hazard anticipation.

On the other hand, the inability to detect latent hazards has been linked to the increased rate of crashes (Horswill & McKenna, 2004; Thomas et al, 2016), making it one of the more critical skills with which to assess VR headsets. In one study, it was reported that out of 1000 crashes reviewed, inexperience and failure to scan for hazards were the main factors contributing to approximately 42.7% of the crashes (McKnight & McKnight, 2003). It was argued that this was due for the most part to the fact that younger drivers are generally inexperienced rather than that they have an increased risk-taking tendency.

To begin the validation of the VR platform for driving simulation purposes, it is vital to replicate results previously validated on another platform. A fixed-based driving simulator was chosen for comparison due to similarities in the manner of simulation and possibility of performance measurement. To validate the VR platform, we will compare the hazard anticipation performance of young and more experienced, middle-aged drivers on the VR headset-based driving simulator and a fixed-based driving simulator. The scenarios used in Pradhan et al. (2005) were redeveloped on a VR headset using Unity 3D, to the closest identifiable approximation. By comparing the two platforms we will determine whether there is a difference in the proportion of latent hazards anticipated by young drivers on the VR headset and fixed-based simulator and correspondingly, whether there is a difference between middle-aged drivers on the two simulator platforms. If the differences are small, this will be important evidence that VR headsets can be used to measure one of the most critical of behaviors, latent hazard anticipation.

1.3. Glance Duration

Glance duration refers to the temporal characteristics (for how long the driver looked) as opposed to the spatial characteristics of latent hazard anticipation glances (where the driver looked) mentioned in the previous section. 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. To validate the VR platform, 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 will also be compared between 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.

1.4. Simulator Sickness

With respect to the second concern about VR headsets, simulator sickness, we gave participants the Simulator Sickness Questionnaire (Kennedy et al, 19930). If VR headsets when used to evaluate hazard anticipation create increased rates of simulator sickness, then the differences should appear in the scores of the VR headset groups when compared with the fixed-base simulator groups.

Simulator sickness is a major obstacle to the use of driving simulators for research, training and driver assessment purposes. Due to a large amount of visual flow associated with virtual environments, visual-temporal lags occur resulting in Simulator Sickness. There is limited scientific literature as to what influences Simulator Sickness and its subsequent effect on the behavior and performance of the user in the virtual environment. Factors such as age, sex, and psychological traits, etc. which increase the likelihood of simulator sickness have been identified. Other factors such as those related to various elements of the virtual environment (curved roads, high speeds, long durations) and those related to the technical setup of the simulator (controls, delay in response) have also been recognized (Classen et al., 2011; Milleville-Pennel & Charron, 2015).

In the past, driving simulation and human factors researchers have employed several measures to limit the problem of simulator sickness. These include various pre-experimental screening questions during the recruitment stages regarding history with motion sickness and preliminary practice drivers to identify and exclude subjects prone to simulator sickness. Despite these measures, it has is seemingly impossible to rule out the chances of a participant experiencing simulator sickness during simulation studies. (Brooks et al.,2010).

In order to validate the VR platform, it is vital to determine whether there is a difference in the simulator sickness questionnaire scores between corresponding driver groups on the two simulator platforms. If the differences are small, this will further help establish VR headsets as a feasible platform for future driving simulation studies.

1.5. Objective of the Thesis

To sum up, the objective of this study is to validate the VR headset-based driving simulator for the following measures: binary-coded hazard anticipation (looked vs not looked), total glance duration (how long did they glance after initial detection) and average glance duration (how long did each of their glances last after initial detection). The results for these variables obtained from participants on the VR headset-based simulator will be compared to those obtained from the fixed-based driving simulator. We hypothesize that these results will identical on both platforms and in-line with past findings, effectively validating the VR platform for measuring anticipatory eye-movements in driving simulation studies. Additionally, we also hypothesize that the symptoms of simulator sickness as calculated from the Simulator Sickness questionnaire will be similar, and that the VR platform will not generate simulator sickness symptoms any more than the fixed-based driving simulator. The methodology and procedures carried out to meet these objectives will be detailed and explained in the following section.

2. METHODOLOGY

2.1. Participant Groups

The study recruited a total of 48 participants, which included 24 young drivers aged 18-21 years; 24 middle-aged drivers aged 30-55 years. There were two drop-outs during the preliminary practice drive due to simulator sickness which were not included in the sample size. For the 48 participants who completed the practice drive without any symptoms of simulator sickness, 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. This resulted in four total groups of drivers, with each group consisting of 12 drivers: young simulator, middle-aged simulator, young headset, and middle-aged headset. The average age and average driving experience of the participants along with their respective standard deviation are listed group wise in Table 2.1. The participant sample according to gender has also been listed in Table 2.1.

Table 2.1. 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

There were no statistically significant differences in the mean ages or years of driving experience of two young simulator groups or the two middle-aged simulator groups. All participants held a valid United States drivers' license, were recruited from the University of Massachusetts Amherst local area and were remunerated for their participation. Due to the difficulty posed by eyeglasses during eye-tracking calibration, participants with eyeglasses were excluded from the study. There were no other inclusion or exclusion criteria in this study.

2.2. Apparatus and Software

The apparatus consists of a fixed-based driving simulator, an eye tracker, a VR headset and vehicle controls. The primary software consists of various programs to create the virtual worlds and coordinate events in these worlds. These are described in more detail below and the differences between the two simulator platforms have been listed in Table 2.2.

2.2.1. Fixed-based Driving Simulator and Eye Tracker

1) RTI Driving Simulator: The Realtime Technologies (RTI) fixed-based driving simulator at the UMass Amherst Arbella Insurance Human Performance Laboratory consists of a fully equipped 2013 Ford Fusion placed in front of five screens with 330-degree field of view (Realtime Technologies Catalog, 2018). The five front and side surrounding screens have a display resolution of 1900 x 1200 dpi, with the sixth rear screen having a resolution of 1400 x 1050 (Figure 2.1.). The cab also 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. All aspects of the simulator are monitored and coordinated on SimCreator which is a PC-based program that launches, controls and collects real-time data from every simulator drive. The scenarios for the driving simulator are designed and developed using software called Internet Scene Assembler (ISA) which contain various commonly used roadway and environmental assets (roads, intersections, buildings, trees, etc.) as well as a user-friendly interface which helps coordinate scripted events in scenarios such as the appearance of a pedestrian at a certain distance from the driver's vehicle.

Figure 2.1. RTI Driving Simulator

2) ASL MobileEye: The Applied Science Laboratories (ASL) MobileEye is a monocular eye tracker consisting of a pair of goggles with one camera focused on the eye, another focused on the scene ahead, and a small reflective monocle for the eye camera to view the eye without obstructing the participant's view (Figure 2.2.). Calibration is conducted using a 9-point calibration screen. Eye movements are recorded at a 30 Hz refresh rate and the gaze cursor is overlaid on the recorded video output. The eye tracker has an accuracy of 0.5 degrees of visual angle. It is used for eye tracking on the fixed-based driving simulator.

Figure 2.2. ASL MobileEye

2.2.2. VR Headset-Based Driving Simulator

The VR Headset-Based Driving Simulator consists of the Tobii Pro Integrated HTC Vive connected to a Logitech G29 Driving Force steering wheel. Unity 3D was used to initiate and execute scenarios. A typical scenario in this experiment would display a virtual avatar of a generic driver with hands on the steering wheel, seated inside a standard sedan class automobile. The virtual cab consisted of shifters, pedals, steering wheel and side/rear view mirrors similar to cabs in the real world. This gave the driver an immersive feel of being seated in an actual car. If the participant moved the steering wheel in the real world, the avatar would also move their hands similarly in the virtual world. Below are individual components of the VR Headset-Based Driving Simulator briefly explained.

1) Tobii Pro Integrated HTC Vive: This virtual reality headset is a retrofitted version of the HTC Vive Business Edition head-mounted display (HMD) which is integrated with Tobii Eye Tracking (Tobii VR Integration, 2018). The headset provides a 110 field-of-view with a display resolution of 1080×1200 at a 90 Hz refresh rate. The eye tracking platform uses the Binocular Dark Pupil Tracking technique (Morimoto & Mimica, 2005) to track the pupil and uses a five-point calibration method to provide eye-tracking with up to 0.5o of visual error at a 120 Hz refresh rate (Figure 2.3., left panel).

