Isolating the effect of off-road glance duration on driving performance: An exemplar study comparing HDD and HUD in different driving scenarios

Missie Smith[ac], Kiran Bagalkotkar[a], Joseph L. Gabbard[a], David R. Large[b], Gary Burnett[b]

[a]Virginia Tech, Grado Department of Industrial & Systems Engineering, 250 Durham Hall (MC 0118), 1145 Perry Street, Blacksburg, VA 24061, USA

[b]Human Factors Research Group, Faculty of Engineering, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK

[c]Corresponding author: missie.smith@gmail.com

Précis: We employed a novel method to control visual attention in two driving environments and evaluated this using a head-down and a head-up display. We found evidence that when visual attention is similar, both display type and driving environment impact driving performance, which has important implications for current display assessment techniques.

Manuscript Type: Research Article

Word Count: 4,740

Acknowledgements: This material is based upon work supported by the National Science Foundation under Grant No. 1816721. The authors wish to thank Amanda Carroll who assisted in data collection and analysis.

Structured Abstract

Objective: We controlled participants' glance behavior while using head-down displays (HDD) and head-up displays (HUD) to isolate driving behavioral changes due to use of different display types across different driving environments. **Background:** Recently, HUD technology has been incorporated into vehicles, allowing drivers to, in theory, gather display information without moving their eyes away from the road. Previous studies comparing the impact of HUD to traditional displays on human performance show differences in both drivers' visual attention and driving performance. Yet no studies have isolated glance from driving behaviors which limits our ability to understand the cause of these differences and resulting impact on display design. **Method:** We developed a novel method to control visual attention in a driving simulator. Twenty experienced drivers sustained visual attention to in-vehicle HDDs and HUDs while driving in both a simple straight and empty roadway environment and a more realistic driving environment which included traffic and turns. **Results:** In the realistic environment, but not the simpler environment, we found evidence of differing driving behaviors between display conditions, even though participants' glance behavior was similar. **Conclusion:** Thus, the assumption that visual attention can be evaluated in the same way for different types of vehicle displays may be inaccurate. Differences between driving environments bring the validity of testing HUDs using simplistic driving environments into question. **Application:** As we move towards the integration of HUD user interfaces into vehicles, it is important that we develop new, sensitive assessment methods to ensure HUD interfaces are indeed safe for driving.

Keywords: Augmented reality, driver behavior, distraction, display assessment.

Introduction

Future in-vehicle displays may provide visual information to users by overlying graphics through the windshield and onto the surrounding environment, advancing potential capability of in-vehicle displays. These advanced HUD interfaces must be assessed for fitness for in-vehicle use to minimize risk to roadway users. Research has identified driver glances away from the road as problematic, and resulting guidelines (e.g. AAM, 2002; ISO, 2006; SAE, 2000) indicate that in-vehicle displays should encourage drivers to return glances back to the road (Metz et al., 2011). Thus, researchers often assess in-vehicle displays by focusing on glance behaviors, such as the duration or frequency of glance fixations on specific areas of the road or surrounding environment.

One established assessment method is Senders' visual occlusion method (Senders et al., 1967) which considers the central visual demands, but disregards information gained using peripheral vision (Burnett et al., 2013; Large & Burnett, 2015). While ignoring peripheral visual cues may be valid for HDD testing, but a key benefit of HUDs is drivers' ability to gather information using peripheral vision while using the display.

Another prevalent assessment method is the National Highway Transportation Safety Administration's (NHTSA's) Eye Glance in a Driving Simulator method (EGDS), in which display acceptability is determined by average display glance duration, percentage of time looking at the display, and total time with the eyes off road (NHTSA, 2012). While glance-based methods of assessing display safety have been validated for use with traditional in-vehicle head-down displays (HDDs), no such validation has taken place for use with novel displays like HUDs. This work explores the implications of applying current NHTSA assessment methods to emerging technologies such as HUDs. The study presented herein is an important step in determining whether two critical elements of common in-vehicle display assessment methods are suitable for HUD interface assessment: (1) glance durations towards the display, and, (2) the driving environment. In order to test these elements, we applied a novel method to systematically control glance duration and visual attention. We then examined the utility of a realistic driving environment as a replacement for national assessment standards, especially given the unique nature of HUD usage.

Visual Attention Toward In-Vehicle Displays

Analyzing visual attention is a fundamental part of understanding driving performance, especially when assessing in-vehicle visual displays (Cotter et al., 2008). Drivers must rapidly process and respond to dynamic visual information and increasingly complex in-vehicle displays contribute additional visual load. Even driving-related information displayed within the vehicle can be dangerous if focusing visual attention toward the display causes drivers to miss roadway hazards or signals. Advanced in-vehicle visual displays can be especially dangerous due to increased information quantity as information already present in the real world must be processed along with added virtual graphics in the case of HUDs, or as graphically rich HDDs provide detailed maps on increasingly large touch-screen displays. These visually rich displays may require more visual attention to process through the information, ultimately increasing the risk of driving accidents (NHTSA, 2010). The risk is especially present when the display requires or encourages sustained off-road visual attention that extends for more than two seconds (referred to in the literature as a “long glance”) (Klauer et al., 2006; NHTSA, 2012; Zwahlen et al., 1988). In this context, a “glance” is defined as an eye movement (saccade) to an area of interest (AOI) combined with all subsequent visual intakes (fixations) and saccades within that AOI (NHTSA, 2012), and may therefore extend for several seconds. A new glance begins when a saccade leaves one AOI (e.g. roadway) and moves into another (e.g. display). Previous findings are based on data collected using HDDs, before the widespread emergence of HUDs. Therefore, the impact of HUD interface design and usage on drivers' behavior and performance is not yet fully understood. Furthermore, researchers haven to yet determined how best to measure visual distraction and resulting safety associated with HUD interfaces.