2) Logitech G29 Driving Force: The steering wheel features a powerful dual-motor force feedback to simulate the force effects required for an accurate response from the driver, along with good steering action. The 900-degree lock-to-lock rotation enables the wheel to be rotated two and a half times. It also consists of a separate floor pedal unit with integrated throttle, brake, and clutch pedals (Figure 2.3., right panel).

Figure 2.3. Tobii Pro Integrated HTC Vive (Left); Logitech G29 Driving Force (Right*)*

3) Unity 3D: Unity is an all-purpose game engine that supports 2D and 3D graphics, drag and drop functionality and scripting through C# (Figure 2.4.). In this study, the Unity 3D engine was used to create graphically-pleasing, realistic environments featuring several on-road elements and hazards. Assets for the various on-road and environmental elements (such as trees, signage, vehicles, etc.) featured in the scenarios were mostly designed from scratch or imported from numerous resources on the Unity Store.

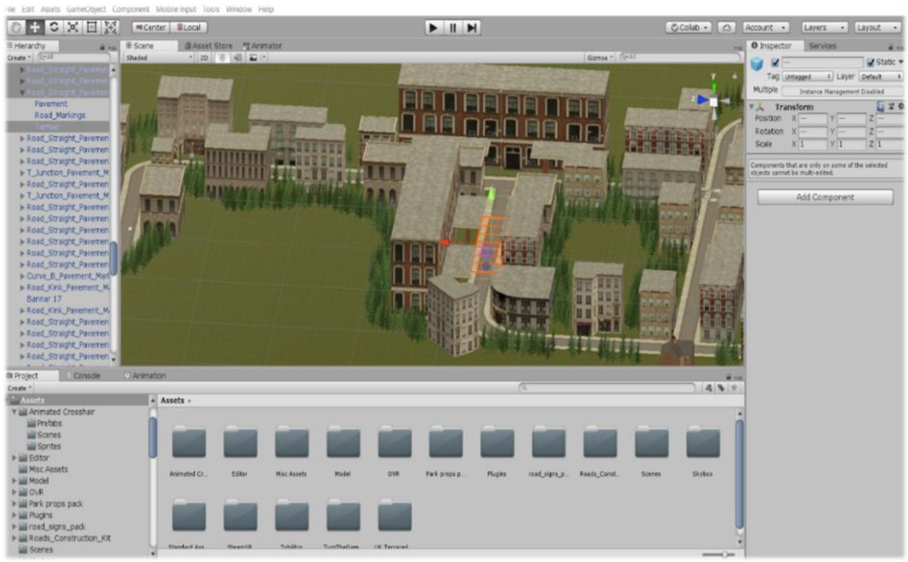


Figure 2.4. Designing virtual worlds using Unity 3D

Table 2.2. Differences between the two simulator/eye tracking platforms

	Fixed-Based RTI Driving Simulator and ASL MobileEye	VR Headset-Based Driving Simulator
Fidelity	High	Low
Vehicle Measures Output	Speed, Lane Deviation, Steering wheel offset, Acceleration, etc.	None, but can be programmed to collect the desired output
Eye-tracking Output	Gaze point, Gaze direction, Blink Rate, Horizontal & Vertical Dispersion	Gaze point and direction are available by default. Other features are programmable.
Eye-tracking Refresh Rate	30 Hz, monocular tracking	120 Hz, binocular tracking
Field of View	330 degrees (Fixed)	110 degrees (Relative to the user's head position)

2.3. Experimental Scenarios

Using Unity 3D and SimCreator, 6 unique scenarios were designed respectively for the VR headset-based driving simulator and the fixed-based driving simulator respectively, in order to examine the driver’s ability to anticipate latent hazards. The design and layout of roadways as well as the latent hazard zones featured in these scenarios were identical on both platforms. The signage, traffic control, and lane markings were similar on both platforms. The six scenarios were similar to those used in the Risk Awareness and Perception Training (RAPT) program which were also evaluated in Pradhan et al. (2005). The posted speed limit for the 'Right Turn', 'Obscured Crosswalk' and 'Obscuring Vegetation' scenarios was 30 mph, while the posted speed limit for the 'Left Turning Truck', 'Pedestrian Island' and 'Stop Ahead' scenarios was 45mph. The scenarios have been listed in Table 2.3. and Table 2.4.

Table 2.3. Descriptions and Plan Views for Scenarios 1-3 (Note: Driver is the red car)

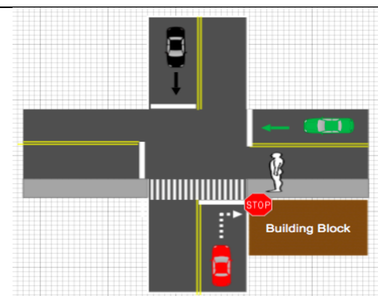
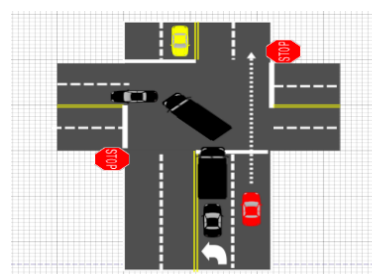
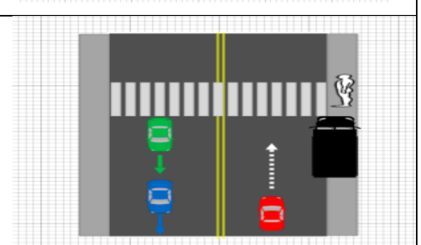
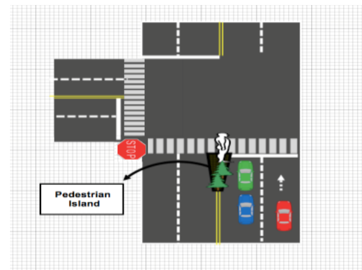
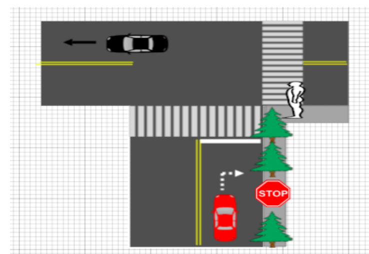
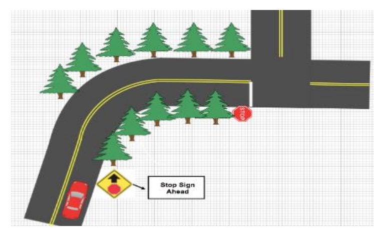
Scenario Description	Required Action	Plan View
1. Right Turn: The driver approaches a stop-sign controlled four-way intersection with a travel lane in either direction. The driver is expected to turn right at the intersection. There is a crosswalk at the intersection and a pedestrian approaching the crosswalk is obscured by a block of buildings on the right.	The driver should scan the obscured area on the right before reaching the intersection to detect any hazards that may arise from the area or to yield to pedestrians that may attempt to cross at the crosswalk.	
2. Left Turning Truck: The driver 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.	The driver should glance at the right occluding edge of the truck to detect any emerging hazards from obscured areas of the roadways.	
3. Obscured Crosswalk: There is a truck parked on the right side of a two-lane roadway right before a crosswalk. As the driver nears the truck and tries to pass from its left side, a vehicle approaches in the opposing lane.	The driver should scan the left side of the crosswalk now obscured by the approaching vehicle and also the left front edge of the truck on the right.	

Table 2.4. Descriptions and Plan Views for Scenarios 4-6 (Note: Driver is the red car)

Scenario Description	Required Action	Plan View
4. Pedestrian Island: The driver 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.	The driver should scan towards the front right edge of the first vehicle in the line of vehicles waiting to turn left to detect any obscured pedestrians who may be attempting to cross.	
5. Obscuring Vegetation: The driver 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.	At the intersection, the driver should continuously scan towards the obscured area on his or her right side while attempting to turn right in order to detect any potential hazards emerging from the obscured area.	
6. Stop Ahead: The driver is traveling on a road curving to the right and approaching a stop sign controlled intersection. 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.	The driver should glance at the Stop Ahead sign and then correctly identify the Stop Sign and stop at the intersection.	