HUDs allow drivers to receive information while still looking toward the road, maximizing the benefit of close spatial proximity, which is an important consideration for in-vehicle display design (Wittmann et al., 2006). It is possible that extended glances toward HUD graphics affect driving performance less than extended glances toward HDDs – most likely because drivers using HUDs may leverage peripheral vision for lane keeping and other basic visual tasks associated with driving (Horrey & Wickens, 2004a). As such, traditional methods of assessing visual attention might even characterize HUD glances as "on-road" since these glances are in the direction of the driving scene. Yet, peripheral vision alone is insufficient to safely drive because drivers must also attend and respond to roadway events (Horrey & Wickens, 2004a). In this case, glances toward HUDs could be considered "off-road" because drivers must verge and accommodate away from the road scene and onto the focal plane of the HUD; this is likely to result in both visual and cognitive distraction. A recent study suggests that even when HUD graphics are presented at the same focal depth as the real-world reference (e.g., a lead vehicle), there is a cognitive cost to switching between the graphic and real-world reference (Gabbard et al., 2018). Therefore, throughout this work, we consider glances to the graphics on the HUD to be "off-road" rather than on-road.

Indeed, changes in drivers' glance and driving behavior while using HUDs has been mixed (Donkor, 2012). Researchers have employed a variety of tasks reflecting potential use cases for

HUDs including visual search tasks (Smith et al., 2015, 2016, 2017), navigation tasks (Bolton et al., 2015; Liu & Wen, 2004), verbal response tasks (Horrey & Wickens, 2004b), and hazard identification/response (Horrey & Wickens, 2004c; Kim et al., 2013; Liu & Wen, 2004). Yet, none of these examples employed tasks that systematically demanded drivers' visual attention, such that eyes-off-road time, or glance duration, was managed within the study design. In studies where visual attention was analyzed, results frequently showed that participants distributed road and display glances differently when using HUDs as compared to HDDs (Bolton et al., 2015; Horrey & Wickens, 2004b; Smith et al., 2016, 2017). Because roadway glances and driving behavior are empirically linked, previous findings of differing driving behaviors may have been caused in part by changes in adopted glance behaviors. Additional research is needed to understand underlying causes of changes to driving performance and the implications of these changes for assessing new HUD interfaces for safe, on-road use.

Driving Environment

In driving simulator-based research, the driving *environment* includes the driving scene and roadway elements, which can affect research outcomes (Large et al., 2015; Teh et al., 2014). However, driving environment is not frequently the focus of experiments, as widely accepted standards have been adopted. For example, research examining the suitability of in-vehicle displays is often conducted under non-binding NHTSA guidelines, whereby participants follow a single lead car traveling at a constant 50mph on a straight, two-lane road with little or no other traffic (NHTSA, 2012). However, past research on traffic complexity and driving performance indicated driver workload increased with increased traffic flow, affecting speed control, headway, and lane keeping (Teh et al., 2014). Further, driving environment can impact glance behaviors, and the simple NHTSA-specified scenario may not elicit authentic driving behavior

(Large et al., 2015). Thus, while glance patterns while using HUDs and HDDs will likely change across different driving environments, it is unclear whether these changes maintain similar patterns. Because physiological indicators like glance allocation are used to predictor changes in workload (Ayaz et al., 2012) and, ultimately, driving behavior, researchers must understand and validate these glance-based assumptions for HUDs. If changes in glance and driving behavior while using HUDs differ from changes found while using HDDs, then there is further evidence for establishing new methods of assessment.

Hypotheses

The goal for this work was to explore how participants' driving behavior and vehicle control changes when glance duration varies while using different in-vehicle displays. A secondary goal was to examine the impact of driving environment when using these different displays. Therefore, we examined driving behaviors while participants used HDDs and HUDs to complete a visually demanding task in two different environments. We tested two hypotheses for this work:

- **H1.** As the duration of focused visual attention toward a display increases, driving performance deteriorates more quickly when using HDDs compared to HUDs.
- **H2.** Simple driving environments (e.g. NHTSA-prescribed) are less likely to reveal differences between display types than driving environments which include dynamic elements (e.g. curves and other vehicles).

Methods

The study took place at the University of Nottingham, UK, and was approved by the University's Faculty of Engineering Ethics Committee and the Institutional Review Board at Virginia Tech (#17-563); informed consent was obtained from each participant.

Participants

Five female and fifteen male experienced drivers (M = 6357.5 miles per year) with a valid driver's license for at least two years (M = 14.75 years) participated in the study. Participants were aged 18 – 65 years old (M = 33.95 years) and self-reported that they had normal or corrected-to-normal vision. No participants reported previous experience using windshield-based HUDs.

Driving Task

Participants completed a series of driving tasks using the car-following paradigm (Brookhuis et al., 1994; NHTSA, 2012) in our UK-based driving simulator, while complying with UK driving laws. The lead car remained in the left lane of the road throughout all drives but exhibited different driving behavior depending on the driving environment, described below.

Conventional Environment

Our *conventional* driving environment adhered to NHTSA guidelines specifying that the lead car travel at a constant speed of 50 mph on a straight, two lane road (NHTSA, 2012). The conventional environment included no traffic, turns, or other stimuli to divert visual attention away from the focused visual attention task. Participants initially drove for approximately 20-seconds, after which a lead car appeared on the road directly in front of participants' simulated car. Participants continued to drive, following the lead car at a safe distance, while completing secondary (focused visual attention) tasks. The conventional environment allowed drivers to anticipate and respond to the behavior of the lead car and the roadway.