2.4. Experimental Design

The experimental design was a 2 x 2 x 6 mixed with platform (fixed-based driving simulator or VR headset-based driving simulator) and age (young or middle-aged) as the two between-subject factors and scenario as the within-subject factor. A power analysis was performed to determine the sufficiency of the sample sizes (Cohen, 2013). A sample size of 12 young drivers and 12 middle-aged drivers, both assigned to drive on the fixed-based simulator, gave a statistical power equal to 93% with an alpha level of .05 and effect size of 0.6. The same sample size of young and middle-aged drivers on the VR headset-based driving simulator also yielded a statistical power equal to 93% with an alpha level of .05 and effect size of 0.6. The between-subject design for platform was chosen due to the fact that the scenarios were conceptually identical on both platforms and in a between-subject design, there would not be an instance where any learning effects experienced by participants after their first exposure to a specific scenario would transfer to their second exposure. Between-subject designs are valid, as long as the participants are assigned randomly to different conditions (Campbell & Stanley, 1963). The four groups of participants navigated six scenarios overall on their assigned platform. The order of the scenarios presented to participants was counterbalanced across and within groups using a balanced Latin Square method (Williams, 1949).

2.5. Simulator Sickness Questionnaire

The Simulator Sickness Questionnaire is the most widely used tool to measure simulator sickness (Stoner et al, 2011). In our experiment, we computed the total score calculated from the participants' responses on the SSQ for each of the age groups (Young and Middle-Aged) on both platforms along with the weighted nausea, oculomotor and disorientation scores.

2.6. 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). In our study, we computed the average score for each subscale based on each participants' responses for each of the age groups (Young and Middle-Aged) on both platforms.

2.7. Post-Study Questionnaire

The Post-Study Questionnaire (PSQ) was developed for this study to compare several user-experience-based attributes of the VR headset-based driving simulator and the fixed-based driving simulator. The average rating for each attribute was computed for each participant for each of the age groups (Young and Middle-aged) on both platforms along with the overall rating by each participant which is the average score of all the attributes' rating for each participant.

2.8. 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 given basic instructions such as to follow on-screen/audio instructions and maintain the posted speed limit. Eye-tracking 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. Both sets of participants then drove through a preliminary practice drive for the next five minutes. The purpose of this practice drive was to familiarize the participants with the virtual world and also the controls of the cab. The virtual world featured in the practice drive was a closed loop roadway consisting of several left/right turns, curves, intersections, and straight roads. While navigating through the practice drive, they were pointed out the rear and side view mirrors and were asked to brake, accelerate and make left/right turns. Once they concluded the practice drive, participants were permitted to continue to the experimental scenarios if they felt confident to drive and maneuver through the simulation. A set of six counterbalanced scenarios were then introduced to the participants with a gap of 30 seconds between loading each scenario. This session lasted for approximately 45 minutes. After concluding the driving session, a Simulator Sickness Questionnaire was administered to track any symptoms of Simulator Sickness (Kennedy et al, 1993). The Post-Study questionnaire was also filled out by the participants.

2.9. Analysis Techniques

The dependent variables considered for this experiment were binary scored latent hazard anticipation (whether the driver detected the latent hazard or not), glance duration (how long the driver scanned for the latent hazard), simulator sickness severity and user experience-based attributes of the simulator platform. To analyze these variables eye-tracking data was decoded from the recorded videos of each participants' drivers through each of the six scenarios and their responses on the simulator sickness questionnaire and post-study questionnaire were also analyzed. In addition to these variables, driver behavior questionnaire responses were also analyzed to wean out anomalies that may arise during a between-subject design experiment.

As mentioned earlier, in order to examine the drivers' latent hazard anticipation behavior (looked or did not look), the eye-tracking data from the recorded videos were binary scored (0 or 1). A set of 'launch zones' and 'target zones' were predetermined for each scenario based on previous studies (Muttart, 2013; 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 begin scan towards the target zone to successfully identify the presence of any potential threats. Participants who successfully glanced at the target zone while in the launch zone in a given scenario were scored '1' while those who failed to do so were scored '0'. The concept of 'Launch Zones' and 'Target Zones' is perhaps better understood with an example. Let us consider the scenario 'Obscured Crosswalk'. Figure 2.5. shows the launch zone and target zones for this scenario. The launch zone starts from a point which is 5 seconds before the crosswalk lying ~50 ft before the crosswalk. The target zones are the two obscured sides of the crosswalk, where potential threats can emerge. To be scored '1', the participant will need to scan both the target zones at least one time after entering the launch zone.

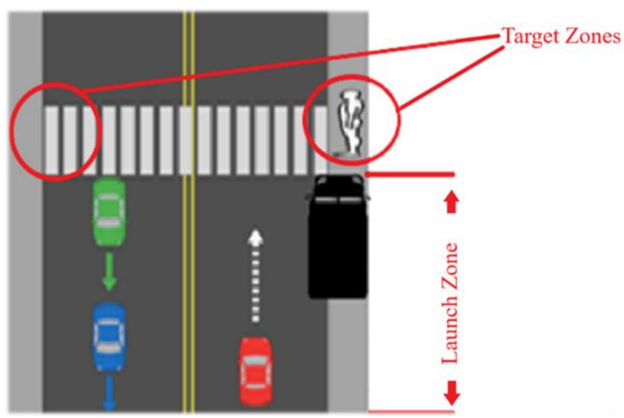


Figure 2.5. Launch Zone and Target Zones for the 'Obscured Crosswalk' scenario

Figure 2.6. and Figure 2.7. show the 'Obscured Crosswalk' scenario, from the driver's point-of-view on both platforms. The drivers in both instances have successfully identified the target zones in the scenario.

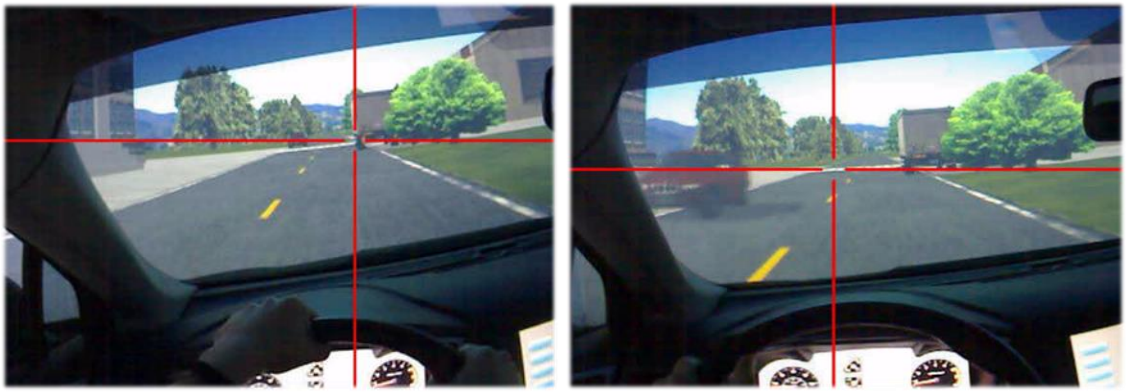


Figure 2.6. Successful Detection for the 'Obscured Crosswalk' scenario on the Fixed-based Driving Simulator

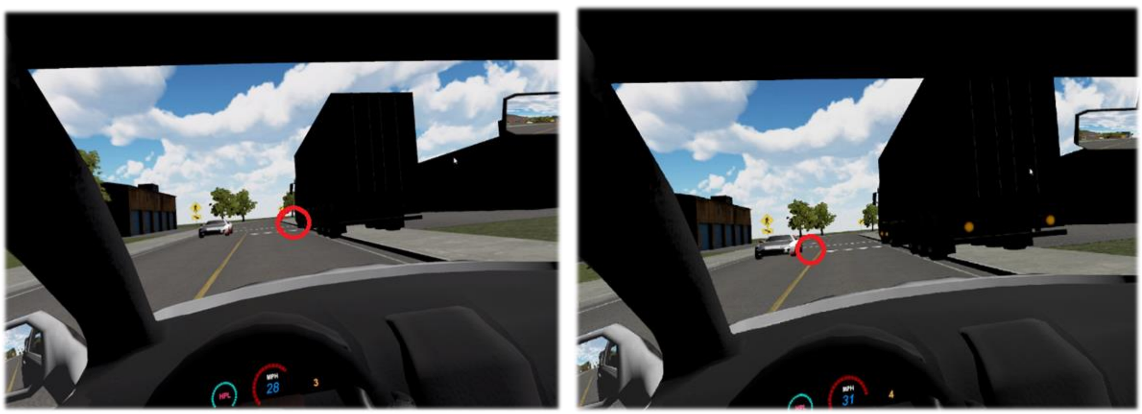


Figure 2.7. Successful Detection for the 'Obscured Crosswalk' scenario on the VR Headset-Based Driving Simulator

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.