Realistic Environment

In our *realistic* driving environment, participants followed a variable-speed lead car on a multi-lane road with slight curvature and additional traffic traveling in the same and opposite directions, with the UK national speed limit of 70mph, appropriate to this type of roadway (Large et al., 2015). The environment included varied speeds, additional road curvature, and increased volume of other cars to provide more realistic driving conditions. With the exception of intermittent lead car "comfort braking" (Large et al., 2018; Pampel et al., 2019), which occurred up to five times during a drive, the lead car drove at the same speed as participants, meaning that the lead car speed was variable and determined by the speed at which participants drove (but they did not know that this was occurring).

Focused Visual Attention Task

At the beginning of each drive, we verbally instructed participants to maintain safe control of the vehicle and follow the lead car at a safe driving distance (primary task) while completing *focused visual attention tasks* to control drivers' off-road glance behavior (secondary task). To complete these tasks, participants focused visual attention on the selected display and watched a single white letter changing every 0.1s until it randomly paused for 0.4s, at which point participants read aloud the paused letter. This method encouraged participants to maintain foveal attention directed to the display for a predetermined glance time. To successfully complete the task, participants could not look away from the stimuli until the task ended and the screen changed to a blank screen (HDD) or became fully transparent (HUD).

We selected durations of 1s, 2s, 5s, 10s, and 20s for the focused visual attention task. However, during pilot tests for this study, HDD glances longer than 5s resulted in crashes often enough that data loss became a concern. Thus, we excluded HDD focused visual attention task durations exceeding 5s to avoid crashes and resulting data loss. Three repetitions of each glance duration (HDD-1s, 2s, 5s; HUD-1s, 2s, 5s, 10s, 20s) were randomly ordered within each drive such that participants were unable to predict the length of the next task. We allocated short breaks between tasks so participants could refocus on driving. When a new task began, a car horn sound alerted participants to stimulus appearance, but participants did not know the duration. Participants wore eye-tracking glasses (ETG) to enable us to validate their visual behavior.

Equipment

We conducted the study in a medium-fidelity, fixed-base simulator in the Human Factors Research Group Lab at University of Nottingham (UK). The simulator included a 270-degree forward field of view curved projection with rear and side mirror displays. Participants drove in both environments in a right-hand drive Audi TT car. We fitted the Audi with a Pioneer CyberNavi HUD (780x260 pixels) with a focal depth of approximately 3 meters and with a Microsoft Surface Pro 4 Tablet model 1724 (HDD) (2736x1824 pixels) which was mounted using the suction cup mount seen in Figure 1. We displayed the focused visual attention task in white font on the displays using time embedded slides in PowerPoint, collecting participants' binocular gaze location and forward-facing view using SensoMotoric Instruments (SMI) eye-tracking glasses, sampled at 60Hz. We matched the visual angle for the tasks such that it was approximately 0.9 degrees, and text for both displays was larger than the suggested 0.25" for in-vehicle displays (Green et al., 1993).

Procedure

After participants consented, we seated them in the driving simulator and helped them adjust the seat to their preferred position, fitted the eye-tracking glasses, and calibrated the software. We then vertically and horizontally aligned letters projected on the HUD with boxes on the curved projection wall and confirmed that the position was correct through the eye-tracking video feed (Figure 1). The purpose of the calibration was to ensure that participants viewed the projected letters at the same location relative to the lead car in their field-of-view.

Figure 1. This eye tracking glasses image shows the calibration guide (blue box) used to properly align the HUD graphics display via Pioneer CyberNavi HUD in front of participants.

After calibration, participants undertook a practice drive in the simulator. We instructed participants to drive 70mph (the U.K. national speed limit) in the realistic environment and 50mph in the conventional environment (in line with NHTSA recommendations). Once participants were familiar with driving in the simulator, we verbally explained the focused visual attention task. Participants subsequently undertook a second practice drive while simultaneously doing the focused visual attention task. Participants then completed six drives (counterbalanced): three in realistic and three in conventional environments. Participants drove with no display (baseline), HUD, and HDD. During the baseline drive, participants drove for five minutes with no secondary task. Between drives, participants took a break, if desired. All participants were compensated with a £10 Amazon voucher.

Analysis

We analyzed participants' glance behavior using sematic gaze mapping with the data obtained from the ETG to validate our method and found no significant differences in average glance duration, glance duration frequency, and total glance time allocated to each AOI, i.e., the road, display (HUD or HDD), or other vehicle instruments (e.g. mirrors and speedometer). Therefore, the method elicited similar visual behavior and division of visual attention regardless of display type, something that has until now not been systematically demonstrated in HUD driving research.

To assess the effect of HUD and HDD on driving performance, we collected lateral and longitudinal vehicle control data. We calculated *lane position* (LP) according to SAE J2944 10.1.1.1 (Option A), meaning that the lateral position was determined relative to lane center (Green, 2013). *Standard deviation of lane position* (SDLP) was derived from lane position (Cotter et al., 2008). Because the lead car drove different speeds in the conventional and realistic driving environments, we used *minimum distance to collision* (MDC) to assess longitudinal vehicle control.

We analyzed three data sets: (1) 20s of data for HUD drives, denoted as HUD-20, (2) 5s of data for HDD drives, denoted as HDD-5, and (3) the first 5s of data from each of those datasets to compare HUD to HDD (Combined-5). Thus, we analyzed the longest focused visual attention duration for each display type individually and compared the first 5s across displays. To conduct our analysis, we subdivided each data set into sequential epochs of 1s duration. Since participants did not know the focused visual attention duration when they began each task, they could not predict how long they would need to look at the display. Therefore, we expected that corresponding epochs (e.g. the first second) for all glance tasks would have similar characteristics for a given display type, regardless of the total focused visual attention duration.

Results

For lateral and longitudinal data in the HUD-20 and HDD-5 data sets, we conducted a repeated-measures ANOVA with focused visual attention duration (5s or 20s), sequential time (1-5s or 1-20s), and driving environment (realistic or conventional) as our independent variables and included replication order effects in the model. In the Combined-5 data set, we conducted a repeated-measures ANOVA (as above) with display type (HUD or HDD) as an additional independent variable. We determined differences to be significant when $p<0.05$.

Lane Position

We found main effects of presentation order on lane position in all three datasets, with the third repetition resulting in a lane position closest to center than the second repetition, which was further to the right for all datasets (Table 1). There were no other main effects.

Table 1. ANOVA Results for Lane Position

ANOVA	*F*	*p*	Post hoc differences
Combined-5			
Display	0.002	0.964	
Environment	2.207	0.138	
Sequential Time	0.387	0.818	
Order	3.323	**0.036***	*3>2*

Display*Sequential Time	0.209	0.933	
Environment*Sequential Time	0.310	0.872	
Environment*Display	15.046	**0.000***	*Realistic-HDD>Conventional-HDD*
HUD-20			
Environment	0.011	0.917	
Sequential Time	0.477	0.972	
Order	5.169	**0.006***	*3>2*
Environment*Sequential Time	0.539	0.9464	
HDD-5			
Environment	2.087	0.149	
Sequential Time	0.369	0.831	
Order	3.744	**0.024***	*3>2*
Environment*Sequential Time	0.490	0.743	

*Note: Differences between levels found in post hoc testing is indicated by “level 1>level 2”, where the level with the larger mean is listed first.

There was an interaction effect of environment and display in the Combined-5 dataset. While all conditions resulted in lane positions slightly right of center, post hoc testing showed that when using HDDs, participants drove further to the right (more negative) in the conventional environment than the realistic environment (Figure 2).

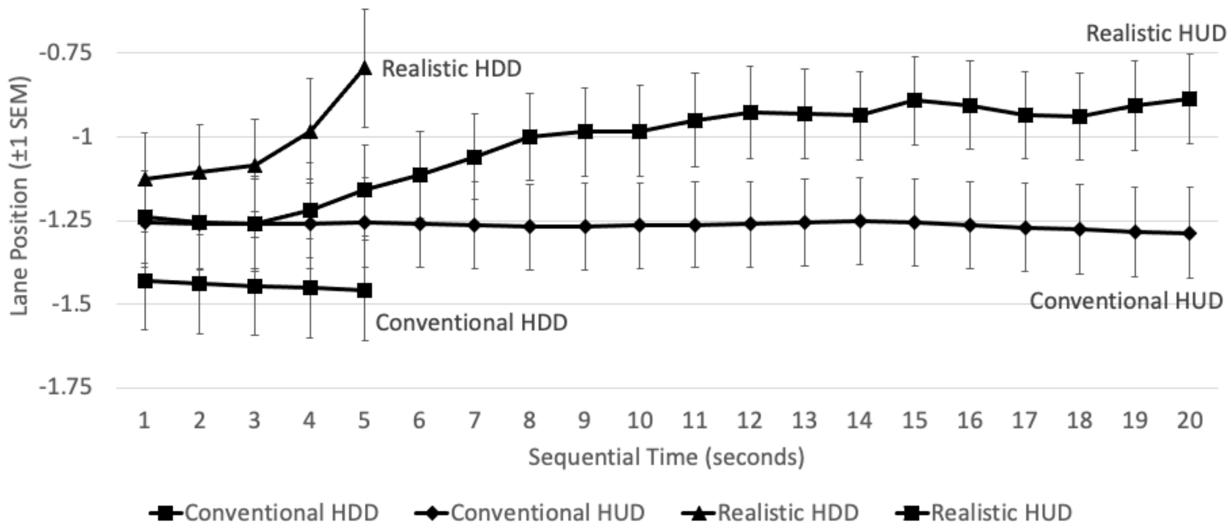

Figure 2. Mean lane position for each epoch is plotted for each display and environment combination. Increasingly negative values indicate driving further to the right.

Standard Deviation of Lane Position

In all three datasets, we found main effects of environment and sequential time on the SDLP (Table 2).

Table 2. ANOVA Results for Standard Deviation of Lane Position

Source	***F***	***p***	**Post hoc differences**
Combined-5			
Display	0.049	0.825	
Environment	142.048	**0.000***	*Realistic>Conventional*
Sequential Time	2.764	**0.026***	*5>1*
Order	2.870	0.057	
Display*Sequential Time	13.303	**0.000***	*HDD-5>(HUD-1|2|3|4|5, HDD-1/2/3)*
			HDD-4>(HUD-1|2|3|4|5, HDD-1|2)
			HDD-3>(HUD-4/5)
Environment*Sequential Time	0.356	0.840	

<table>
<tr><td>Environment*Display</td><td>60.326</td><td>0.000*</td><td>Realistic-HDD>(all others)
Realistic-HUD>(Conventional-HDD|HUD)
Conventional-HUD>Conventional-HDD</td></tr>
<tr><td>HUD-20</td><td></td><td></td><td></td></tr>
<tr><td>Environment</td><td>60.636</td><td>0.000*</td><td>Realistic>Conventional</td></tr>
<tr><td>Sequential Time</td><td>2.204</td><td>0.002*</td><td>2>17|15|8</td></tr>
<tr><td>Order</td><td>3.074</td><td>0.046*</td><td>2>3</td></tr>
<tr><td>Environment*Sequential Time</td><td>1.569</td><td>0.055</td><td></td></tr>
<tr><td>HDD-5</td><td></td><td></td><td></td></tr>
<tr><td>Environment</td><td>29.182</td><td>0.000*</td><td>Realistic>Conventional</td></tr>
<tr><td>Sequential Time</td><td>6.557</td><td>0.000*</td><td>5>1|2|3</td></tr>
<tr><td>Order</td><td>0.271</td><td>0.763</td><td></td></tr>
<tr><td>Environment*Sequential Time</td><td>5.126</td><td>0.001*</td><td>Realistic-(1,2,3,4,5)>Conventional-(1,2,3,4,5),
Realistic-5>Realistic-(1,2,3)
Realistic-4>Realistic-(1,2)</td></tr>
</table>