For the scenarios where the participant successfully glanced at the target zone(s), the total and average glance duration were calculated. The process of calculating the glance duration is illustrated below with the help of figures.

1) At frame #3452, the participant upon entering the launch zone has not yet scanned the target zone. (Figure 2.8)

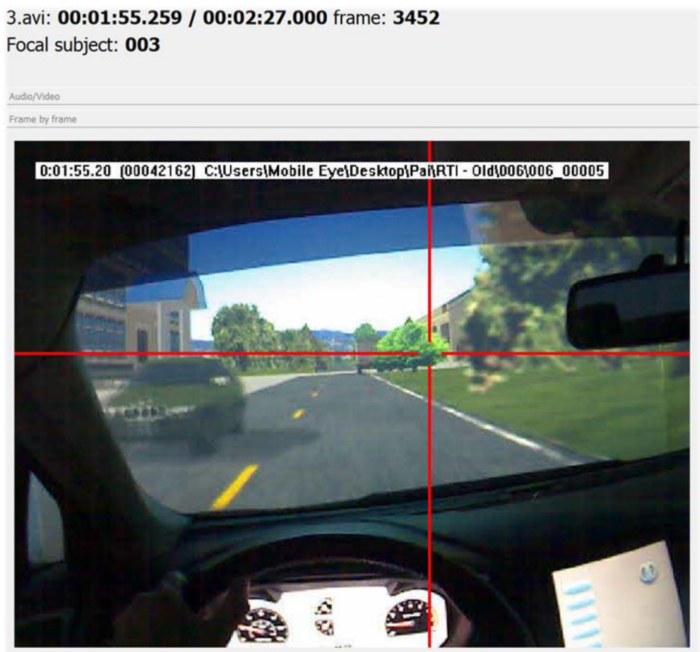


Figure 2.8. Participant is yet to scan the target zone(s)

2) At frame #3453, the participant scans the left side of the crosswalk which is one of the target zones and continues scanning that zone until frame #3474. (Figure 2.9.)

Figure 2.9. Participant begins scanning a target zone (Left); Participant stops scanning the target zone (Right)

Since each frame is 33 milliseconds each, the amount of the time the participant spent glancing at the target zone, i.e. glance duration is 3473 – 3453 = 20 * 33 = 660 milliseconds.

3) At frame #3477, the participant begins scanning the left edge of the truck on the right side of the crosswalk, which the other target zone in this scenario. He/she continues to do so until frame #3496. (Figure 2.10.)

Figure 2.10. Participant begins scanning another target zone (Left); Participant stops scanning the target zone (Right)

The amount of the time the participant spent glancing at the target zone, i.e. glance duration is 3496 – 3477 = 19 * 33 = 627 milliseconds. Considering these two glances at the target zones, the total glance duration would be the sum of the glance duration which is 1287 milliseconds or 1.28 seconds. The average glance duration would be 643.5 milliseconds or 0.64 seconds.

3. RESULTS

3.1. Latent Hazard Anticipation

In order to analyze the binary scored, binomially distributed eye-tracking data, a logistic regression model within the framework of Generalized Estimation Equations (GEE) was used. The model included age (younger and older) and the two platforms (VR headset and fixed-based driving simulator) as the between-subject factors, while scenario type was considered as a within-subject factor. The significance level was set at .05 and the participants were included as a random effect in the model. The model was used to determine whether there was a significant difference between the proportion of latent hazards detected by participants across two groups (young vs middle-aged) and two platforms (Fixed-based driving simulator vs VR headset-based driving simulator) as well as whether there was an interaction between scenario type and platform.

A backward elimination procedure was used to eliminate any non-significant higher order interactions. The final model revealed a highly significant main effect of age [Wald $\chi^2 = 28.72$; $p < 0.001$] which is consistent with the results from Pradhan et al. (2005) as well as our expected results. There was no significant effect of the platform [Wald $\chi^2 = 0.117$; $p > 0.05$]. The second order interaction between age and platform was not significant. There was a significant effect of scenario type [Wald $\chi^2 = 4871.61$; $p < 0.001$], but the second-order interaction between scenario type and platform was not significant.

For both platforms, the proportion of latent hazards detected was smaller for young driver groups when compared to their middle-aged driver counterparts on the same platform. On the fixed-based driving simulator, middle-aged drivers anticipated 92% of the latent hazards compared to only 64% for the young drivers. Similarly, on the virtual reality headset-based simulator, the middle-aged drivers anticipated 90% of the latent hazards compared to 62% for the young drivers (Figure 3.1.).

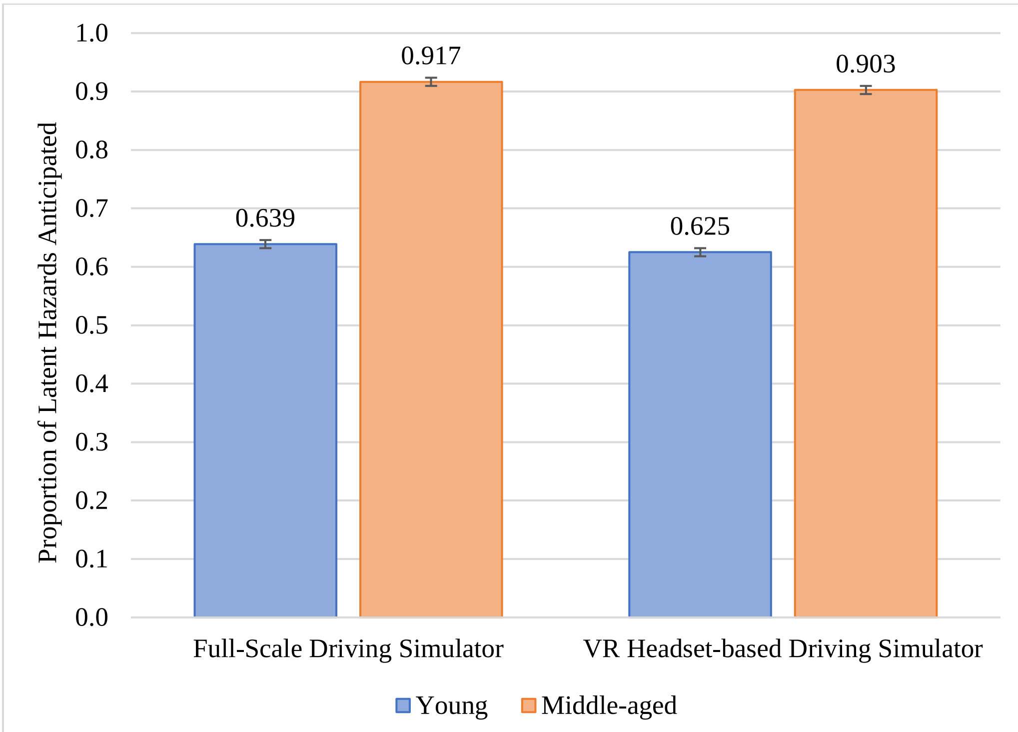


Figure 3.1. Proportion of Latent Hazards Anticipated by each group

3.2. Glance Duration

A 2 × 2 factorial [2 age 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, $\alpha = 0.05$.

3.2.1. Total Glance Duration

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 age and platform ($F = 2.733$; p-value = 0.1; $\eta^2 = 0.012$). There was a main effect of age ($F = 19.9$; p-value < 0.005; $\eta^2 = 0.084$).

3.2.2. Average Glance Duration

For average glance duration, there was no interaction between age 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 age ($F = 3.429$; p-value = 0.065; $\eta^2 = 0.015$).

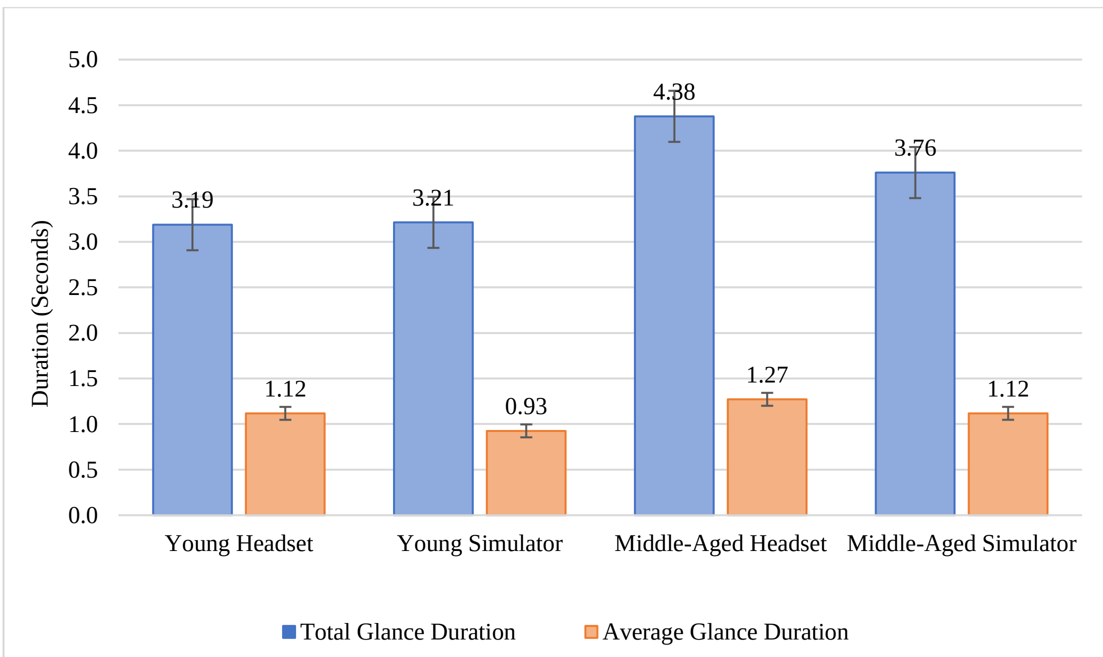


Figure 3.2. The mean average glance duration and mean total glance duration for each driver group

3.3. Simulator Sickness Questionnaire

Data from the simulator sickness questionnaire was collected and processed. While all drivers assigned to the fixed-based driving simulator groups completed their drives, two drivers assigned to the VR headset-based driving simulator (one Young, one Middle-aged) dropped out during or right after the preliminary practice drive and were immediately withdrawn from the study.

A 2 × 2 factorial [2 age groups: Young & Middle-Aged; 2 Platforms: VR headset-based simulator and fixed-based driving simulator] ANOVA was performed for the SSQ total scores as well as for the individual weighted scores for the three subscales (nausea, oculomotor and disorientation) for the non-dropout participants, $n = 48$, $\alpha = 0.05$.

3.3.1. Nausea

No interaction between age and platform ($F = 1.348$; p-value = 0.252; $\eta^2 = 0.030$) or main effect of platform ($F = 0.84$; p-value = 0.773; $\eta^2 = 0.002$) were observed, although there was a main effect of age ($F = 7.207$; p-value = 0.010; $\eta^2 = 0.141$), with middle-aged drivers scoring higher on the SSQ scaled score for Nausea than their younger counterparts.

3.3.2. Oculomotor

No interaction between age and platform ($F = 1.179$; p-value = 0.284; $\eta^2 = 0.026$) or main effect of platform ($F = 0.354$; p-value = 0.555; $\eta^2 = 0.008$) were observed, although there was a main effect of age ($F = 4.269$; p-value = 0.045; $\eta^2 = 0.088$), with middle-aged drivers scoring higher on the SSQ scaled score for Oculomotor than their younger counterparts.

3.3.3. Disorientation

No interaction between age and platform (F = 2.928; p-value = 0.094; $\eta^2 = 0.062$) or main effect of platform (F = 0.007; p-value = 0.932; $\eta^2 = 0.000$) were observed, although there was a main effect of age (F = 7.973; p-value = 0.007; $\eta^2 = 0.153$), with middle-aged drivers scoring higher on the SSQ scaled score for Disorientation than their younger counterparts.

3.3.4. Total Severity

No interaction between age and platform (F = 0.322; p-value = 0.573; $\eta^2 = 0.007$) or main effect of platform (F = 0.688; p-value = 0.411; $\eta^2 = 0.015$) were observed, although there was a main effect of age (F = 14.641; p-value = 0.0004; $\eta^2 = 0.25$), with middle-aged drivers scoring higher on the Total Severity score than their younger counterparts.

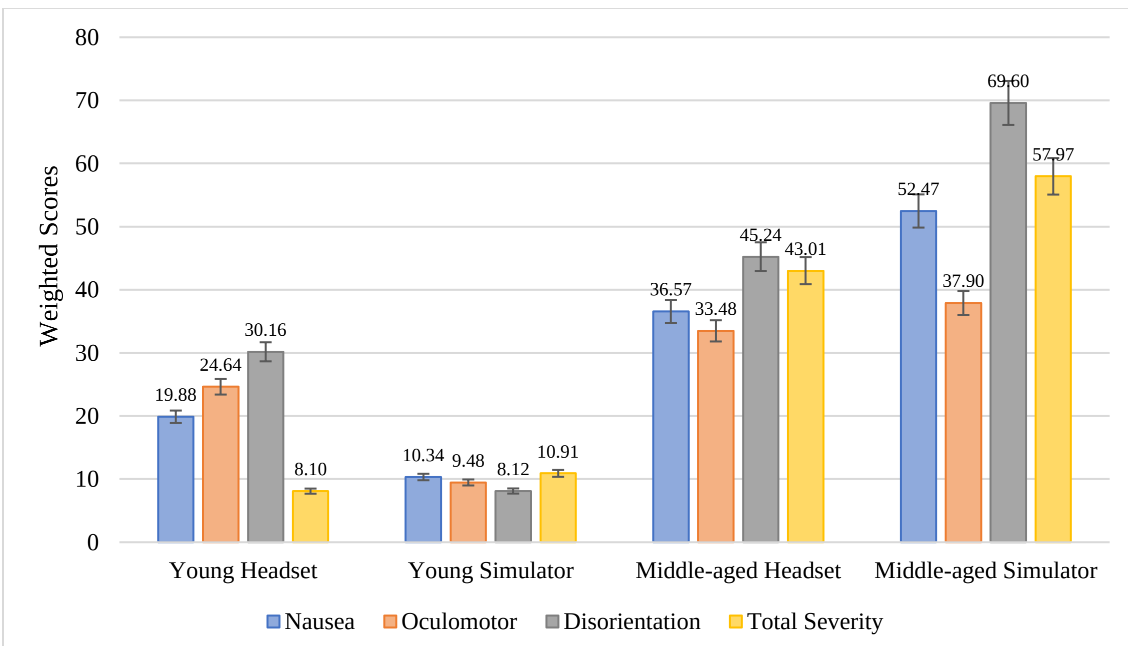


Figure 3.3. The weighted Simulator Sickness scores for each driver group

3.4. Driver Behavior Questionnaire

A 2 × 2 factorial [2 age 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, Lapse and Violation, $n = 48$, $\alpha = 0.05$.