Post hoc testing and Figure 3 show that the realistic environment resulted in higher SDLP than the conventional environment for all three datasets. In the Combined-5 and HDD-5 datasets, the fifth epoch was associated with higher SDLP than the first epochs, and in the HDD-5 dataset, the fifth epoch was also associated with higher SDLP than the second and third epochs. In the HUD-20 dataset, the second epoch was associated with higher SDLP than several other epochs (8s, 5s, 17s).

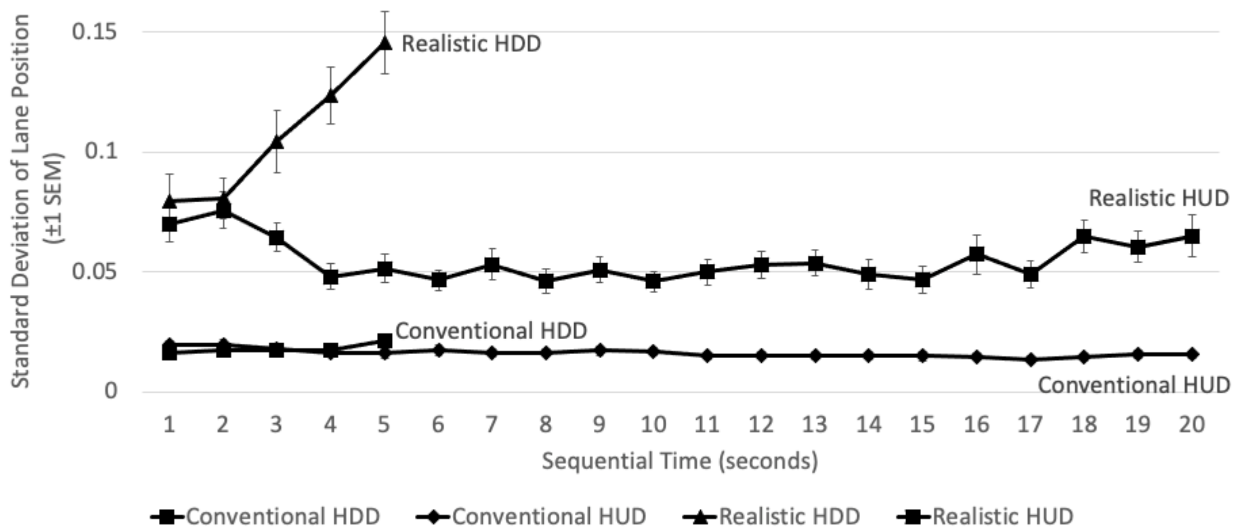

Figure 3. Mean Standard Deviation of Lane Position (SDLP) for display type (HUD and HDD) and environment (realistic and conventional) combination.

There was also an interaction effect of display and sequential time, with the fourth and fifth HDD epochs associated with higher SDLP than all five HUD epochs and the first two HDD epochs. The third HDD epoch was associated with higher SDLP than the fourth and fifth HUD epochs. Thus, there was a pattern of participants' SDLP increasing with passing time when using the HDD, however, these effects were not present with the HUD.

Minimum Distance to Collision

For all three data sets, the conventional environment resulted in longer MDC than the realistic (Table 3, Figure 4). In the Combined-5 data set, HDD use resulted in longer MDC than HUD use.

Table 3. ANOVA Results for Minimum Distance to Collision

ANOVA	*F*	*p*	Post hoc differences
Combined-5			

Display	1.535	0.216	
Environment	64.978	**0.000***	*Conventional>Realistic*
Sequential Time	0.004	1.000	
Order	1.773	0.170	
Display*Sequential Time	0.011	0.999	
Environment*Sequential Time	0.007	0.999	
Environment*Display	0.052	0.820	
HUD-20			
Environment	32.807	**0.000***	*Conventional>Realistic*
Sequential Time	0.456	0.979	
Order	12.836	**0.000***	*3>1|2*
Environment*Sequential Time	0.434	0.984	
HDD-5			
Environment	33.279	**0.000***	*Conventional>Realistic*
Sequential Time	0.008	0.999	
Order	12.473	**0.001***	*2>3|1, 1>3*
Environment*Sequential Time	0.023	0.999	