3.4.1 Error

No interaction between age and platform ($F = 0.22$; p-value = 0.641; $\eta^2 = 0.005$) or main effect of platform ($F = 0.74$; p-value = 0.394; $\eta^2 = 0.017$) or of age ($F = 0.055$; p-value = 0.816; $\eta^2 = 0.001$) were observed.

3.4.2. Lapse

No interaction between age and platform ($F = 1.914$; p-value = 0.174; $\eta^2 = 0.042$) or main effect of platform ($F = 0.733$; p-value = 0.396; $\eta^2 = 0.016$) or of age ($F = 1.254$; p-value = 0.269; $\eta^2 = 0.028$) were observed.

3.4.3. Violation

No interaction between age and platform ($F = 0.561$; p-value = 0.458; $\eta^2 = 0.013$) or main effect of platform ($F = 0.773$; p-value = 0.384; $\eta^2 = 0.017$) or of age ($F = 0.027$; p-value = 0.871; $\eta^2 = 0.001$) were observed.

3.5. Post-Study Questionnaire

A 2 × 2 factorial [2 age groups: Young and Middle-aged; 2 Platforms: VR headset-based simulator and fixed-based driving simulator] ANOVA was performed for the average scores of each attribute to check for main effects or an interaction effect. Apart from Driving Controls ($F = 5.038$; p-value = 0.03; $\eta^2 = 0.103$), no other attribute had a significant main effect of age. Among all the attributes analyzed, only Navigation ($F = 6.856$; p-value $= 0.012$; $\eta^2 = 0.135$) and Driving Controls ($F = 36.52$; p-value < 0.005); $\eta^2 = 0.454$) had a significant main effect of Platform. There was an interaction effect between age and platform for Graphics ($F = 6.707$; p-value $= 0.013$; $\eta^2 = 0.132$), while no interaction effect between age and platform was found for any attributes. The mean scores for each of the attributes for all groups are listed below in Table 3.1. and illustrated in Figure 3.4.

Table 3.1. Mean Post-Study Questionnaire Scores for each group (Scaled 1 to 5)

	Young Headset	Young Simulator	Middle-aged Headset	Middle-aged Simulator
Navigation	3.75	4.25	3.33	4.08
Driving Controls	2.75	4.08	2.08	3.67
Graphical Quality	3.25	3.92	3.92	3.33
Sense of Realism	3.58	3.75	3.50	3.67
Audio Quality	4.00	4.00	3.67	3.83
Wearable Equipment	4.08	3.83	4.17	3.50
Seating Comfort	4.08	4.42	4.17	4.50
Overall Rating	3.64	4.04	3.55	3.80

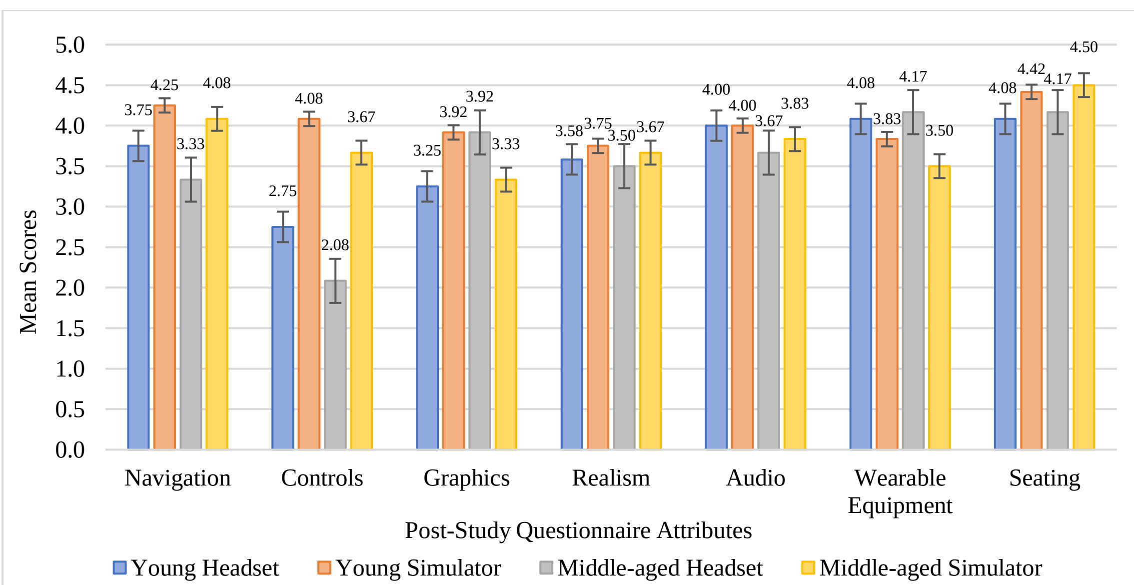


Figure 3. 4. Mean Scores for Post-Study Questionnaire attributes for each driver group

4. DISCUSSION

VR headsets are much less expensive than fixed-based driving simulators and therefore could greatly extend the power of simulation. Yet, even if valid as a way to measure something like latent hazard anticipation, they have produced documented evidence of simulator sickness (Munafo et al, 2017; Jensen & Konradsen, 2018). Thus, it is important to understand not only whether VR headsets are valid, but also whether they can be put to practical use. The current study sought to fill a gap in the literature by examining the validity of VR headsets at measuring driver performance (hazard anticipation ability) compared to a fixed-based driving simulator. While we could have chosen other metrics of performance to validate the platform, we chose to measure latent hazard anticipation ability both because it has been demonstrated to be significantly higher for middle-aged drivers compared to young drivers, on fixed-based driving simulators and on the open road (Lee et al, 2008; Romoser et al, 2013; Pradhan et al, 2005) and because it is linked to crashes.

4.1. Latent Hazard Anticipation

Consistent with our expected results, the results of the current study showed that the proportion of latent hazards anticipated by the middle-aged drivers was significantly more than that anticipated by young drivers on both the VR headset-based driving simulator (90% for middle-aged vs 62.5% for young – a difference of 27.8 percentage points) and the fixed-based driving simulator (91.7% for middle-aged vs 64% for young – a difference of 27.7 percentage points). This result was also in line with results from previous research conducted on driving simulators and in the field that demonstrated that middle-aged drivers anticipate a significantly greater proportion of latent hazards than young drivers (Pradhan et al, 2005). The result from the mixed-effect logistic regression model showed that there was no impact of platform on performance for either the young or the middle-aged drivers.

4.2. Glance Duration

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 (Urwyler et al, 2015; Crundall et al, 2012), there may only be marginal or no differences in terms of their average glance duration when compared to the younger drivers (Chapman & Underwood, 1998; Chan et al, 2010). 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.

4.3. Simulator Sickness

Driving simulator-based studies have always presented difficulties associated with high attrition rates due to simulator sickness or simulator adaptation syndrome for both young and old drivers (Helland et al, 2016). Virtual reality headsets have also been associated with such difficulties, with several studies reporting a high attrition rate among users due to motion sickness (Munafo et al, 2017; Jensen & Konradsen, 2018). Hence, the current study also examined the effect of simulator sickness on both platforms by comparing data collected from a standard Simulator Sickness Questionnaire (Kennedy et al, 1993). Consistent with this, two drivers assigned to the VR headset group dropped out of the study, but none in the fixed-based simulator group dropped out of the study. This drop-out rate is less than 10% and, for most studies, may not pose a serious limitation. Importantly, the weighted subscale scores and total simulator sickness scores among those who completed the experiment were compared between all driver groups on both platforms. The results indicated that there was no significant difference between simulator sickness scores on both platforms. There was a significant main effect of age on both platforms, with middle-aged drivers having significantly higher severity scores compared to young drivers. This is generally consistent with previous literature which states that older drivers are more prone to the symptoms of simulator sickness when compared to younger drivers (Brooks et al, 2010; Keshavarz et al, 2018). Furthermore, the lack of significance for any second-order interaction between age and platform indicated that the difference between the simulator sickness scores of middle-aged and young drivers was similar on both platforms.

4.4. Driver Behavior Questionnaire

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 age groups.

4.5. Post-study Questionnaire

The objective behind administering a Post-Study Questionnaire was to identify the various attributes we could improve the VR headset-based driving simulator based solely on a user experience standpoint. Analysis of the participants' responses on the questionnaire indicated that although several attributes are already on par with the fixed-based driving simulator, a few attributes such as 'Navigation' and 'Driving Controls' can be improved on the VR headset-based driving simulator, since the VR simulator received 15% and 38% lower rating on said attributes when compared to the fixed-based simulator. 'Driving Controls' were also perceived differently by the younger drivers and middle-aged drivers, where younger drivers were 12% more likely to rate the controls favorably than the middle-aged drivers.

4.6. Limitations and future work

The study has several important limitations as noted here. First, the current study used a between-design experiment to address the hypothesis that drivers would perform similarly on a VR-based driving simulator and fixed-based driving simulator. In these kinds of experiments, it is difficult to maintain complete homogeneity across the groups despite random assignment. It would be useful to consider a within-subject design with matching or block randomization techniques to eliminate confounds. In such a case, it would be worth looking into the possibility of integrating the VR headset to the controls of the fixed-based simulator in order to improve the comparison between the two platforms. Second, this study validated the virtual reality platform based only on the hazard anticipation skills of the young and middle-aged drivers. Future studies should also consider investigating other crash avoidance skills such as hazard mitigation and attention maintenance. Third, other measures of driving performance may also be considered for validation of a platform (e.g., various vehicle measures such as the standard deviation of lane position, or other eye movement measures such as horizontal and vertical gaze dispersion, physiological variables such as percentage of eye closure and blink rate or perhaps even workload metrics). Fourth, while the two platforms were found to differ in terms of dropout rates, there were no statistically significant differences in terms of severity of simulator sickness among those who completed the experiment. Evaluation of older drivers aged 65 years and above needs to be considered to measure true effectiveness. Fifth, the recruited population was imbalanced with regards to gender and the implications of this imbalance have not been explored. To further examine if gender had any effect on the hazard anticipation performance of the participants, gender was included in the logistic regression model along with two age groups and two platform groups. Results revealed that there was no significant effect of gender [Wald $\chi^2 = 0.150$; $p = 0.699$] on the latent hazard anticipation performance of the participants. Additionally, there was no second order interaction between age and gender [Wald $\chi^2 = 0.380$; $p = 0.537$] or between gender and platform [Wald $\chi^2 = 0.019$; p =

0.890]. A future study could focus on balancing the recruited population by gender and compare the performance between the two gender groups.

4.7. Conclusion

In summary, the current study showed that VR headsets may be used to effectively measure driver performance, specifically spatial characteristics of latent hazard anticipation behaviors and also the temporal characteristics. 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). VR headsets also appear, at least with hazard anticipation scenarios, not to generate more than minimal simulator sickness. VR headsets offer promise as an alternative to conventional simulators especially as a platform that can easily accommodate multiple users. The range of applications in which VR headset-based driving simulators could now be employed is greatly expanded. Multiple-vehicle conflicts involving multiple drivers or road users is one research theme that may be suitably addressed using VR headset-based simulators, 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.

APPENDIX A.

SIMULATOR SICKNESS QUESTIONNAIRE

Among the 16 (out of 29) symptoms highlighted on the Simulator Sickness Questionnaire, there were sets of symptoms that were correlated and three subscales were identified: Nausea (N), Oculomotor problems (O), and Disorientation (D). Each participant rated a symptom score of 0, 1, 2, or 3. For example, let's say a participant rates the seven symptoms under disorientation as, 2, 2, 3, 2, 3, 2, 1. The unweighted disorientation factor score will be 15 and the weighted disorientation score will be 15 × 7.58. Similarly, the weights for N and D are 9.54 and 13.92. The total score will be equal to the sum, N + O + D × 3.74.

#	Symptom	Severity Value 0=none, 1=slight 2=moderate, 3=severe
1	General discomfort	
2	Fatigue	
3	Boredom	
4	Drowsiness	
5	Headache	
6	Eye strain	
7	Difficulty focusing	
8a	Salivation increased	
8b	Salivation decreased	
9	Sweating	
10	Nausea	
11	Difficulty comcentrating	
12	Mental depression	
13	Fullness of head	
14	Blurred vision	
15a	Dizziness with eyes open	
15b	Dizziness with eyes closed	
16	Vertigo	
17	Visual flashbacks	
18	Faintness	
19	Aware of breathing	
20	Stomach awareness	
21	Loss of appetite	
22	Increased appetite	
23	Desire to move bowels	
24	Confusion	
25	Burping	
26	Vomiting	
27	Other	
	Subtotal	

APPENDIX B.

DRIVER BEHAVIOR QUESTIONNAIRE

Three subscales were identified for the 24 items listed in the Driver Behavior 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.

Statements	Rarely 0	1	2	3	4	Always 5
Try to pass another car that is signaling a left turn.						
Select the wrong turn lane when approaching an intersection.						
Fail to "Stop" or "Yield" at a sign, almost hitting a car that has right of way.						
Misread signs and miss your exit.						
Fail to notice pedestrians crossing when turning onto a side street.						
Drive very close to a car in front of you as a signal that they should go faster or get out of the way.						
Forget where you parked your car in a parking lot.						
When preparing to turn from a side road onto a main road, you pay too much attention to the traffic on the main road so that you nearly hit the car in front of you.						
When you backup, you hit something that you did not observe before but was there.						
Pass through an intersection even though you know that the traffic light has turned yellow and may go red.						
When making a turn, you almost hit a cyclist or pedestrian who has come up on your right side.						
Ignore speed limits late at night or very early in the morning.						
Forget that your lights are on high beam until another driver flashes his headlights at you.						
Fail to check your rear-view mirror before pulling out and changing lanes.						
Have a strong dislike of a particular type of driver, and indicate your dislike by any means that you can.						
Become impatient with a slow driver in the left lane and pass on the right.						
Underestimate the speed of an oncoming vehicle when passing.						
Switch on one thing, for example, the headlights, when you meant to switch on something else, for example, the windshield wipers.						
Brake too quickly on a slippery road, or turn your steering wheel in the wrong direction while skidding.						
You intend to drive to destination A, but you "wake up" to find yourself on the road to destination B, perhaps because B is your more usual destination.						
Drive even though you realize that your blood alcohol may be over the legal limit.						
Get involved in spontaneous, spur-of-the-moment, races with other drivers.						
Realize that you cannot clearly remember the road you were just driving on.						
You get angry at the behavior of another driver and you chase that driver so that you can give him/her a piece of your mind.						

APPENDIX C.

POST-STUDY QUESTIONNAIRE

The Post Study Questionnaire listed the following attributes of the simulator which were to be rated on a scale of 1 to 5 (Very Bad to Very Good): ‘Navigation’, ‘Driving Controls’, ‘Graphical Quality’, ‘Sense of Realism’, ‘Audio Quality’, ‘Wearable Equipment’, and ‘Seating Comfort’.

	Very Bad 1	**Bad 2**	**Neither Good nor Bad 3**	**Good 4**	**Very Good 5**
1: Navigating the Environment	1	2	3	4	5
2: Driving Controls	1	2	3	4	5
3: Graphics	1	2	3	4	5
4: Sense of Realism	1	2	3	4	5
5: Audio Quality	1	2	3	4	5
6: Wearable Equipment	1	2	3	4	5
7: Seating Comfort	1	2	3	4	5

BIBLIOGRAPHY

Agrawal, R., Knodler, M., Fisher, D. L., & Samuel, S. Virtual Reality Headset Training: Can It Be Used to Improve Young Drivers' Latent Hazard Anticipation and Mitigation Skills. *Transportation Research Record: Journal of the Transportation Research Board,* 2018. 0361198118758311.

Borowsky, A., & Oron-Gilad, T. (2013). Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. Accident Analysis & Prevention, 59, 548-565.

Brooks, J. O., Goodenough, R. R., Crisler, M. C., Klein, N. D., Alley, R. L., Koon, B. L., ... & Wills, R. F. (2010). Simulator sickness during driving simulation studies. *Accident Analysis & Prevention*, *42*(3), 788-796.