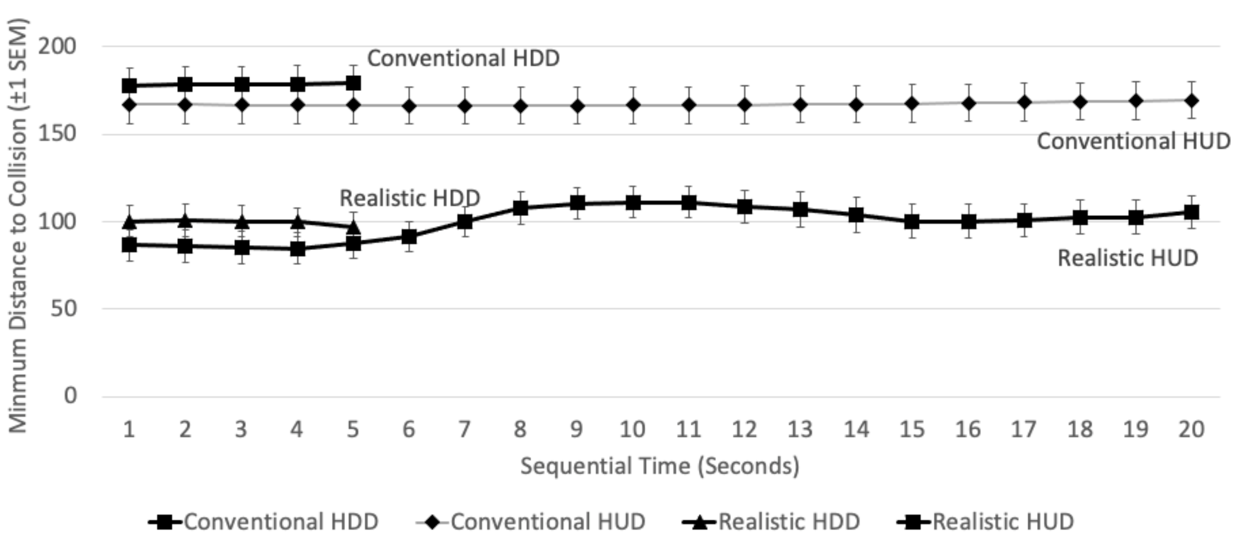

Figure 4. Minimum distance to collision for display type (HUD and HDD) and environment (realistic and conventional) combination.

Discussion

The purpose of this study was to examine two assumptions underpinning current glance-based display assessments: (1) glance duration can be used to predict driving behavior, and, (2) HUDs and HDDs affect drivers similarly across different driving environments. To achieve this, we systematically controlled focused visual attention towards displays and examined the impact of more realistic driving environments on drivers performing visually demanding tasks. In general, we found that both display type and driving environment affected participants' driving behavior when visual attention was controlled.

Durations

As we systematically controlled participants' focused visual attention duration, we expected to find quicker and more significant driving performance deterioration associated with HDDs compared to HUDs (**H1**). We found no significant differences in lane position, but HDD use was associated with increasing SDLP over time which was higher than when using HUDs. The trend in the HDD data suggests SDLP may increase until intervention occurs (e.g. looking back to the road). When controlling for visual attention duration toward HUDs, participants showed no marked increase in SDLP over the first sequential epoch at any time. Conversely, participants using the HDD showed increased SDLP as the task duration increased, especially after 2s. In particular, the third, fourth, and fifth seconds driving while using HDDs were all associated with higher SDLP (i.e. degraded lateral vehicle control) than the same epochs with HUD use. These findings provide evidence of changes in both lateral and longitudinal vehicle control measures between displays, with HDD use resulting in more rapid and diminished driving performance than HUD use.

The Combined-5 data showed participants using HDDs allowed more distance between their car and the lead vehicle compared to HUD use, which is indicative of more conservative driving (Brookhuis et al., 1994). Because this finding was true across both environments, it suggests that participants were less comfortable extending glances toward HDDs, and is evidence of deteriorated driving performance relative to HUD use.

These differences in lateral and longitudinal vehicle control support **H1**, suggesting that drivers may sustain longer visual attention toward HUDs without experiencing as much deterioration in driving performance. There are many potential causes for the vehicle control differences between display types, including increased use of peripheral vision when using HUDs and prior exposure to HDDs. However, two theories provide possible explanations for systematic changes in lateral vehicle control. First, Senders (1967) posits that time looking away from the road, and in this case toward HDDs, results in increased *uncertainty* which impacts drivers' behavior (Senders et al., 1967). As participants maintained glances toward the displays, their visual uncertainty about the state of the road may have increased more rapidly during HDD tasks because participants could not leverage their peripheral vision as they could when using the HUD. As uncertainty increased, drivers may have been less aware of their lane position resulting in over- or under-compensation for changes in lane position, ultimately impacting their SDLP. A second theory concerns *gaze concentration*. Specifically, situations in which drivers primarily focus on one point in the road (their gaze concentration) can result in decreased lateral lane position variation (Li et al., 2018), supporting the decreased SDLP evident with HUD use. While both theories provide plausible explanations, they may have vastly different implications for drivers. Senders' theory would support HUD use in vehicles because degraded lateral vehicle control was lower due to lower uncertainty when participants used HUDs. However, if HUD use indeed causes increased gaze concentration and cognitive tunneling, HUDs may negatively impact drivers' ability to respond to roadway events – as seen when using AR applications in other domains (Kerr et al., 2011). While these two theories may result in conflicting recommendations for which display is *safer*, it is important to note that both theories may be evident in this study. It is possible that HUDs can simultaneously introduce benefits to drivers while also causing new problems. Therefore, further work is required to more explicitly test these theories and to determine design implications.

Driving Environment

Characteristics of the driving environment can impact driving performance (Horrey & Wickens, 2004b; Senders et al., 1967), yet some assessment methods, such as EGDS (NHTSA, 2012) specify one type of driving environment. We therefore examined the impact of realistic and conventional driving environments on driver performance and hypothesized that we would find more rapid driving performance decrements in the more realistic environment (**H2**).

We found no significant main effect of driving environment on lane position, but the realistic environment resulted in a different lane position than the conventional environment during HDD use. The road geometry slightly differed between environments (3 lanes in the realistic and 2 lanes in the conventional), which may have influenced participants' perception of space and the resulting position they adopted. Nevertheless, the absolute difference between positions were small (less than one foot), so the real-world implications are likely minimal.

In all three data sets, the realistic environment resulted in higher SDLP (lateral instability) than the conventional environment, suggesting the realistic environment was more challenging to drive. Additionally, we found interaction effects of sequential time and display in the HDD-5 dataset only. Specifically, HDD use in the realistic environment was associated with higher SDLP than in the conventional environment for all epochs. Moreover, later epochs in the realistic environment were associated with higher SDLP than early epochs, showing an increase in SDLP over time – this was only present when participants used HDDs. This supports **H2** in part, because driving performance deteriorated more quickly in our realistic driving environment than in the conventional, but only when using the HDD. Thus, participants' ability to maintain lateral vehicle control differed between the two displays.

Assessment Methods

Because many in-vehicle display assessments are based on glance behaviors (e.g. NHTSA, 2012), extended visual attention towards HUDs might be assumed to have a similar negative impact on driving behavior as extended glances towards HDDs. Yet, currently accepted assessment techniques were developed using data collected from HDDs. While participants in our study drove similarly when using both displays in the conventional environment, they exhibited different driving behaviors in the realistic environment. Specifically, when visual attention was controlled, participants' driving behaviors changed differently, depending on the display. In other words, not all driving behavior differences between HUDs and HDDs in prior research can be attributed to differences in participants' *selected* glance behaviors. This is important because while the NHTSA EGDS method is commonly used to assess HDDs, it only includes one type of driving environment that is not representative of all, or arguably any, real-world scenario. Because HUDs and HDDs impact users differently in different driving environments, we cannot assume that results from a simple environment will generalize to real-world driving. Assessing glance behavior in simple environments, like our conventional environment, may under-emphasize potential benefits of HUD use, namely, drivers' ability to more effectively use their ambient or peripheral vision, as evidenced by the driving performance measures. In other words, even when the duration of focused visual attention was the same, driving performance differed between HUD and HDD. Thus, assessing HUDs based on extant assumptions about glance and driving behavior developed with HDDs may be inadequate. Instead, we must develop new methods that are valid for each display type.

Long Glances

Prior research into drivers' glance behavior indicates that there is a two-second threshold for glances away from the road, above which the likelihood of crash increases significantly (Klauer et al., 2006). In our study, the HDD was most in-keeping with these findings – we found degradation in SDLP for HDD after two seconds of focused visual attention. Thus, our study suggests that one contributor to increased crash risk at two seconds could be the result of increased lateral instability. However, we did not find similar degradation in SDLP when using the HUD, which may suggest that HUDs are a safer alternative to HDDs because they permit glances without hindering lateral vehicle control. It might also mean that drivers using HUDs are able to maintain lateral control for longer than the widely accepted two-seconds, and new "safety" thresholds could be established for HUDs. While it is not possible to determine a new threshold from our results, it appears that visual attention focused on HUDs could potentially extend beyond 20 seconds in some situations.

Limitations

While the findings are compelling, this driving simulator study included a relatively small sample size (n=20). Future work should be done to validate these findings with more participants as well as on-road studies.

Conclusions and Future Work

This work has uniquely contributed to driving-related research by providing a systematic method to control "off-road" visual attention duration ("off road glances"). Applying this, we found that driving performance differed between HUD and HDD usage even when visual attention did not. Further, simplistic driving environments commonly used in research failed to reveal any differences between display type, whereas a more realistic driving environment uncovered nuanced differences in vehicle control. Thus, measures implying that driving performance can be determined based on glance pattern alone in simple environments are likely flawed. As such, common methods like the NHTSA EGDS test may provide poor recommendations when assessing HUDs. Because of this, we must pursue other methods of assessing driver behavior and performance to ensure safe on-road interactions. Assessing HUDs in visually rich environments may be required to provide realistic feedback on drivers' potential performance while using this type of display. Further, standard recommendations, such as the widely accepted two-second rule, should be evaluated for HUDs in future work to help designers quickly assess potential dangers of using these displays.

Key Points

- Visual attention has been closely linked with driving behavior and is commonly used to assess in-vehicle visual displays.

- Augmented reality head-up display (HUD) usage is associated with different glance and driving behaviors than traditional in-vehicle displays.
- Even when glance behavior is controlled, HUD use may result in different driving behaviors relative to traditional (head-down) displays.
- Different types of driving environments affect driver behaviors differently when using HUDs and traditional in-vehicle displays.

References

Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., & Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage*, *59*(1), 36–47.

Bolton, A., Burnett, G., & Large, D. R. (2015). An investigation of augmented reality presentations of landmark-based navigation using a head-up display. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15*, 56–63. https://doi.org/10.1145/2799250.2799253

Brookhuis, K., Waard, D. de, & Mulder, B. E. N. (1994). Measuring driving performance by car-following in traffic. *Ergonomics*, *37*(3), 427–434.

Burnett, G., Neila, N., Crundall, E., Large, D. R., Lawson, G., Skrypchuk, L., & Thompson, S. (2013). How do you assess the distraction of in-vehicle information systems? A comparison of occlusion, lane change task and medium-fidelity driving simulator methods. *Proceedings of DDI2013*.

Cotter, S., Stevens, A., Popken, A., & Gelau, C. (2008). Development of innovative methodologies to evaluate ITS safety and usability: HUMANIST TF E. *Proceedings of European Conference on Human Centred Design for Intelligent Transport Systems*, 55.

Donkor, G. E. (2012). Evaluating the impact of Head-Up Display complexity on peripheral detection performance: a driving simulator study. *Advances in Transportation Studies*, *28*.

Gabbard, J., Mehra, D. G., & Swan II, J. E. (2018). Effects of AR display context switching and focal distance switching on human performance. *IEEE Transactions on Visualization and Computer Graphics*.

Green, P. (2013). Standard definitions for driving measures and statistics: overview and status of recommended practice J2944. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 184–191.

Green, P., Levison, W., Paelke, G., & Serafin, C. (1993). *Suggested human factors design guidelines for driver information systems*. Citeseer.