Burns, M. Immersive Learning for Teacher Professional Development. Retrieved from https://elearnmag.acm.org/featured.cfm?aid=2181208. Accessed July 21, 2018

Campbell, D., & Stanley, J. Experimental and quasi-experimental designs for research, 1963. Boston: Houghton Mifflin.

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.

Classen, S., Bewernitz, M., & Shechtman, O. (2011). Driving simulator sickness: an evidence-based review of the literature. *American journal of occupational therapy*, *65*(2), 179-188.

Cohen, J. (2013). Statistical power analysis for the behavioral sciences. Routledge.

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.

Darken, R. P., & Sibert, J. L. Wayfinding strategies and behaviors in large virtual worlds. In *Proceedings of the SIGCHI conference on Human factors in computing systems,* 1996. 142-149.

DeLuca, M., & Deluca, J. (2003). *U.S. Patent No. 6,559,813*. Washington, DC: U.S. Patent and Trademark Office.

Desai, P. R., Desai, P. N., Ajmera, K. D., & Mehta, K. (2014). A review paper on oculus rift-a virtual reality headset. arXiv preprint arXiv:1408.1173.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human factors*, 37(1), 65-84.

Fisher, D. L., Laurie, N. E., Glaser, R., Connerney, K., Pollatsek, A., Duffy, S. A., & Brock, J. The use of an advanced driving simulator to evaluate the effects of training and experience on drivers' behavior in risky traffic scenarios. *Human Factors*, 2002. 44(2): 287-302.

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.

Godley, S. T., Triggs, T. J., & Fildes, B. N. Driving simulator validation for speed research. *Accident analysis & prevention*, 2002. 34(5): 589-600.

Helland, A., Lydersen, S., Lervåg, L. E., Jenssen, G. D., Mørland, J., & Slørdal, L. Driving simulator sickness: Impact on driving performance, influence of blood alcohol concentration, and effect of repeated simulator exposures. *Accident Analysis & Prevention*, 2016. 94: 180-187.

Hilfert, T., & König, M. (2016). Low-cost virtual reality environment for engineering and construction. *Visualization in Engineering*, *4*(1), 2.

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.

Jensen, L., & Konradsen, F. A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies,* 2018. 23(4): 1515-1529.

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.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology,* 1993. 3(3): 203-220.

Keshavarz, B., Ramkhalawansingh, R., Haycock, B., Shahab, S., & Campos, J. L. Comparing simulator sickness in younger and older adults during simulated driving under different multisensory conditions. *Transportation research part F: traffic psychology and behaviour*, 2018. 54, 47-62.

Lee, J. D., Caven, B., Haake, S., & Brown, T. L. Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human factors*, 2001. 43(4): 631-640.

Lee, S. E., Klauer, S. G., Olsen, E. C. B., Simons-Morton, B. G., Dingus, T. A., Ramsey, D. J., Ouimet, M. C. Detection of road hazards by novice teen and experienced adult drivers. *Transportation Research Record*: *Journal of the Transportation Research Board*, 2008. 2078: 26-32.

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.

Marks, S., Estevez, J. E., & Connor, A. M. (2014, November). Towards the Holodeck: fully immersive virtual reality visualisation of scientific and engineering data. In *Proceedings of the 29th International Conference on Image and Vision Computing New Zealand* (pp. 42-47). ACM.

Mayhew, D. R. and Simpson, H. M. *The role of driving experience: Implications for training and licensing of new drivers*, Insurance Bureau of Canada, Toronto, Canada, 1995.
McComas, J., MacKay, M., & Pivik, J. (2002). Effectiveness of virtual reality for teaching pedestrian safety. *CyberPsychology & Behavior*, *5*(3), 185-190.

McDonald, C. C., Goodwin, A. H., Pradhan, A. K., Romoser, M. R., & Williams, A. F. A review of hazard anticipation training programs for young drivers. *Journal of Adolescent Health*, 2015. 57(1): 15-23.

McKnight, A. J., & McKnight, A. S. Young novice drivers: careless or clueless? *Accident Analysis & Prevention*, 2003. 35(6): 921-925.

Mehranian, H. (2013). Novice drivers: Development and evaluation of training program for hazard anticipation, hazard mitigation and attention maintenance skills in complex driving scenarios. University of Massachusetts Amherst.

Milleville-Pennel, I., & Charron, C. (2015). Do mental workload and presence experienced when driving a real car predispose drivers to simulator sickness? An exploratory study. *Accident Analysis & Prevention*, *74*, 192-202.

Morimoto, C. H., & Mimica, M. R. Eye gaze tracking techniques for interactive applications. *Computer vision and image understanding*, 2005. 98(1): 4-24.

Munafo, J., Diedrick, M., & Stoffregen, T. A. The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental brain research*, 2017. 235(3): 889-901.

Muttart, J. W. Identifying hazard mitigation behaviors that lead to differences in the crash risk between experienced and novice drivers (Doctoral dissertation, University of Massachusetts Amherst), 2013.

Nichols, S. (1999). Physical ergonomics of virtual environment use. *Applied Ergonomics*, *30*(1), 79-90.

Oagaz, H., Sable, A., Choi, M. H., Xu, W., & Lin, F. VRInsole: An unobtrusive and immersive mobility training system for stroke rehabilitation. In *Wearable and Implantable Body Sensor Networks (BSN), 2018 IEEE 15th International Conference,* 2018. pp. 5-8.

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.

Pulijala, Y., Ma, M., Pears, M., Peebles, D., & Ayoub, A. Effectiveness of immersive virtual reality in surgical training—a randomized control trial. *Journal of Oral and Maxillofacial Surgery*, 2018. 76(5): 1065-1072.

Realtime Technologies Catalog. Realtime Technologies Product Catalog. Retrieved from https://bit.ly/2RsFvfr. Accessed October 8, 2018.

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.

Reimer, B., D'Ambrosio, L. A., Gilbert, J., Coughlin, J. F., Biederman, J., Surman, C., ... & Aleardi, M. (2005). Behavior differences in drivers with attention deficit hyperactivity disorder: The driving behavior questionnaire. *Accident Analysis & Prevention*, *37*(6), 996-1004.

Roenker, D. L., Cissell, G. M., Ball, K. K., Wadley, V. G., & Edwards, J. D. Speed-of-processing and driving simulator training result in improved driving performance. *Human factors*, 2003. 45(2): 218-233.

Romoser, Matthew RE, Alexander Pollatsek, Donald L. Fisher, and Carrick C. Williams. Comparing the glance patterns of older versus younger experienced drivers: Scanning for hazards while approaching and entering the intersection. *Transportation research part F: traffic psychology and behavior*, 2013. 16: 104-116.

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.

Stoner, H., Fisher, D., & Mollenhaurer, M. J. Simulator and Scenario Factors Influencing Simulator Sickness. In *Handbook of Driving Simulation for Engineering, Medicine and Psychology*, 2011, pp. 14-1:14-23. Boca Rator: CRC.

Tan, C. T., Leong, T. W., Shen, S., Dubravs, C., & Si, C. (2015, October). Exploring gameplay experiences on the oculus rift. In Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (pp. 253-263). ACM.

Thomas, F. D., Rilea, S. L., Blomberg, R. D., Peck. R. C., & Korbelak, K. T. *Evaluation of the safety benefits of the risk awareness and perception training program for novice teen* *drivers*, Report No. DOT HS 812 235. Washington, DC: National Highway Traffic Safety Administration, 2016.

Tobii VR Integration. Eye tracking Specifications. Retrieved from https://www.tobiipro.com/product-listing/vr-integration/#Specifications. Accessed October 8, 2018.

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.

Velev, D., & Zlateva, P. Virtual reality challenges in education and training. *International Journal of Learning and Teaching,* 2017. 3(1): 33-37.

Vlakveld, W. Hazard anticipation of young novice drivers: Assessing and enhancing the capabilities of young novice drivers to anticipate latent hazards in road and traffic situations. Doctoral dissertation, University of Groningen, 2011.

Williams, E. J. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 1949. 2(2): 149-168.

Winn, W. Learning in virtual environments: A theoretical framework and considerations for design. *Education Media International*, 1999. 36(4): 271-279.