Horrey, W. J., & Wickens, C. D. (2004a). Driving and side task performance: The effects of display clutter, separation, and modality. *Human Factors*, *46*(4), 611–624.

Horrey, W. J., & Wickens, C. D. (2004b). Driving and Side Task Performance: The Effects of Display Clutter, Separation, and Modality. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(4), 611–624. https://doi.org/10.1518/hfes.46.4.611.56805

Horrey, W. J., & Wickens, C. D. (2004c). Focal and ambient visual contributions and driver visual scanning in lane keeping and hazard detection. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*(19), 2325–2329.

Kerr, S. J., Rice, M. D., Teo, Y., Wan, M., Cheong, Y. L., Ng, J., Ng-Thamrin, L., Thura-Myo, T., & Wren, D. (2011). Wearable mobile augmented reality: evaluating outdoor user experience. *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*, 209–216.

Kim, H., Wu, X., Gabbard, J. L., & Polys, N. F. (2013). Exploring head-up augmented reality interfaces for crash warning systems. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13*, 224–227. https://doi.org/10.1145/2516540.2516566

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., Ramsey, D. J., & others. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data*.

Large, D. R., & Burnett, G. E. (2015). An overview of occlusion versus driving simulation for assessing the visual demands of in-vehicle user-interfaces. *4th International Conference on Driver Distraction and Inattention*, *November*, 1–12.

Large, D. R., Pampel, S. M., Burnett, G. E., Thompson, S., Skrypchuk, L., & others. (2018). *Exploring Drivers' Visual Behaviour During Take-Over Requests*.

Large, D. R., van Loon, E., Burnett, G., & Pournami, S. (2015). Applying NHTSA task acceptance criteria to different simulated driving scenarios. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive*

Vehicular Applications, 117–124.

Li, P., Markkula, G., Li, Y., & Merat, N. (2018). Is improved lane keeping during cognitive load caused by increased physical arousal or gaze concentration toward the road center? *Accident Analysis & Prevention*, *117*, 65–74.

Liu, Y.-C., & Wen, M.-H. (2004). Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. *International Journal of Human-Computer Studies*, *61*(5), 679–697.

Metz, B., Schömig, N., & Krüger, H.-P. (2011). Attention during visual secondary tasks in driving: Adaptation to the demands of the driving task. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*(5), 369–380.

NHTSA. (2010). Fatality analysis reporting system (FARS) encyclopedia. *Online:[Http://Wwwfars. Nhtsa. Dot. Gov/Main/Index. Aspx], Accessed:[July 12, 2014]*.

NHTSA. (2012). Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. *Washington, DC: National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT)*.

Pampel, S. M., Large, D. R., Burnett, G., Matthias, R., Thompson, S., & Skrypchuk, L. (2019). Getting the driver back into the loop: the quality of manual vehicle control following long and short non-critical transfer-of-control requests: TI: NS. *Theoretical Issues in Ergonomics Science*, *20*(3), 265–283.

Senders, J. W., Kristofferson, A. B., Levison, W. H., Dietrich, C. W., & Ward, J. L. (1967). *The attentional demand of automobile driving*. Bolt, Beranek and Newman Incorporated.

Smith, M., Gabbard, J. L., Burnett, G., & Doutcheva, N. (2017). The effects of augmented reality head-up displays on drivers' eye scan patterns, performance, and perceptions. *International Journal of Mobile Human Computer Interaction*, *9*(2). https://doi.org/10.4018/IJMHCI.2017040101

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

Smith, M., Streeter, J., Burnett, G., & Gabbard, J. L. (2015). Visual search tasks: the effects of head-up displays on driving and task performance. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15*, 80–87. https://doi.org/10.1145/2799250.2799291

Teh, E., Jamson, S., Carsten, O., & Jamson, H. (2014). Temporal fluctuations in driving demand: The effect of traffic complexity on subjective measures of workload and driving performance. *Transportation Research Part F: Traffic Psychology and Behaviour*, *22*, 207–217.

Wittmann, M., Kiss, M., Gugg, P., Steffen, A., Fink, M., Pöppel, E., & Kamiya, H. (2006). Effects of display position of a visual in-vehicle task on simulated driving. *Applied Ergonomics*, *37*(2), 187–199. https://doi.org/10.1016/j.apergo.2005.06.002

Zwahlen, H. T., Adams, C. C., & DeBals, D. P. (1988). Safety Aspects of CRT Touch Panel Controls in Automobiles. *Vision in Vehicles--II : Proceedings of the Second International Conference on Vision in Vehicles*, 335–344. https://trid.trb.org/view/927092

Biographies

Missie Smith earned her BS and MS from Mississippi State University in 2010 and 2012, respectively. She earned her Ph.D. in Industrial and Systems Engineering in 2018 from Virginia Tech. Dr. Smith researches the impact of technology on users' perception, performance, and behaviors.

Kiran Bagalkotkar is a Junior Software/Systems with Planned Systems International, Inc. supporting the U.S. Army Night Vision and Electronic Sensors Directorate (NVESD). She received her M.Eng. in Industrial and Systems Engineering with a certificate in Human Computer Interaction in 2020 from Virginia Tech.

Joseph L. Gabbard is an Associate Professor of Human Factors in the Grado Department of Industrial and Systems Engineering at Virginia Tech and Director of the Cognitive Engineering for Novel Technology lab. He received his Ph.D. in Computer Science from Virginia Tech in 2008.

David R. Large is a Senior Research Fellow with the Human Factors Research Group at the University of Nottingham. He holds a PhD in Human Factors (2013) from the University of Nottingham.

Gary Burnett is a professor of Transport Human Factors at the University of Nottingham. He received his Ph.D. from Loughborough University in 1998.